Theoretical Physics Reference

Release 0.5

Ondřej Čertík

Dec 11, 2019

CONTENTS

CHAPTER

ONE

INTRODUCTION

1.1 Preface

I have a very bad memory. I am able to memorize quite a lot of things short term, but I am not able to remember most formulas from quantum mechanics over the long term (e.g. like over the summer). I don't remember formulas for perturbation theory (neither time dependent or time independent), I don't remember Feynman rules in quantum field theory, I don't even remember the Dirac equation exactly (where the i should be, if there is m or m^2, ...). The thing about quantum field theory is not that some particular steps are difficult, but that there are so many of them and one has to master all of them at once, in order to really "get it".

I never got QFT, because once I mastered one part sufficiently, I forgot some other part and it took so much time to master that other part that I forgot the first part again. However, I was determined that I would get it. In order to do so, I realized I need to keep notes of things I understood, written in my own way. Then, when I relearn some parts that I forgot, it just takes me a few minutes to go over my reference notes to get into it quickly. My own style of understanding is that the notes should be complete (no need to consult external books), yet very short and getting directly to the point, and also with every single calculation carried out explicitly.

See also the preface to the QFT part.

If you want to study physics, learn math the physics way (as opposed to the usual mathematics way of a definition, theorem, proof, ...). When I was beginning my undergrad physics studies (and even on a high school), I also had this common misconception, that I need to study math and understand every proof and then I'll be somehow prepared for physics. I was very wrong. I used to study calculus by myself and then trying to learn the proofs, and Lebesgue integral and I was learning that from the mathematics books. At the university, I always did all my math exams first (as far as I remember, I always got A from those), hoping that would be a good start for the physics exams, but I always found out that it was mostly useless.

Now I know that the only way to study physics is to go and do physics directly and learn the math on the way as needed. The math section of this book reviews all the math, that is necessary for studying theoretical physics (graduate level).

There are actually quite a lot of good math books written by physicists as well as many excellent physics books, covering everything that I cover here. But I really like to have all the theoretical physics and the corresponding math explained in one book, and to keep it as short as possible. Also everyone has a bit different style and amount of rigor and I have not found a book that would perfectly suite my own style, thus I wrote one.

1.2 Introduction

The Theoretical Physics Reference is an attempt to derive all theoretical physics equations (that are ever needed for applications) from the general and special relativity and the standard model of particle physics.

The goals are:

- All calculations are very explicit, with no intermediate steps left out.

- Start from the most general (and correct) physical theories (general relativity or standard model) and derive the specialized equations from them (e.g. the Schrödinger equation).
- Math is developed in the math section (not in the physics section).
- Theory should be presented as short and as explicitly as possible. Then there should be arbitrary number of examples, to show how the theory is used.
- There should be just one notation used throughout the book.
- It should serve as a reference to any physics equation (exact derivation where it comes from) and the reader should be able to understand how things work from this book, and be ready to understand specialized literature.

This is a work in progress and some chapters don't conform to the above goals yet. Usually first some derivation is written, as we understood it, then the mathematical tools are extracted and put into the math section, and the rest is fit where it belongs. Sometimes we don't understand some parts yet, then those are currently left there as they are.

There are many excellent books about theoretical physics, that one can consult about particular details. The goal of this book (when completed) is to show where things come from and serve as a reference to any particular field, so that one doesn't get lost when reading specialized literature.

Here is an incomplete list of some of the best books in theoretical physics (we only picked those that we actually read):

1. Landau, L. D.; Lifshitz, E. M: Course of Theoretical Physics
2. Richard Feynman: The Feynman Lectures on Physics
3. Walter Greiner: "Classical Theoretical Physics" series of texts
4. Herbert Goldstein: Classical Mechanics
5. J.D. Jackson: Classical Electrodynamics
6. Charles W. Misner, Kip S. Thorne, John Wheeler: Gravitation
7. Bernard Schutz: A First Course in General Relativity
8. Carrol S.: The Lecture Notes on General Relativity
9. J.J. Sakurai: Advanced Quantum Mechanics
10. Brown L. S.: Quantum Field Theory
11. Mark Srednicki: Quantum Field Theory
12. Claude Itzykson, Jean-Bernard Zuber: Quantum Field Theory
13. Zee A.: Quantum Field Theory in a Nutshell
14. Steven Weinberg: The Quantum Theory of Fields
15. L.H. Ryder: Quantum Field Theory
16. Jiří Hořejší: Fundamentals of Electroweak Theory
17. Michele Maggiore: A Modern Introduction to Quantum Field Theory
18. M.E. Peskin & D.V. Schroeder: An Introduction to Quantum Field Theory
19. J.W. Negele, H. Orland: Quantum Many-Particle Systems
20. X-G. Wen: Quantum Field Theory of Many-Body Systems
21. Dirac, P.A.M.: General Theory of Relativity

CHAPTER

TWO

CONTRIBUTORS

A list of people who contributed patches or text for the book in the order of the date of contribution:

- Ondřej Čertík
- Karel Výborný
- Jean Jordaan
- Lei Ma
- 14. Sukumar

CHAPTER

THREE

MATHEMATICS

3.1 Integration

This chapter doesn't assume any knowledge about differential geometry. The most versatile way to do integration over manifolds is explained in the differential geometry section.

3.1.1 General Case

We want to integrate a function f over a k-manifold in $\mathbf{R}^n$, parametrized as:

$$\varphi : \mathbf{R}^k \to \mathbf{R}^n \quad \varphi(t_1, t_2, \ldots, t_k) = \begin{pmatrix} \varphi_1(t_1, t_2, \ldots, t_k) \\ \varphi_2(t_1, t_2, \ldots, t_k) \\ \vdots \\ \varphi_n(t_1, t_2, \ldots, t_k) \end{pmatrix}$$

then the integral of $f(x_1, x_2, \ldots, x_n)$ over φ is:

$$\int_M f(x_1, x_2, \ldots, x_n)\,\mathrm{d}S = \int_{\mathbf{R}^n} f(\varphi(t_1, t_2, \ldots, t_k))\sqrt{\det \mathbf{G}}\,\mathrm{d}t_1\mathrm{d}t_2\cdots\mathrm{d}t_k$$

where $\mathbf{G}$ is called a Gram matrix and $\mathbf{J}$ is a Jacobian:

$$(\mathbf{G})_{ij} = (\mathbf{J}^T\mathbf{J})_{ij} = J_{ik}J_{kj} = \frac{\partial\varphi_k}{\partial t_i}\frac{\partial\varphi_k}{\partial t_j}$$

$$(\mathbf{J})_{ij} = \frac{\partial\varphi_i}{\partial t_j} = \begin{pmatrix} \frac{\partial\varphi}{\partial t_1} & \frac{\partial\varphi}{\partial t_2} & \cdots & \frac{\partial\varphi}{\partial t_k} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

The idea behind this comes from the fact that the volume of the k-dimensional parallelepiped spanned by the vectors

$$\frac{\partial\varphi}{\partial t_1}, \ldots, \frac{\partial\varphi}{\partial t_k}$$

is given by

$$V = \sqrt{\det \mathbf{J}^T\mathbf{J}}$$

where $\mathbf{J}$ is an $n \times k$ matrix having those vectors as its column vectors.

Example

Let's integrate a function $f(x, y, z)$ over the surface of a sphere in 3D (e.g. $k = 2$ and $n = 3$):

$$\begin{aligned}
\varphi(\theta, \phi) &= \begin{pmatrix} r\sin\theta\cos\phi \\ r\sin\theta\sin\phi \\ r\cos\theta \end{pmatrix} \\
\mathbf{J} &= \begin{pmatrix} -r\sin\theta\sin\phi & r\cos\theta\cos\phi \\ r\sin\theta\cos\phi & r\cos\theta\sin\phi \\ 0 & -r\sin\theta \end{pmatrix} \\
\mathbf{G} = \mathbf{J}^T\mathbf{J} &= \begin{pmatrix} -r\sin\theta\sin\phi & r\sin\theta\cos\phi & 0 \\ r\cos\theta\cos\phi & r\cos\theta\sin\phi & -r\sin\theta \end{pmatrix} \begin{pmatrix} -r\sin\theta\sin\phi & r\cos\theta\cos\phi \\ r\sin\theta\cos\phi & r\cos\theta\sin\phi \\ 0 & -r\sin\theta \end{pmatrix} = \begin{pmatrix} r^2\sin^2\theta & 0 \\ 0 & r^2 \end{pmatrix} \\
\det\mathbf{G} &= r^4\sin^2\theta \\
\sqrt{\det\mathbf{G}} &= r^2\sin\theta \\
\int_M f(x,y,z)\mathrm{d}S &= \int_{\mathbf{R}^n} f(r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta)\, r^2\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi = \\
&= \int_0^\pi \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi\, f(r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta)\, r^2\sin\theta
\end{aligned}$$

Let's say we want to calculate the surface area of a sphere, so we set $f(x, y, z) = 1$ and get:

$$\int_M \mathrm{d}S = \int_0^\pi \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi\, r^2\sin\theta = 2\pi r^2 \int_0^\pi \mathrm{d}\theta\sin\theta = 4\pi r^2$$

3.1.2 Special Cases

k = n

$$\begin{aligned}
\det\mathbf{G} = \det\mathbf{J}^R\mathbf{J} &= (\det\mathbf{J})^2 \\
\mathrm{d}S &= |\det\mathbf{J}|\,\mathrm{d}t_1\,\mathrm{d}t_2\cdots\mathrm{d}t_k
\end{aligned}$$

k = 1

$$\begin{aligned}
\det\mathbf{G} = \det\left(\left(\frac{\mathrm{d}\varphi_1}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}\varphi_2}{\mathrm{d}t}\right)^2 + \cdots\right) &= \left|\frac{\mathrm{d}\varphi}{\mathrm{d}t}\right|^2 \\
\mathrm{d}S &= \left|\frac{\mathrm{d}\varphi}{\mathrm{d}t}\right|\mathrm{d}t
\end{aligned}$$

k = n - 1

$$\det \mathbf{G} = \det \mathbf{J}^R \mathbf{J} =$$
$$= \det(\cdots)^2 + \det(\cdots)^2 + \cdots + \det(\dots)^2 =$$
$$= \left| \det \begin{pmatrix} \frac{\partial \varphi}{\partial t_1} & \frac{\partial \varphi}{\partial t_2} & \cdots & \frac{\partial \varphi}{\partial t_k} & \mathbf{e}_1 \\ \vdots & \vdots & \vdots & \vdots & \mathbf{e}_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \mathbf{e}_n \end{pmatrix} \right|^2 \equiv |\omega_\varphi|^2$$
$$\mathrm{d}S = |\omega_\varphi| \, \mathrm{d}t_1 \, \mathrm{d}t_2 \cdots \mathrm{d}t_k$$

ω_φ is a generalization of a vector cross product. The $\det(\cdots)$ symbol means a determinant of a matrix with one row removed (first term in the sum has first row removed, second term has second row removed, etc.).

k = 2, n = 3

$$\det \mathbf{G} = \left| \frac{\partial \varphi}{\partial t_1} \times \frac{\partial \varphi}{\partial t_2} \right|^2$$
$$\mathrm{d}S = \left| \frac{\partial \varphi}{\partial t_1} \times \frac{\partial \varphi}{\partial t_2} \right| \mathrm{d}t_1 \, \mathrm{d}t_2$$

y = f(x, z)

$$\det \mathbf{G} = 1 + \left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial z} \right)^2$$
$$\mathrm{d}S = \sqrt{1 + \left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial z} \right)^2} \, \mathrm{d}x \, \mathrm{d}z$$

in general for $x_j = f(x_1, x_2, \dots, x_n)$ we get:

$$\det \mathbf{G} = 1 + \left(\frac{\partial f}{\partial x_1} \right)^2 + \left(\frac{\partial f}{\partial x_2} \right)^2 + \cdots$$
$$\mathrm{d}S = \sqrt{1 + \left(\frac{\partial f}{\partial x_1} \right)^2 + \left(\frac{\partial f}{\partial x_2} \right)^2 + \cdots} \, \mathrm{d}x_1 \, \mathrm{d}x_2 \cdots \mathrm{d}x_n$$

The "x_j" term is missing in the sums above.

Implicit Surface

For a surface given implicitly by

$$F(x_1, x_2, ..., x_n) = 0$$

we get:

$$\mathrm{d}S = |\nabla F| \left| \frac{\partial F}{\partial x_n} \right| \mathrm{d}x_1 \cdots \mathrm{d}x_{n-1}$$

Orthogonal Coordinates

If the coordinate vectors are orthogonal to each other:

$$\frac{\partial\varphi}{\partial t_i}\cdot\frac{\partial\varphi}{\partial t_i}=0 \quad \text{for } i\neq j$$

we get:

$$\mathrm{d}S=\left|\frac{\partial\varphi}{\partial t_1}\right|\left|\frac{\partial\varphi}{\partial t_2}\right|\cdots\left|\frac{\partial\varphi}{\partial t_k}\right|\mathrm{d}t_1\cdots\mathrm{d}t_k$$

3.1.3 Motivation

Let the k-dimensional parallelepiped P be spanned by the vectors

$$\frac{\partial\varphi}{t_1},\ldots,\frac{\partial\varphi}{t_k}$$

and let $\mathbf{J}$ is $n\times k$ matrix having these vectors as its column vectors. Then the area of P is

$$V=\sqrt{\det\mathbf{J}^T\mathbf{J}}$$

so the definition of the integral over a manifold is just approximating the surface by infinitesimal parallelepipeds and integrating over them.

3.1.4 Example

Let's calculate the total distance traveled by a body in 1D, whose position is given by $s(t)$:

$$\begin{aligned}l=\int_\gamma \mathrm{d}s=\int_{t_1}^{t_2}\left|\frac{\mathrm{d}s}{\mathrm{d}t}\right|\mathrm{d}t=\\ =\int_{t_1}^{t'}\left|\frac{\mathrm{d}s}{\mathrm{d}t}\right|\mathrm{d}t+\int_{t'}^{t''}\left|\frac{\mathrm{d}s}{\mathrm{d}t}\right|\mathrm{d}t+\cdots+\int_{t''''\cdots}^{t_2}\left|\frac{\mathrm{d}s}{\mathrm{d}t}\right|\mathrm{d}t=\\ =|s(t')-s(t_1)|+|s(t'')-s(t')|+\cdots+|s(t_2)-s(t''''\cdots)|\end{aligned}$$

where t', t'', ... are all the points at which $\left|\frac{\mathrm{d}s}{\mathrm{d}t}\right|=0$, so each of the integrals in the above sum has either positive or negative integrand.

3.2 Complex Numbers

We start by defining $\arg(z)$ by its principal value, then everything else follows from this definition. We could have also used any other branch, but then most results in this chapter would need to be updated with the new convention.

Then we define exponential, logarithm, power and so on using simple natural formulas. From these definitions, everything else follows using a very simple algebra manipulation, all the "messy" features are hidden in the definition and properties of the real $\mathrm{atan}\,2$ function. In the derivation of each formula, only formulas introduced before (above) are used.

Every formula in this chapter holds for all complex numbers, unless explicitly specified otherwise.

3.2.1 Real and Imaginary Part

A complex number z can be written using its real and imaginary parts:

$$z = \operatorname{Re} z + i \operatorname{Im} z$$

The absolute value $|z|$ is defined as:

$$|z| = \sqrt{\operatorname{Re}^2 z + \operatorname{Im}^2 z}$$

3.2.2 Argument Function

Principal value of $\arg(z)$ is defined as

$$\arg z = \operatorname{atan} 2(\operatorname{Im} z, \operatorname{Re} z)$$

Thus we have $-\pi < \arg z \leq \pi$. All operations with $\arg z$ are then derived using the properties of the real $\operatorname{atan} 2$ function.

3.2.3 Exponential

Exponential is defined using:

$$e^z = e^{\operatorname{Re} z + i \operatorname{Im} z} = e^{\operatorname{Re} z}(\cos \operatorname{Im} z + i \sin \operatorname{Im} z)$$

It follows:

$$\begin{aligned} e^{a+b} &= e^{\operatorname{Re}(a+b)}(\cos \operatorname{Im}(a+b) + i \sin \operatorname{Im}(a+b)) = \\ &= e^{\operatorname{Re} a} e^{\operatorname{Re} b} (\cos(\operatorname{Im} a)\cos(\operatorname{Im} b) - \sin(\operatorname{Im} a)\sin(\operatorname{Im} b) + i \sin(\operatorname{Im} a)\cos(\operatorname{Im} b) + i \cos(\operatorname{Im} a)\sin(\operatorname{Im} b)) = \\ &= e^{\operatorname{Re} a}(\cos \operatorname{Im} a + i \sin \operatorname{Im} a) e^{\operatorname{Re} b}(\cos \operatorname{Im} b + i \sin \operatorname{Im} b) = \\ &= e^a e^b \end{aligned}$$

Any complex number can be written in a polar form as follows:

$$\begin{aligned} z = \operatorname{Re} z + i \operatorname{Im} z &= |z| \left(\frac{\operatorname{Re} z}{|z|} + i \frac{\operatorname{Im} z}{|z|} \right) = \\ &= |z| \left(\frac{\operatorname{Re} z}{\sqrt{\operatorname{Re}^2 z + \operatorname{Im}^2 z}} + i \frac{\operatorname{Im} z}{\sqrt{\operatorname{Re}^2 z + \operatorname{Im}^2 z}} \right) = \\ &= |z| \left(\cos \operatorname{atan} 2(\operatorname{Im} z, \operatorname{Re} z) + i \sin \operatorname{atan} 2(\operatorname{Im} z, \operatorname{Re} z) \right) = \\ &= |z| \left(\cos \arg z + i \sin \arg z \right) = \\ &= |z| e^{i \arg z} \end{aligned}$$

The following formula holds:

$$\begin{aligned} \arg e^z = \arg e^{\operatorname{Re} z} e^{i \operatorname{Im} z} &= \arg e^{i \operatorname{Im} z} = \\ &= \arg(\cos \operatorname{Im} z + i \sin \operatorname{Im} z) = \\ &= \operatorname{atan} 2(\sin \operatorname{Im} z, \cos \operatorname{Im} z) = \\ &= \operatorname{Im} z + 2\pi \left\lfloor \frac{\pi - \operatorname{Im} z}{2\pi} \right\rfloor \end{aligned}$$

and also:

$$\begin{aligned}
\arg ab = \arg(|a|e^{i\arg a}|b|e^{i\arg b}) = \\
= \arg(|a||b|e^{i(\arg a+\arg b)}) = \\
= \arg(e^{i(\arg a+\arg b)}) = \\
= \arg(\cos(\arg a+\arg b)+i\sin(\arg a+\arg b)) = \\
= \operatorname{atan2}(\sin(\arg a+\arg b),\cos(\arg a+\arg b)) = \\
= \arg a+\arg b+2\pi\left\lfloor\frac{\pi-\arg a-\arg b}{2\pi}\right\rfloor
\end{aligned}$$

and

$$\arg\frac{1}{z} = -\arg z+2\pi\left\lfloor\frac{\pi+\arg z}{2\pi}\right\rfloor$$

and

$$\begin{aligned}
\arg\frac{a}{b} = \arg\left(|a|e^{i\arg a}\left|\frac{1}{b}\right|e^{i\arg\frac{1}{b}}\right) = \\
= \arg\left(|a|\left|\frac{1}{b}\right|e^{i(\arg a-\arg b)+2\pi i\left\lfloor\frac{\pi+\arg b}{2\pi}\right\rfloor}\right) = \\
= \arg(e^{i(\arg a-\arg b)}) = \\
= \arg(\cos(\arg a-\arg b)+i\sin(\arg a-\arg b)) = \\
= \operatorname{atan2}(\sin(\arg a-\arg b),\cos(\arg a-\arg b)) = \\
= \arg a-\arg b+2\pi\left\lfloor\frac{\pi-\arg a+\arg b}{2\pi}\right\rfloor
\end{aligned}$$

3.2.4 Logarithm

The logarithm is defined as:

$$\log z = \log|z|+i\arg z \tag{3.2.4.1}$$

The motivation is from the following formula:

$$z = |z|e^{i\arg z} = e^{\log|z|}e^{i\arg z} = e^{\log|z|+i\arg z}$$

which using our definition becomes:

$$z = e^{\log|z|+i\arg z} = e^{\log z} \tag{3.2.4.2}$$

so a logarithm is an inverse function to an exponential. The formula (3.2.4.2) would be satisfied even if we add a factor of $2\pi in$ (where n is an integer) to the right hand side of (3.2.4.1). However, the convention is to define logarithm using the equation (3.2.4.1) exactly.

We can now derive a few important formulas:

$$\log|e^z| = \log|e^{\operatorname{Re}z}e^{i\operatorname{Im}z}| = \log|e^{\operatorname{Re}z}| = \operatorname{Re}z$$

$$\log e^z = \log|e^z|+i\arg e^z = \operatorname{Re}z+i\left(\operatorname{Im}z+2\pi\left\lfloor\frac{\pi-\operatorname{Im}z}{2\pi}\right\rfloor\right) = z+2\pi i\left\lfloor\frac{\pi-\operatorname{Im}z}{2\pi}\right\rfloor$$

and

$$\begin{aligned}\log ab = \log|ab| + i\arg ab = \\ = \log|a| + \log|b| + i\arg a + i\arg b + 2\pi i\left\lfloor\frac{\pi - \arg a - \arg b}{2\pi}\right\rfloor = \\ = \log a + \log b + 2\pi i\left\lfloor\frac{\pi - \arg a - \arg b}{2\pi}\right\rfloor\end{aligned} \tag{3.2.4.3}$$

and

$$\begin{aligned}\log\frac{a}{b} = \log\left|\frac{a}{b}\right| + i\arg\frac{a}{b} = \\ = \log|a| - \log|b| + i\arg a - i\arg b + 2\pi i\left\lfloor\frac{\pi - \arg a + \arg b}{2\pi}\right\rfloor = \\ = \log a - \log b + 2\pi i\left\lfloor\frac{\pi - \arg a + \arg b}{2\pi}\right\rfloor\end{aligned} \tag{3.2.4.4}$$

3.2.5 Power

A power of two complex numbers is defined as:

$$z^a = e^{a\log z}$$

From above we can also write the power z^a in two different ways:

$$z^a = \left(e^{\log z}\right)^a = e^{\log z^a}$$

But these two cannot be used as a definition of a power, because both require the knowledge of x^a, which we are trying to define, where $x = z$ or $x = e^{\log z}$.

It follows:

$$\log x^a = \log e^{a\log x} = a\log x + 2\pi i\left\lfloor\frac{\pi - \operatorname{Im} a\log x}{2\pi}\right\rfloor \tag{3.2.5.1}$$

and

$$\begin{aligned}(x^a)^b = e^{b\log x^a} = e^{b\left(a\log x + 2\pi i\left\lfloor\frac{\pi - \operatorname{Im} a\log x}{2\pi}\right\rfloor\right)} = \\ = e^{ab\log x}e^{2\pi ib\left\lfloor\frac{\pi - \operatorname{Im} a\log x}{2\pi}\right\rfloor} = \\ = x^{ab}e^{2\pi ib\left\lfloor\frac{\pi - \operatorname{Im} a\log x}{2\pi}\right\rfloor}\end{aligned} \tag{3.2.5.2}$$

As a special case for $x = e$ one gets:

$$(e^a)^b = e^{ab}e^{2\pi ib\left\lfloor\frac{\pi - \operatorname{Im} a}{2\pi}\right\rfloor} \tag{3.2.5.3}$$

Similarly:

$$\begin{aligned}(xy)^a = e^{a\log xy} = e^{a\log x + a\log y + 2\pi ia\left\lfloor\frac{\pi - \arg x - \arg y}{2\pi}\right\rfloor} = \\ = x^a y^a e^{2\pi ia\left\lfloor\frac{\pi - \arg x - \arg y}{2\pi}\right\rfloor}\end{aligned}$$

3.2.6 Examples

For integer n we get from (3.2.5.2):

$$(x^a)^n = x^{an} e^{2\pi i n \left\lfloor \frac{\pi - \operatorname{Im} a \log x}{2\pi} \right\rfloor} = x^{an}$$

Using (3.2.5.2):

$$\sqrt{x^2} = (x^2)^{\frac{1}{2}} = x^{\frac{1}{2}\cdot 2} e^{2\pi i \frac{1}{2} \left\lfloor \frac{\pi - \operatorname{Im} 2 \log x}{2\pi} \right\rfloor} = x e^{\pi i \left\lfloor \frac{\pi - 2 \arg x}{2\pi} \right\rfloor} = (-1)^{\left\lfloor \frac{\pi - 2 \arg x}{2\pi} \right\rfloor} x$$

Using (3.2.5.3):

$$\sqrt{e^x} = (e^x)^{\frac{1}{2}} = e^{\frac{x}{2}} e^{\pi i \left\lfloor \frac{\pi - \operatorname{Im} x}{2\pi} \right\rfloor} = (-1)^{\left\lfloor \frac{\pi - \operatorname{Im} x}{2\pi} \right\rfloor} e^{\frac{x}{2}}$$

Using (3.2.4.3):

$$0 = \log 1 = \log(-1)(-1) = \log(-1) + \log(-1) + 2\pi i \left\lfloor \frac{\pi - \pi - \pi}{2\pi} \right\rfloor =$$
$$= i\pi + i\pi + 2\pi i \left\lfloor -\tfrac{1}{2} \right\rfloor = i\pi + i\pi - 2\pi i = 0$$

Code:

```
>>> from math import floor, pi
>>> from cmath import log
>>> log((-1)*(-1))
0j
>>> log(-1)+log(-1)+2*pi*1j*floor((pi-pi-pi)/(2*pi))
0j
```

Another example:

$$i^i = e^{i \log i} = e^{i^2 \arg i} = e^{-\frac{\pi}{2}}$$

Code:

```
>>> from math import exp, pi
>>> 1j**1j
(0.20787957635076193+0j)
>>> exp(-pi/2)
0.20787957635076193
```

Another example, using (3.2.5.1):

$$\log(\sqrt{z}) = \log(z^{\frac{1}{2}}) = \tfrac{1}{2} \log z + 2\pi i \left\lfloor \frac{\pi - \operatorname{Im} \frac{1}{2} \log z}{2\pi} \right\rfloor =$$
$$= \tfrac{1}{2} \log z + 2\pi i \left\lfloor \frac{\pi - \frac{1}{2} \arg z}{2\pi} \right\rfloor = \tfrac{1}{2} \log z$$

and

$$\log(z^2) = 2 \log z + 2\pi i \left\lfloor \frac{\pi - \operatorname{Im} 2 \log z}{2\pi} \right\rfloor =$$
$$= 2 \log z + 2\pi i \left\lfloor \frac{\pi - 2 \arg z}{2\pi} \right\rfloor$$

and

$$\log\left(\frac{1}{z}\right) = \log(z^{-1}) = -\log z + 2\pi i \left\lfloor \frac{\pi - \operatorname{Im}(-1)\log z}{2\pi} \right\rfloor = \\ = -\log z + 2\pi i \left\lfloor \frac{\pi + \arg z}{2\pi} \right\rfloor$$

Another example, following from (3.2.4.1) and (3.2.4.4):

$$\arg z = \frac{1}{i}(\log z - \log|z|) = \frac{1}{i}\left(\log\frac{z}{|z|} - 2\pi i \left\lfloor \frac{\pi - \arg z + \arg|z|}{2\pi} \right\rfloor\right) = \frac{1}{i}\log\frac{z}{|z|}$$

3.2.7 Complex Conjugate

The complex conjugate is defined by:

$$z = \operatorname{Re} z + i \operatorname{Im} z$$
$$\bar{z} = \operatorname{Re} z - i \operatorname{Im} z$$

Now we can solve for $\operatorname{Re} z$ and $\operatorname{Im} z$:

$$\operatorname{Re} z = \frac{1}{2}(z + \bar{z})$$
$$\operatorname{Im} z = \frac{i}{2}(-z + \bar{z})$$

Any complex function f can be written using $\operatorname{Re} z$ and $\operatorname{Im} z$, i.e. $f = f(\operatorname{Re} z, \operatorname{Im} z)$ or using z and $\bar{z}$, i.e. $f = f(z, \bar{z})$.

Examples

$$|z| = \sqrt{\operatorname{Re}^2 z + \operatorname{Im}^2 z} = \sqrt{\left(\frac{1}{2}(z+\bar{z})\right)^2 + \left(\frac{i}{2}(-z+\bar{z})\right)^2} = \sqrt{z\bar{z}}$$

$$|\bar{z}| = \sqrt{\operatorname{Re}^2 \bar{z} + \operatorname{Im}^2 \bar{z}} = \sqrt{\operatorname{Re}^2 z + (-\operatorname{Im} z)^2} = |z|$$

$$\arg z = \operatorname{atan2}(\operatorname{Im} z, \operatorname{Re} z) = \operatorname{atan2}\left(\frac{i}{2}(-z+\bar{z}), \frac{1}{2}(z+\bar{z})\right) = \operatorname{atan2}\left(i(-z+\bar{z}), z+\bar{z}\right)$$

$$\arg \bar{z} = \operatorname{atan2}(-\operatorname{Im} z, \operatorname{Re} z) = -\operatorname{atan2}(\operatorname{Im} z, \operatorname{Re} z) + 2\pi \left\lfloor \frac{\operatorname{atan2}(\operatorname{Im} z, \operatorname{Re} z) + \pi}{2\pi} \right\rfloor = -\arg z + 2\pi \left\lfloor \frac{\arg z + \pi}{2\pi} \right\rfloor$$

$$\overline{\log z} = \log|z| - i\arg z = \log|\bar{z}| + i\arg\bar{z} - 2\pi i \left\lfloor \frac{\arg z + \pi}{2\pi} \right\rfloor = \log\bar{z} - 2\pi i \left\lfloor \frac{\arg z + \pi}{2\pi} \right\rfloor$$

$$\overline{\sqrt{z}} = \overline{z^{\frac{1}{2}}} = \overline{e^{\frac{1}{2}\log z}} = \overline{e^{\frac{1}{2}(\log|z| + i\arg z)}} = e^{\frac{1}{2}(\log|z| - i\arg z)} = e^{\frac{1}{2}\left(\log|\bar{z}| + i\arg\bar{z} - 2\pi i\left\lfloor\frac{\arg z+\pi}{2\pi}\right\rfloor\right)} = \\ = (\bar{z})^{\frac{1}{2}} e^{-\pi i\left\lfloor\frac{\arg z+\pi}{2\pi}\right\rfloor} = (-1)^{\left\lfloor\frac{\arg z+\pi}{2\pi}\right\rfloor}\sqrt{\bar{z}}$$

3.2.8 Complex Derivatives

The complex derivative is defined by

$$\frac{\mathrm{d}f}{\mathrm{d}z} = \lim_{h\to 0}\frac{f(z+h) - f(z)}{h} \tag{3.2.8.1}$$

Let's calculate the complex derivative in the direction θ, i.e. we use $h = te^{i\theta}$ with real t and we introduce $f = f(x, y)$ with $x = \operatorname{Re} z$, $y = \operatorname{Im} z$ to simplify the notation:

$$\begin{aligned}
\frac{\mathrm{d}f}{\mathrm{d}z} &= \lim_{t\to 0} \frac{f(z + te^{i\theta}) - f(z)}{te^{i\theta}} = \\
&= \lim_{t\to 0} \frac{f(x + t\cos\theta, y + t\sin\theta) - f(x, y)}{t} e^{-i\theta} = \\
&= \frac{\mathrm{d}}{\mathrm{d}t} f(x + t\cos\theta, y + t\sin\theta) e^{-i\theta} = \\
&= \left(\frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta\right) e^{-i\theta} = \\
&= \left(\frac{\partial f}{\partial x}\frac{e^{i\theta} + e^{-i\theta}}{2} + \frac{\partial f}{\partial y}\frac{e^{i\theta} - e^{-i\theta}}{2i}\right) e^{-i\theta} = \\
&= \frac{\partial f}{\partial x}\frac{1 + e^{-2i\theta}}{2} + \frac{\partial f}{\partial y}\frac{1 - e^{-2i\theta}}{2i} = \\
&= \tfrac{1}{2}\left(\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y}\right) + \tfrac{1}{2}\left(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}\right) e^{-2i\theta} = \\
&= \frac{\partial f}{\partial z} + \frac{\partial f}{\partial \bar{z}} e^{-2i\theta}
\end{aligned}$$

In the last step we have expressed the derivatives with respect to x, y in terms of derivatives with respect to z, $\bar{z}$, using the relations:

$$\begin{aligned}
\frac{\partial f}{\partial z} &= \frac{\partial x}{\partial z}\frac{\partial f}{\partial x} + \frac{\partial y}{\partial z}\frac{\partial f}{\partial y} = \\
&= \frac{1}{2}\frac{\partial f}{\partial x} - \frac{i}{2}\frac{\partial f}{\partial y} = \\
&= \tfrac{1}{2}\left(\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y}\right)
\end{aligned} \tag{3.2.8.2}$$

$$\begin{aligned}
\frac{\partial f}{\partial \bar{z}} &= \frac{\partial x}{\partial \bar{z}}\frac{\partial f}{\partial x} + \frac{\partial y}{\partial \bar{z}}\frac{\partial f}{\partial y} = \\
&= \frac{1}{2}\frac{\partial f}{\partial x} + \frac{i}{2}\frac{\partial f}{\partial y} = \\
&= \tfrac{1}{2}\left(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}\right)
\end{aligned} \tag{3.2.8.3}$$

Let's repeat the important result:

$$\frac{\mathrm{d}f(z, \bar{z})}{\mathrm{d}z} = \frac{\partial f(z, \bar{z})}{\partial z} + \frac{\partial f(z, \bar{z})}{\partial \bar{z}} e^{-2i\theta} \tag{3.2.8.4}$$

The equation (3.2.8.4) states that the complex derivative along the direction θ of any function can be calculated, but the result in general depends on θ. The derivatives for all possible angles θ lie on a circle, with the center $\frac{\partial f}{\partial z}$ and the radius $\left|\frac{\partial f}{\partial \bar{z}}\right|$. When the derivative has different values for different θ, i.e. when $\frac{\partial f}{\partial \bar{z}} \neq 0$, it means that the complex limit (3.2.8.1) does not exist. On the other hand, if the derivative does not depend on θ, i.e. when $\frac{\partial f}{\partial \bar{z}} = 0$, then the complex limit (3.2.8.1) exists, and the function has a complex derivative — such functions are called analytic. Analytic functions thus do not depend on $\bar{z}$ and we can write just $f = f(z)$ for those.

The $\frac{\partial f}{\partial z}$ and $\frac{\partial f}{\partial \bar{z}}$ are called Wirtinger derivatives.

We can see that the function is analytic (i.e. has a complex derivative) if and only if:

$$\frac{\partial f}{\partial \bar{z}} = \tfrac{1}{2}\left(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}\right) = 0$$

We can write $f = u + iv$:

$$\begin{aligned}\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y} &= 0\\ \frac{\partial(u+iv)}{\partial x} + i\frac{\partial(u+iv)}{\partial y} &= 0\\ \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) + i\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right) &= 0\end{aligned}$$

both the real and imaginary parts must be equal to zero:

$$\begin{aligned}\frac{\partial u}{\partial x} &= \frac{\partial v}{\partial y}\\ \frac{\partial u}{\partial y} &= -\frac{\partial v}{\partial x}\end{aligned}$$

These are called the Cauchy-Riemann equations.

We can derive the chain rule:

$$\begin{aligned}\frac{\mathrm{d}f(g)}{\mathrm{d}z} = \frac{\partial f(g)}{\partial z} + \frac{\partial f(g)}{\partial \bar{z}}e^{-2i\theta} =\\ = \left(\frac{\partial f}{\partial g}\frac{\partial g}{\partial z} + \frac{\partial f}{\partial \bar{g}}\frac{\partial \bar{g}}{\partial z}\right) + \left(\frac{\partial f}{\partial g}\frac{\partial g}{\partial \bar{z}} + \frac{\partial f}{\partial \bar{g}}\frac{\partial \bar{g}}{\partial \bar{z}}\right)e^{-2i\theta} =\\ = \frac{\partial f}{\partial g}\left(\frac{\partial g}{\partial z} + \frac{\partial g}{\partial \bar{z}}e^{-2i\theta}\right) + \frac{\partial f}{\partial \bar{g}}\left(\frac{\partial \bar{g}}{\partial z} + \frac{\partial \bar{g}}{\partial \bar{z}}e^{-2i\theta}\right) =\\ = \frac{\partial f}{\partial g}\frac{\mathrm{d}g}{\mathrm{d}z} + \frac{\partial f}{\partial \bar{g}}\frac{\mathrm{d}\bar{g}}{\mathrm{d}z}\end{aligned} \tag{3.2.8.5}$$

Another useful formula is the derivative of a conjugate function:

$$\begin{aligned}\frac{\mathrm{d}\bar{f}}{\mathrm{d}z} = \frac{\partial \bar{f}}{\partial z} + \frac{\partial \bar{f}}{\partial \bar{z}}e^{-2i\theta} = \overline{\frac{\partial f}{\partial \bar{z}}} + \overline{\frac{\partial f}{\partial z}}e^{-2i\theta} =\\ = \overline{\left(\frac{\partial f}{\partial \bar{z}}e^{-2i\theta} + \frac{\partial f}{\partial z}\right)}e^{-2i\theta} = \overline{\frac{\mathrm{d}f}{\mathrm{d}z}}e^{-2i\theta}\end{aligned} \tag{3.2.8.6}$$

Using (3.2.8.6), the chain rule (3.2.8.5) can also be written as:

$$\frac{\mathrm{d}f(g)}{\mathrm{d}z} = \frac{\partial f}{\partial g}\frac{\mathrm{d}g}{\mathrm{d}z} + \frac{\partial f}{\partial \bar{g}}\frac{\mathrm{d}\bar{g}}{\mathrm{d}z} = \frac{\partial f}{\partial g}\frac{\mathrm{d}g}{\mathrm{d}z} + \frac{\partial f}{\partial \bar{g}}\overline{\frac{\mathrm{d}g}{\mathrm{d}z}}e^{-2i\theta} \tag{3.2.8.7}$$

Which has the advantage that only the $\frac{\mathrm{d}g}{\mathrm{d}z}$ derivative is needed, the rest is just conjugation and multiplication. If f is analytic, then $\frac{\partial f}{\partial \bar{g}} = 0$, the second term vanishes and the chain rule is analogous to real functions.

Examples

$$\frac{\mathrm{d}z}{\mathrm{d}z} = \frac{\partial z}{\partial z} + \frac{\partial z}{\partial \bar{z}} e^{-2i\theta} = 1$$

$$\frac{\mathrm{d}\bar{z}}{\mathrm{d}z} = \frac{\partial \bar{z}}{\partial z} + \frac{\partial \bar{z}}{\partial \bar{z}} e^{-2i\theta} = e^{-2i\theta}$$

$$\frac{\mathrm{d}\,\mathrm{Re}\, z}{\mathrm{d}z} = \frac{\mathrm{d}\frac{1}{2}(z+\bar{z})}{\mathrm{d}z} = \frac{\partial \frac{1}{2}(z+\bar{z})}{\partial z} + \frac{\partial \frac{1}{2}(z+\bar{z})}{\partial \bar{z}} e^{-2i\theta} = \tfrac{1}{2} + \tfrac{1}{2} e^{-2i\theta}$$

$$\frac{\mathrm{d}\,\mathrm{Im}\, z}{\mathrm{d}z} = \frac{\mathrm{d}\frac{i}{2}(-z+\bar{z})}{\mathrm{d}z} = \frac{\partial \frac{i}{2}(-z+\bar{z})}{\partial z} + \frac{\partial \frac{i}{2}(-z+\bar{z})}{\partial \bar{z}} e^{-2i\theta} = -\frac{i}{2} + \frac{i}{2} e^{-2i\theta}$$

$$\frac{\mathrm{d}|z|}{\mathrm{d}z} = \frac{\mathrm{d}\sqrt{z\bar{z}}}{\mathrm{d}z} = \frac{\partial \sqrt{z\bar{z}}}{\partial z} + \frac{\partial \sqrt{z\bar{z}}}{\partial \bar{z}} e^{-2i\theta} = \frac{\bar{z} + z e^{-2i\theta}}{2\sqrt{z\bar{z}}} = \frac{\bar{z} + z e^{-2i\theta}}{2|z|}$$

$$\frac{\mathrm{d}|f(z)|}{\mathrm{d}z} = \frac{\partial |f|}{\partial f}\frac{\mathrm{d}f}{\mathrm{d}z} + \frac{\partial |f|}{\partial \bar{f}}\frac{\mathrm{d}\bar{f}}{\mathrm{d}z} = \frac{\bar{f}\frac{\mathrm{d}f}{\mathrm{d}z} + f\frac{\mathrm{d}\bar{f}}{\mathrm{d}z}}{2|f|}$$

$$\frac{\mathrm{d}\arg z}{\mathrm{d}z} = \frac{\mathrm{d}\ \mathrm{atan}\,2\left(i(-z+\bar{z}), z+\bar{z}\right)}{\mathrm{d}z} = \frac{\partial\ \mathrm{atan}\,2\left(i(-z+\bar{z}), z+\bar{z}\right)}{\partial z} + \frac{\partial\ \mathrm{atan}\,2\left(i(-z+\bar{z}), z+\bar{z}\right)}{\partial \bar{z}} e^{-2i\theta} =$$

$$= \frac{(z+\bar{z})(-i) - i(-z+\bar{z})}{4z\bar{z}} + \frac{(z+\bar{z})i - i(-z+\bar{z})}{4z\bar{z}} e^{-2i\theta} =$$

$$= \frac{i}{2}\left(-\frac{1}{z} + \frac{1}{\bar{z}} e^{-2i\theta}\right) = \frac{i}{2}\left(\frac{-\bar{z} + z e^{-2i\theta}}{|z|^2}\right)$$

$$\frac{\mathrm{d}\log|z|}{\mathrm{d}z} = \frac{1}{|z|}\frac{\bar{z} + z e^{-2i\theta}}{2|z|} = \frac{\bar{z} + z e^{-2i\theta}}{2|z|^2}$$

$$\frac{\mathrm{d}\log z}{\mathrm{d}z} = \frac{\mathrm{d}(\log|z| + i\arg z)}{\mathrm{d}z} = \frac{\bar{z} + z e^{-2i\theta}}{2|z|^2} + i\frac{i}{2}\left(\frac{-\bar{z} + z e^{-2i\theta}}{|z|^2}\right) = \frac{\bar{z}}{|z|^2} = \frac{\bar{z}}{z\bar{z}} = \frac{1}{z}$$

$$\frac{\mathrm{d}\overline{\log z}}{\mathrm{d}z} = \frac{\partial \overline{\log z}}{\partial z} + \frac{\partial \overline{\log z}}{\partial \bar{z}} e^{-2i\theta} = \overline{\frac{\partial \log z}{\partial \bar{z}}} + \overline{\frac{\partial \log z}{\partial z}} e^{-2i\theta} = \frac{1}{\bar{z}} e^{-2i\theta}$$

$$\frac{\mathrm{d}|\log z|}{\mathrm{d}z} = \frac{\overline{\log z}\frac{\mathrm{d}\log z}{\mathrm{d}z} + \log z\frac{\mathrm{d}\overline{\log z}}{\mathrm{d}z}}{2|\log z|} = \frac{\frac{1}{z}\overline{\log z} + \frac{1}{\bar{z}}(\log z)e^{-2i\theta}}{2|\log z|} = \frac{\bar{z}\overline{\log z} + z(\log z)e^{-2i\theta}}{2z\bar{z}|\log z|}$$

Note that if z is real, i.e. $z = \bar{z}$, we recover the real derivative results by setting $\theta = 0$, i.e. taking the derivative along

the x-axis:

$$\begin{aligned}
\frac{\mathrm{d}x}{\mathrm{d}x} &= 1 \\
\frac{\mathrm{d}\,\mathrm{Re}\,x}{\mathrm{d}x} &= \tfrac{1}{2} + \tfrac{1}{2} = 1 \\
\frac{\mathrm{d}\,\mathrm{Im}\,x}{\mathrm{d}x} &= -\frac{i}{2} + \frac{i}{2} = 0 \\
\frac{\mathrm{d}|x|}{\mathrm{d}x} &= \frac{x+x}{2|x|} = \frac{x}{|x|} \\
\frac{\mathrm{d}|f(x)|}{\mathrm{d}x} &= \frac{f\frac{\mathrm{d}f}{\mathrm{d}x} + f\frac{\mathrm{d}f}{\mathrm{d}z}}{2|f|} = \frac{f\frac{\mathrm{d}f}{\mathrm{d}x}}{|f|} \\
\frac{\mathrm{d}\,\arg x}{\mathrm{d}x} &= \frac{i}{2}\left(-\frac{1}{x} + \frac{1}{x}\right) = 0 \\
\frac{\mathrm{d}\log|x|}{\mathrm{d}x} &= \frac{x+x}{2|x|^2} = \frac{x}{|x|^2} \\
\frac{\mathrm{d}\log x}{\mathrm{d}x} &= \frac{1}{x} \\
\frac{\mathrm{d}|\log x|}{\mathrm{d}x} &= \frac{x\log x + x\log x}{2x^2|\log x|} = \frac{\log x}{x|\log x|}
\end{aligned}$$

The above approach to first express things in terms of z and $\bar{z}$ and then differentiate is probably the easiest, but we can do things in any order we want. For example the derivative of $|z|$ can also be calculated in this (arguably more complicated) way:

$$\begin{aligned}
\frac{\mathrm{d}|z|}{\mathrm{d}z} = \frac{\mathrm{d}\sqrt{\mathrm{Re}^2\,z + \mathrm{Im}^2\,z}}{\mathrm{d}z} = \frac{\mathrm{Re}\,z\frac{\mathrm{d}\,\mathrm{Re}\,z}{\mathrm{d}z} + \mathrm{Im}\,z\frac{\mathrm{d}\,\mathrm{Im}\,z}{\mathrm{d}z}}{|z|} = \\
= \frac{\left(\frac{1}{2}(z+\bar{z})\right)\left(\frac{1}{2} + \frac{1}{2}e^{-2i\theta}\right) + \left(\frac{i}{2}(-z+\bar{z})\right)\left(-\frac{i}{2} + \frac{i}{2}e^{-2i\theta}\right)}{|z|} = \\
= \frac{\bar{z} + ze^{-2i\theta}}{2|z|}
\end{aligned}$$

3.2.9 Testing Identities Using Computer Code

All the complex identities in this chapter can be tested using the following codes. `test_complex.py`:

```
def arg(x):
    """
    The argument function.
    """
    from cmath import log
    return log(x).imag

def generate_values():
    """
    Create values to test the function at.
    """
    from math import sin, cos, pi
    from random import random
    # Generate 3 circles in complex plane, with diameters 0.5, 1 and 2. We
    # avoid special values like -1, +/- i, etc., because they typically send
    # the numerical values close to the branch cut, and numerical errors then
```

(continues on next page)

(continued from previous page)

```
    # flip the sign, e.g.:
    #    In [1]: sqrt((-0.5j)**2)
    #    Out[1]: -0.5j
    #
    #    In [2]: (-0.5j)**2
    #    Out[2]: (-0.25-0j)
    #
    #    In [3]: sqrt(-0.25)
    #    Out[3]: 0.5j
    # Here [3] is the correct value and [1] is incorrect, but that happens due
    # to the round off errors in [2] (the small negative imaginary part makes
    # sqrt() return -0.5j instead of +0.5j).
    #
    # For this reason, we chose N=7.
    N = 7
    circle = []
    for n in range(N): circle.append(cos(2*pi*n/N)+1j*sin(2*pi*n/N))
    values = []
    for n in range(N): values.append(0.5*circle[n])
    for n in range(N): values.append(1.0*circle[n])
    for n in range(N): values.append(2.0*circle[n])
    # Add some random points:
    for n in range(30):
        values.append((random()-0.5)*20 + 1j*(random()-0.5)*20)
    return values
values = generate_values()

def feq(a, b, max_relative_error=1e-12, max_absolute_error=1e-12):
    """
    Returns True if a==b to the given relative and absolute errors, otherwise
    False.
    """
    # if the numbers are close enough (absolutely), then they are equal
    if abs(a-b) < max_absolute_error:
        return True
    # if not, they can still be equal if their relative error is small
    if abs(b) > abs(a):
        relative_error = abs(a-b)/abs(b)
    else:
        relative_error = abs(a-b)/abs(a)
    #print abs(a-b), relative_error
    return relative_error <= max_relative_error

def test_zero1(lhs, rhs):
    """
    Tests that a complex function f(x) of one complex variable is zero.
    """
    for x in values:
        r = feq(lhs(x), rhs(x))
        if not r:
            print "x lhs(x) rhs(x)"
            print x, lhs(x), rhs(x)
            assert False

def test_zero2(lhs, rhs):
    """
    Tests that a complex function f(x, y) of two complex variables is zero.
```

(continues on next page)

(continued from previous page)

```
    """
    for x in values:
        for y in values:
            r = feq(lhs(x, y), rhs(x, y))
            if not r:
                print "x y lhs(x, y) rhs(x, y)"
                print x, y, lhs(x, y), rhs(x, y)
                assert False

def test_zero3(lhs, rhs):
    """
    Tests that a complex function f(x, y, z) of three complex variables is zero.
    """
    for x in values:
        for y in values:
            for z in values:
                r = feq(lhs(x, y, z), rhs(x, y, z))
                if not r:
                    print "x y z lhs(x, y, z) rhs(x, y, z)"
                    print x, y, z, lhs(x, y, z), rhs(x, y, z)
                    assert False

from math import floor, pi
from cmath import sqrt, exp, log
I = 1j

# Test the various identities
test_zero1(lambda x: sqrt(x**2), lambda x: (-1)**floor((pi-2*arg(x))/(2*pi))*x)
test_zero1(lambda x: sqrt(exp(x)), lambda x: (-1)**floor((pi-x.imag)/(2*pi))*exp(x/2))
test_zero1(lambda x: log(exp(x)), lambda x: x+2*pi*I*floor((pi-x.imag)/(2*pi)))
test_zero1(lambda x: log(abs(exp(x))), lambda x: x.real)
test_zero1(lambda z: z, lambda z: abs(z)*exp(I*arg(z)))
test_zero1(lambda z: arg(exp(z)), lambda z: z.imag + 2*pi*floor((pi-z.imag)/(2*pi)))
test_zero1(lambda z: sqrt(z).conjugate(), lambda z: (-1)**floor((arg(z)+pi)/
↪(2*pi))*sqrt(z.conjugate()))
test_zero1(lambda z: arg(z.conjugate()), lambda z: -arg(z) + 2*pi*floor((arg(z)+pi)/
↪(2*pi)))

test_zero2(lambda a,b: exp(a)**b, lambda a,b: exp(a*b)*exp(2*pi*I*b*floor((pi-a.imag)/
↪(2*pi))))
test_zero2(lambda x,a: log(x**a), lambda x,a: a*log(x)+2*pi*I*floor((pi-(a*log(x)).
↪imag)/(2*pi)))
test_zero2(lambda a,b: log(a*b), lambda a,b: log(a)+log(b)+2*pi*I*floor((pi-arg(a)-
↪arg(b))/(2*pi)))
test_zero2(lambda a,b: arg(a*b), lambda a,b: arg(a)+arg(b)+2*pi*floor((pi-arg(a)-
↪arg(b))/(2*pi)))

test_zero3(lambda x,a,b: (x**a)**b, lambda x,a,b: x**(a*b)*exp(2*pi*I*b*floor((pi-
↪(a*log(x)).imag)/(2*pi))))
test_zero3(lambda x,y,a: (x*y)**a, lambda x,y,a: (x**a)*(y**a)*exp(2*pi*I*a*floor((pi-
↪arg(x)-arg(y))/(2*pi))))
```

test_complex_diff.py:

```
def arg(x):
    """
    The argument function.
```

(continues on next page)

(continued from previous page)

```
    """
    from cmath import log
    return log(x).imag

def generate_values():
    """
    Create values to test the function at.
    """
    from math import sin, cos, pi
    from random import random
    # Generate 3 circles in complex plane, with diameters 0.5, 1 and 2. We
    # avoid special values like -1, +/- i, etc., because they typically send
    # the numerical values close to the branch cut, and numerical errors then
    # flip the sign, e.g.:
    #    In [1]: sqrt((-0.5j)**2)
    #    Out[1]: -0.5j
    #
    #    In [2]: (-0.5j)**2
    #    Out[2]: (-0.25-0j)
    #
    #    In [3]: sqrt(-0.25)
    #    Out[3]: 0.5j
    # Here [3] is the correct value and [1] is incorrect, but that happens due
    # to the round off errors in [2] (the small negative imaginary part makes
    # sqrt() return -0.5j instead of +0.5j).
    #
    # For this reason, we chose N=7.
    N = 7
    circle = []
    for n in range(N): circle.append(cos(2*pi*n/N)+1j*sin(2*pi*n/N))
    values = []
    for n in range(N): values.append(0.5*circle[n])
    for n in range(N): values.append(1.0*circle[n])
    for n in range(N): values.append(2.0*circle[n])
    # Add some random points:
    for n in range(30):
        values.append((random()-0.5)*20 + 1j*(random()-0.5)*20)
    return values
values = generate_values()

def feq(a, b, max_relative_error=1e-12, max_absolute_error=1e-12):
    """
    Returns True if a==b to the given relative and absolute errors, otherwise
    False.
    """
    # if the numbers are close enough (absolutely), then they are equal
    if abs(a-b) < max_absolute_error:
        return True
    # if not, they can still be equal if their relative error is small
    if abs(b) > abs(a):
        relative_error = abs(a-b)/abs(b)
    else:
        relative_error = abs(a-b)/abs(a)
    #print abs(a-b), relative_error
    return relative_error <= max_relative_error

def test_zero1(lhs, rhs):
```

(continues on next page)

(continued from previous page)

```
    """
    Tests that a complex function f(x) of one complex variable is zero.
    """
    for x in values:
        r = feq(lhs(x), rhs(x))
        if not r:
            print "x lhs(x) rhs(x)"
            print x, lhs(x), rhs(x)
            assert False

def test_zero2(lhs, rhs):
    """
    Tests that a complex function f(x, y) of two complex variables is zero.
    """
    for x in values:
        for y in values:
            r = feq(lhs(x, y), rhs(x, y))
            if not r:
                print "x y lhs(x, y) rhs(x, y)"
                print x, y, lhs(x, y), rhs(x, y)
                assert False

def test_zero3(lhs, rhs):
    """
    Tests that a complex function f(x, y, z) of three complex variables is zero.
    """
    for x in values:
        for y in values:
            for z in values:
                r = feq(lhs(x, y, z), rhs(x, y, z))
                if not r:
                    print "x y z lhs(x, y, z) rhs(x, y, z)"
                    print x, y, z, lhs(x, y, z), rhs(x, y, z)
                    assert False

def diff(f, z0, theta, eps=1e-8):
    h = eps*exp(I*theta)
    return (f(z0+h)-f(z0)) / h

def diff2(dfdz, dfdconjz, z0, theta):
    return dfdz(z0) + dfdconjz(z0)*exp(-2*I*theta)

def test_zero(f, dfdz, dfdconjz, z0, theta, eps=1e-8):
    assert feq(diff(f, z0, theta, eps), diff2(dfdz, dfdconjz, z0, theta),
            max_relative_error=eps*1e2, max_absolute_error=eps*1e2)

from math import floor, pi
from cmath import sqrt, exp, log
I = 1j

angles = [0, pi/7, pi/4, pi/2, 3*pi/4, pi]

for x in values:
    for theta in angles:
        test_zero(lambda x: abs(x), lambda x: x.conjugate()/(2*abs(x)),
                lambda x: x/(2*abs(x)), x, theta)
        test_zero(lambda x: log(x), lambda x: 1/x, lambda x: 0, x, theta)
```

(continues on next page)

(continued from previous page)

```
        test_zero(lambda x: log(exp(x-x.conjugate())), lambda x: 1,
                lambda x: -1, x, theta)
        test_zero(lambda x: sqrt(x**2), lambda x: sqrt(x**2)/x, lambda x: 0, x, theta)
        test_zero(lambda x: sqrt(x**2), lambda x: x/sqrt(x**2), lambda x: 0, x, theta)
```

3.3 Residue Theorem

The Residue Theorem says that a contour integral of an analytic function f over a closed curve γ (loop) is equal to the sum of residues $\mathrm{Res}_{z_k} f(z)$ of the function at all singularities z_k inside the loop:

$$\int_\gamma f(z)\mathrm{d}z = 2\pi i \sum_{z_k} \mathrm{Res}_{z=z_k} f(z)$$

Residue $\mathrm{Res}_{z_0} f(z)$ is defined as the contour integral around z_0 divided by $2\pi i$:

$$\mathrm{Res}_{z=z_0} f(z) = \frac{1}{2\pi i} \int_{|z-z_0|=\epsilon} f(z)\mathrm{d}z$$

and it is equal to the coefficient of $\frac{1}{z-z_0}$ in the Laurent series of $f(z)$ around the point z_0, as can be easily calculated:

$$\begin{aligned}\mathrm{Res}_{z=z_0} f(z) = \frac{1}{2\pi i} \int\limits_{|z-z_0|=\epsilon} f(z)\mathrm{d}z = \frac{1}{2\pi i} \int\limits_{|z-z_0|=\epsilon} \sum_{n=-\infty}^{\infty} c_n (z-z_0)^n \mathrm{d}z = \\ = \sum_{n=-\infty}^{\infty} c_n \frac{1}{2\pi i} \int\limits_{|z-z_0|=\epsilon} (z-z_0)^n \mathrm{d}z = \sum_{n=-\infty}^{\infty} c_n \delta_{n,-1} = c_{-1}\end{aligned}$$

where we used the result of the following integral (we integrate over the curve $z = z_0 + \epsilon e^{i\varphi}$, $0 \le \varphi < 2\pi$, so $\mathrm{d}z = i\epsilon e^{i\varphi}\mathrm{d}\varphi$):

$$\begin{aligned}\frac{1}{2\pi i} \int\limits_{|z-z_0|=\epsilon} (z-z_0)^n \mathrm{d}z = \frac{1}{2\pi i} \int_0^{2\pi} (z_0 + \epsilon e^{i\varphi} - z_0)^n i\epsilon e^{i\varphi} \mathrm{d}\varphi = \frac{\epsilon^{n+1}}{2\pi} \int_0^{2\pi} e^{i\varphi(n+1)} \mathrm{d}\varphi = \\ = \begin{cases} \frac{\epsilon^{n+1}}{2\pi} \left[\frac{e^{i\varphi(n+1)}}{i(n+1)}\right]_0^{2\pi} = 0 & \text{for } n \neq -1 \\ \frac{1}{2\pi} \int_0^{2\pi} \mathrm{d}\varphi = 1 & \text{for } n = -1 \end{cases} = \delta_{n,-1}\end{aligned}$$

3.3.1 Computation of Residues

One has to calculate the c_{-1} coefficient in the Laurent series. One way to do that is to write $f(z)$ as:

$$f(z) = \frac{H(z)}{(z-z_0)^m}$$

where $H(z)$ is analytic in the vicinity of z_0, e.g. $f(z)$ has a pole of order m at z_0. Then:

$$\mathrm{Res}_{z=z_0} f(z) = c_{-1} = \frac{1}{(m-1)!} \left.\frac{\mathrm{d}^m H(z)}{\mathrm{d}z^m}\right|_{z=z_0}$$

in particular for $m = 1$:

$$\mathrm{Res}_{z=z_0} f(z) = H(z_0) = \lim_{z\to z_0}(z - z_0)f(z)$$

for $m = 2$:

$$\mathrm{Res}_{z=z_0} f(z) = H'(z_0) = \lim_{z\to z_0}\frac{\mathrm{d}}{\mathrm{d}z}[(z - z_0)^2 f(z)]$$

f has a pole of order 1 at z_0, g is analytic at z_0:

$$\mathrm{Res}_{z=z_0} f(z)g(z) = \lim_{z\to z_0}(z - z_0)f(z)g(z) = g(z_0)\lim_{z\to z_0}(z - z_0)f(z) = g(z_0)\,\mathrm{Res}_{z=z_0} f(z)$$

$f(z_0) = 0$, but $f'(z_0) \neq 0$ and g is analytic at z_0:

$$\mathrm{Res}_{z=z_0}\frac{g(z)}{f(z)} = g(z_0)\lim_{z\to z_0}\frac{z - z_0}{f(z)} = g(z_0)\lim_{z\to z_0}\frac{z - z_0}{f(z) - f(z_0)} = \frac{g(z_0)}{f'(z_0)}$$

3.3.2 Useful Formulas

Jordan's Lemma

For estimating integrals over semicircles Ω ($z = Re^{i\varphi}$, $0 \leq \varphi \leq \pi$), we can use the following estimates:

$$\left|\int_\Omega g(z)\mathrm{d}z\right| \leq \pi R \max_\Omega |g(z)|$$
$$\left|\int_\Omega e^{i\alpha z} g(z)\mathrm{d}z\right| \leq \frac{\pi}{\alpha}\max_\Omega |g(z)| \quad \text{for } \alpha > 0$$

(In the first case the integration path can be extended to the full circle if needed ($0 \leq \varphi \leq 2\pi$), in the second case the semicircle is the maximum path. Also if $\alpha < 0$, we need to integrate over the lower semicircle.) These formulas can be used to make sure the integral over the semicircle goes to zero as $R \to \infty$. Intuitively speaking, in the first case $g(z)$ must vanish faster than $\frac{1}{R}$ (e.g. $\frac{1}{R^2}$ is ok), in the second case it's enough if $g(z)$ just goes to 0 (no matter how fast).

The estimates can be proved easily:

$$\left|\int_\Omega g(z)\mathrm{d}z\right| = \left|\int_0^\pi g(Re^{i\varphi})iRe^{i\varphi}\mathrm{d}\varphi\right| \leq \int_0^\pi \left|g(Re^{i\varphi})\right| R\mathrm{d}\varphi \leq R\max_\Omega |g(z)|\int_0^\pi \mathrm{d}\varphi = \pi R\max_\Omega |g(z)|$$

and

$$\left|\int_\Omega e^{i\alpha z} g(z)\mathrm{d}z\right| = \left|\int_0^\pi e^{i\alpha Re^{i\varphi}} g(Re^{i\varphi})iRe^{i\varphi}\mathrm{d}\varphi\right| \leq$$
$$\leq \int_0^\pi e^{-\alpha R\sin\varphi}\left|g(Re^{i\varphi})\right| R\mathrm{d}\varphi \leq R\max_\Omega |g(z)|\int_0^\pi e^{-\alpha R\sin\varphi}\mathrm{d}\varphi <$$
$$< \frac{\pi}{\alpha}\max_\Omega |g(z)|$$

where we use the following useful estimate for the integral (valid for $\alpha > 0$):

$$\int_0^\pi e^{-\alpha R\sin\varphi}\mathrm{d}\varphi < 2\int_0^{\frac{\pi}{2}} e^{-\alpha R\frac{2}{\pi}\varphi}\mathrm{d}\varphi = 2\left[\frac{e^{-\alpha R\frac{2}{\pi}\varphi}}{-\alpha R\frac{2}{\pi}}\right]_0^{\pi/2} =$$
$$= \frac{2}{-\alpha R\frac{2}{\pi}}\left[e^{-\alpha R} - 1\right] = \frac{\pi}{\alpha R}(1 - e^{-\alpha R}) < \frac{\pi}{\alpha R}$$

Other

Sometimes it is useful to integrate over the arc $z = z_0 + \epsilon e^{i\varphi}$, $\varphi_0 \le \varphi \le \varphi_0 + \alpha$, and let $\epsilon \to 0$ at the end. If the function is analytic, the result is 0. If the function has a pole of order $n > 1$, the result is infinity, unless it's a full circle (in which case the result is 0). The remaining case is if the function has a pole of order one, e.g. it can be written ($H(z)$ is analytic at z_0):

$$f(z) = \frac{H(z)}{z - z_0}$$

Then:

$$\int_\Omega f(z)\mathrm{d}z = \int_\Omega \frac{H(z)}{z - z_0}\mathrm{d}z = \int_{\varphi_0}^{\varphi_0+\alpha} \frac{H(z_0 + \epsilon e^{i\varphi})}{z_0 + \epsilon e^{i\varphi} - z_0}\epsilon i e^{i\varphi}\mathrm{d}\varphi =$$
$$= \int_{\varphi_0}^{\varphi_0+\alpha} H(z_0 + \epsilon e^{i\varphi}) i \mathrm{d}\varphi \to \int_{\varphi_0}^{\varphi_0+\alpha} H(z_0) i \mathrm{d}\varphi = i\alpha H(z_0) = i\alpha \operatorname{Res}_{z=z_0} f(z)$$

3.3.3 Complex Substitution

When substituting in integrals, as long as we just substitute for real functions, we use the regular substitution theorem, e.g. $x = y + 1$ ($f(x)$ can be a complex function):

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = \int_{-\infty}^{\infty} f(y+1)\mathrm{d}y$$

if, on the other hand, we substitute for complex functions, e.g. $x = iy$:

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = \int_{i\infty}^{-i\infty} f(iy) i \mathrm{d}y \to \int_{\infty}^{-\infty} f(iy) i \mathrm{d}y$$

then the first two integrals in the left hand side are equal, however the integral on the right hand side is over a different integration path and we need to use the Residue Theorem to relate those integrals, e.g. in general the two integrals on the LHS and the integral on the RHS are not equal. However the idea is that the integral after the substitution (and changing the limits, e.g. the integration path) is easier to evaluate, so the substitution guides us which integration path to choose for the Residue Theorem.

3.4 Fourier Transform

The 1D Fourier transform is:

$$F[f(x)] \equiv \tilde{f}(\omega) = \int_{-\infty}^{\infty} f(x) e^{-i\omega x}\,\mathrm{d}x$$
$$F^{-1}[\tilde{f}(\omega)] = f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{f}(\omega) e^{+i\omega x}\,\mathrm{d}\omega$$

To show that it works:

$$F^{-1}F[f(x)] = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(x)e^{-i\omega x}\,\mathrm{d}x\right] e^{+i\omega x}\,\mathrm{d}\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(x')e^{-i\omega x'}\,\mathrm{d}x'\right] e^{+i\omega x}\,\mathrm{d}\omega =$$
$$= \int_{-\infty}^{\infty} f(x')\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega(x-x')}\,\mathrm{d}\omega\right]\mathrm{d}x' = \int_{-\infty}^{\infty} f(x')\delta(x-x')\,\mathrm{d}x' = f(x)$$

If x is time (unit $[\mathrm{s}]$), then ω is angular frequency (unit $[\mathrm{rad}/\mathrm{s}]$). One can express the Fourier transform in terms of ordinary frequency ν (unit $[1/\mathrm{s}] = [\mathrm{Hz}]$) by substituting $\omega = 2\pi\nu$:

$$\tilde{f}(\omega) = \tilde{f}(2\pi\nu) \equiv \tilde{f}'(\nu) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i \nu x} \,\mathrm{d}x$$
$$f(x) = \int_{-\infty}^{\infty} \tilde{f}'(\nu) e^{+2\pi i \nu x} \,\mathrm{d}\nu$$

Both transformations are equivalent and only differ in whether we express the transform in terms of ω or ν, the conversion being given by $\tilde{f}(\omega) = \tilde{f}(2\pi\nu) \equiv \tilde{f}'(\nu)$. Third frequently used convention that is however not equivalent to the above is:

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} \,\mathrm{d}x$$
$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{f}(k) e^{+ikx} \,\mathrm{d}k$$

The 3D Fourier transform is:

$$\begin{aligned} F[f(\mathbf{x})] \equiv \tilde{f}(\boldsymbol{\omega}) &= \int_{-\infty}^{\infty} f(\mathbf{x}) e^{-i\boldsymbol{\omega}\cdot\mathbf{x}} \,\mathrm{d}^3x \\ F^{-1}[\tilde{f}(\boldsymbol{\omega})] = f(\mathbf{x}) &= \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} \tilde{f}(\boldsymbol{\omega}) e^{+i\boldsymbol{\omega}\cdot\mathbf{x}} \,\mathrm{d}^3\omega \end{aligned} \tag{3.4.1}$$

With obvious analogs for other conventions and dimensions.

The sign convention in the exponentials $e^{\pm i\omega x}$ is arbitrary, one can as well flip the sign of the direct and inverse transforms. In particular, one often uses both sign conventions in the same equation. Consider a spacetime plane-wave $e^{ik\cdot x} = e^{i(\omega t - \mathbf{k}\cdot\mathbf{x})}$. Then we obtain (using plus sign convention in the $e^{ik\cdot x}$ exponential for the direct transformation):

$$F[f(x)] \equiv \tilde{f}(k) = \int_{-\infty}^{\infty} f(x) e^{ik\cdot x} \,\mathrm{d}^4x = \int_{-\infty}^{\infty} f(x) e^{i(\omega t - \mathbf{k}\cdot\mathbf{x})} \,\mathrm{d}^4x$$
$$F^{-1}[f(k)] \equiv f(x) = \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} \tilde{f}(k) e^{-ik\cdot x} \,\mathrm{d}^4k = \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} \tilde{f}(k) e^{-i(\omega t - \mathbf{k}\cdot\mathbf{x})} \,\mathrm{d}^4k$$

Finally, the equation $k \cdot x = \omega t - \mathbf{k} \cdot \mathbf{x}$ depends on the metric signature, in this case $\mathrm{diag}(1, -1, -1, -1)$. For a signature $\mathrm{diag}(-1, 1, 1, 1)$ we would get $k \cdot x = -\omega t + \mathbf{k} \cdot \mathbf{x}$.

Unlike the normalization convention, where one has to be very careful, the sign convention in Fourier transform is not a problem, one just has to remember to flip the sign for the inverse transform.

3.4.1 Shift Theorem

The Fourier transform of a shifted function, in 3D:

$$\begin{aligned} F[f(\mathbf{x}+\mathbf{b})] &= \int_{-\infty}^{\infty} f(\mathbf{x}+\mathbf{b}) e^{-i\boldsymbol{\omega}\cdot\mathbf{x}} \,\mathrm{d}^3x = \\ &= \int_{-\infty}^{\infty} f(\mathbf{x}) e^{-i\boldsymbol{\omega}\cdot(\mathbf{x}-\mathbf{b})} \,\mathrm{d}^3x = \\ &= e^{i\boldsymbol{\omega}\cdot\mathbf{b}} \int_{-\infty}^{\infty} f(\mathbf{x}) e^{-i\boldsymbol{\omega}\cdot\mathbf{x}} \,\mathrm{d}^3x = \\ &= e^{i\boldsymbol{\omega}\cdot\mathbf{b}} F[f(\mathbf{x})] \end{aligned}$$

3.4.2 Scaling

For $a > 0$:

$$\begin{aligned} F[f(ax)](\omega) &= \int_{-\infty}^{\infty} f(ax)e^{-i\omega x}\,\mathrm{d}x = \\ &= \frac{1}{a}\int_{-\infty}^{\infty} f(y)e^{-i\frac{\omega}{a}y}\,\mathrm{d}y = \\ &= \frac{1}{a}F[f(x)]\left(\frac{\omega}{a}\right) \end{aligned}$$

3.4.3 Derivative

The Fourier transform of a derivative, in 3D:

$$\begin{aligned} F[\partial_i f(\mathbf{x})] &= \int_{-\infty}^{\infty} (\partial_i f(\mathbf{x}))e^{-i\boldsymbol{\omega}\cdot\mathbf{x}}\,\mathrm{d}^3x = \\ &= \left[f(\mathbf{x})e^{-i\boldsymbol{\omega}\cdot\mathbf{x}}\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f(\mathbf{x})\partial_i e^{-i\boldsymbol{\omega}\cdot\mathbf{x}}\,\mathrm{d}^3x = \\ &= -\int_{-\infty}^{\infty} f(\mathbf{x})\partial_i e^{-i\omega_j x^j}\,\mathrm{d}^3x = \\ &= -(-i\omega_i)\int_{-\infty}^{\infty} f(\mathbf{x})e^{-i\boldsymbol{\omega}\cdot\mathbf{x}}\,\mathrm{d}^3x = \\ &= i\omega_i F[f(\mathbf{x})]\,. \end{aligned}$$

An alternative derivation is to start from:

$$f(\mathbf{x}) = F^{-1}[\tilde{f}(\boldsymbol{\omega})] = \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty} \tilde{f}(\boldsymbol{\omega})e^{+i\boldsymbol{\omega}\cdot\mathbf{x}}\,\mathrm{d}^3\omega$$

and differentiate both sides:

$$\begin{aligned} \partial_i f(\mathbf{x}) &= \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty} \tilde{f}(\boldsymbol{\omega})\partial_i e^{+i\boldsymbol{\omega}\cdot\mathbf{x}}\,\mathrm{d}^3\omega \\ \partial_i f(\mathbf{x}) &= \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty} i\omega_i\tilde{f}(\boldsymbol{\omega})e^{+i\boldsymbol{\omega}\cdot\mathbf{x}}\,\mathrm{d}^3\omega\,, \end{aligned}$$

from which:

$$F[\partial_i f(\mathbf{x})] = i\omega_i\tilde{f}(\boldsymbol{\omega}) = i\omega_i F[f(\mathbf{x})]\,.$$

3.4.4 Convolution

The convolution of two functions $f(x)$ and $g(x)$ is defined as:

$$f(x) * g(x) = \int_{-\infty}^{\infty} f(y)g(x-y)\,\mathrm{d}y$$

The Fourier transform of a convolution is:

$$\begin{aligned}
F[f(x) * g(x)](\omega) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(y)g(x-y)\,\mathrm{d}y\, e^{-i\omega x}\mathrm{d}x = \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x-y)e^{-i\omega x}\mathrm{d}x\, f(y)\,\mathrm{d}y = \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(u)e^{-i\omega(u+y)}\mathrm{d}u\, f(y)\,\mathrm{d}y = \\
&= \int_{-\infty}^{\infty} g(u)e^{-i\omega u}\mathrm{d}u \int_{-\infty}^{\infty} f(y)e^{-i\omega y}\mathrm{d}y \\
&= F[f(x)](\omega)\; F[g(x)](\omega)
\end{aligned}$$

And for the inverse transform:

$$\begin{aligned}
F^{-1}[f(\omega) * g(\omega)](x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(y)g(\omega-y)\,\mathrm{d}y\, e^{i\omega x}\mathrm{d}\omega = \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(\omega-y)e^{i\omega x}\mathrm{d}\omega\, f(y)\,\mathrm{d}y = \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(u)e^{ix(u+y)}\mathrm{d}u\, f(y)\,\mathrm{d}y = \\
&= 2\pi\frac{1}{2\pi}\int_{-\infty}^{\infty} g(u)e^{ixu}\mathrm{d}u\frac{1}{2\pi}\int_{-\infty}^{\infty} f(y)e^{ixy}\mathrm{d}y \\
&= 2\pi F^{-1}[f(\omega)](x)\; F^{-1}[g(\omega)](x)
\end{aligned}$$

Fourier transform of a function multiplication is:

$$F[fg] = F[\, F^{-1}[F[f]] \quad F^{-1}[F[g]]\,] = \frac{1}{2\pi}F[F^{-1}[F[f] * F[g]]] = \frac{1}{2\pi}F[f] * F[g]$$

and for the inverse transform:

$$F^{-1}[fg] = F^{-1}[\, F[F^{-1}[f]] \quad F[F^{-1}[g]]\,] = F^{-1}[F[F^{-1}[f] * F^{-1}[g]]] = F^{-1}[f] * F^{-1}[g]$$

3.4.5 Radial Fourier Transform

As a special case when the function $f(\mathbf{x}) = f(r)$ is spherically symmetric, we introduce spherical coordinates such that the z-axis is along the $\boldsymbol{\omega}$ vector and calculate (we use $r = |\mathbf{x}|$ and $\omega = |\boldsymbol{\omega}|$):

$$\begin{aligned}
F[f(\mathbf{x})] \equiv \tilde{f}(\boldsymbol{\omega}) &= \int_{-\infty}^{\infty} f(\mathbf{x})e^{-i\boldsymbol{\omega}\cdot\mathbf{x}}\,\mathrm{d}^3x = \int_{-\infty}^{\infty} f(r)e^{-i\boldsymbol{\omega}\cdot\mathbf{x}}\,\mathrm{d}^3x = \\
&= \int_0^{\infty}\mathrm{d}r\int_0^{\pi}\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi f(r)e^{-i\omega r\cos\theta}\,r^2\sin\theta = \\
&= 2\pi\int_0^{\infty}\mathrm{d}r\int_0^{\pi}\mathrm{d}\theta f(r)e^{-i\omega r\cos\theta}\,r^2\sin\theta = \\
&= 4\pi\int_0^{\infty} f(r)\,\mathrm{sinc}(\omega r)\,r^2\mathrm{d}r = \\
&= 4\pi\int_0^{\infty} f(r)\frac{\sin\omega r}{\omega r}\,r^2\mathrm{d}r = \\
&= \frac{4\pi}{\omega}\int_0^{\infty} r\sin(\omega r)f(r)\,\mathrm{d}r\,,
\end{aligned}$$

where we used:

$$\int_0^\pi e^{-i\omega r\cos\theta}\,\sin\theta \mathrm{d}\theta = \int_{-1}^1 e^{i\omega r u}\mathrm{d}u = \left[\frac{e^{i\omega r u}}{i\omega r}\right]_{-1}^1 = \frac{e^{i\omega r}-e^{-i\omega r}}{i\omega r} = \\ = 2\frac{\sin(\omega r)}{\omega r} = 2\,\mathrm{sinc}(\omega r) = 2j_0(\omega r)\,.$$

So the transform is real and spherically symmetric, since the result only depends on ω.

Similarly, for the inverse transform:

$$F^{-1}[\tilde{f}(\boldsymbol{\omega})] = f(\mathbf{x}) = \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}\tilde{f}(\boldsymbol{\omega})e^{+i\boldsymbol{\omega}\cdot\mathbf{x}}\,\mathrm{d}^3\omega = \\ = \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}\tilde{f}(\omega)e^{+i\boldsymbol{\omega}\cdot\mathbf{x}}\,\mathrm{d}^3\omega = \\ = \frac{1}{(2\pi)^3}\frac{4\pi}{r}\int_0^{\infty}\omega\sin(\omega r)f(\omega)\,\mathrm{d}\omega = \\ = \frac{1}{2\pi^2 r}\int_0^{\infty}\omega\sin(\omega r)f(\omega)\,\mathrm{d}\omega$$

3.4.6 Examples

Rectangular Function

The rectangular function is defined as:

$$\Pi(x) = H(x+\tfrac{1}{2}) - H(x-\tfrac{1}{2})$$

The Fourier transform is:

$$F[\Pi(x)] \equiv \tilde{\Pi}(\omega) = \int_{-\infty}^{\infty}\Pi(x)e^{-i\omega x}\,\mathrm{d}x = \int_{-\frac{1}{2}}^{\frac{1}{2}}e^{-i\omega x}\,\mathrm{d}x = \\ = \left[\frac{e^{-i\omega x}}{-i\omega x}\right]_{x=-\frac{1}{2}}^{\frac{1}{2}} = \frac{e^{i\frac{\omega}{2}}-e^{-i\frac{\omega}{2}}}{i\frac{\omega}{2}} = 2\frac{\sin(\frac{\omega}{2})}{\frac{\omega}{2}} = 2\,\mathrm{sinc}\left(\frac{\omega}{2}\right)\,.$$

Dirichlet Kernel

The Dirichlet kernel $D_N(x)$ is a partial sum of complex exponentials:

$$D_N(x) = \frac{1}{2\pi}\sum_{n=-N}^{N}e^{inx} = \\ = \frac{1}{2\pi}\left(1+2\sum_{n=1}^{N}\cos(nx)\right) = \\ = \frac{1}{2\pi\sin\left(\frac{x}{2}\right)}\left(\sin\left(\frac{x}{2}\right)+2\sum_{n=1}^{N}\cos(nx)\sin\left(\frac{x}{2}\right)\right) = \\ = \frac{1}{2\pi\sin\left(\frac{x}{2}\right)}\left(\sin\left(\frac{x}{2}\right)+\sum_{n=1}^{N}\left(\sin\left(\left(n+\tfrac{1}{2}\right)x\right)-\sin\left(\left(n-\tfrac{1}{2}\right)x\right)\right)\right) = \\ = \frac{\sin\left(\left(N+\frac{1}{2}\right)x\right)}{2\pi\sin\left(\frac{x}{2}\right)}$$

From the definition, it is a periodic function with period 2π.

Integral of it is equal to one:

$$\int_{-\pi}^{\pi} D_N(x)\mathrm{d}x = \int_{-\pi}^{\pi} \frac{1}{2\pi}\left(1 + 2\sum_{n=1}^{N}\cos(nx)\right)\mathrm{d}x = \\ = 1 + \frac{1}{\pi}\sum_{n=1}^{N}\int_{-\pi}^{\pi}\cos(nx)\,\mathrm{d}x = 1$$

also

$$\int_{-\pi}^{\pi} D_N(x-y)\mathrm{d}y = 1$$

The Dirichlet kernel $D_N(x)$ converges towards a train of delta functions (called Dirac comb, see the equation (3.4.6.2) in the next section):

$$\frac{1}{2\pi}\sum_{n=-\infty}^{\infty} e^{inx} = \lim_{N\to\infty}\frac{1}{2\pi}\sum_{n=-N}^{N} e^{inx} = \lim_{N\to\infty} D_N(x) = \\ = \lim_{N\to\infty}\frac{\sin\left(\left(N+\frac{1}{2}\right)x\right)}{2\pi\sin\left(\frac{x}{2}\right)} = \sum_{n=-\infty}^{\infty}\delta(x-2\pi n) \tag{3.4.6.1}$$

Let us do the crucial step in more details using distributions:

$$\int_{-\infty}^{\infty}\lim_{N\to\infty}\frac{\sin\left(\left(N+\frac{1}{2}\right)x\right)}{2\pi\sin\left(\frac{x}{2}\right)}\varphi(x)\,\mathrm{d}x = \\ = \sum_{n=-\infty}^{\infty}\lim_{N\to\infty}\int_{-\pi}^{\pi}\frac{\sin\left(\left(N+\frac{1}{2}\right)(x+2\pi n)\right)}{2\pi\sin\left(\frac{x+2\pi n}{2}\right)}\varphi(x+2\pi n)\,\mathrm{d}x = \\ = \sum_{n=-\infty}^{\infty}\varphi(2\pi n) = \\ = \int_{-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\delta(x-2\pi n)\varphi(x)\,\mathrm{d}x$$

Where we used the fact that

$$\left[\lim_{N\to\infty}\int_{-\pi}^{\pi}\frac{\sin\left(\left(N+\frac{1}{2}\right)(x+2\pi n)\right)}{2\pi\sin\left(\frac{x+2\pi n}{2}\right)}\varphi(x+2\pi n)\,\mathrm{d}x\right] - \varphi(2\pi n) = \\ = \left[\lim_{N\to\infty}\int_{-\pi}^{\pi} D_N(x+2\pi n)\varphi(x+2\pi n)\,\mathrm{d}x\right] - \varphi(2\pi n) = \\ = \lim_{N\to\infty}\int_{-\pi}^{\pi} D_N(x+2\pi n)\left(\varphi(x+2\pi n) - \varphi(2\pi n)\right)\,\mathrm{d}x = \\ = \lim_{N\to\infty}\int_{-\pi}^{\pi}\frac{\varphi(x+2\pi n)-\varphi(2\pi n)}{2\pi\sin\left(\frac{x+2\pi n}{2}\right)}\sin\left(\left(N+\frac{1}{2}\right)(x+2\pi n)\right)\,\mathrm{d}x = \\ = 0$$

Dirac Comb (Shah) Function

The Dirac comb function, also called the Shah function, is defined as:

$$\mathrm{Ш}(x) = \sum_{n=-\infty}^{\infty}\delta(x-n)$$

It has the following scaling property:

$$\text{Ш}(ax) = \sum_{n=-\infty}^{\infty} \delta(ax - n) = \sum_{n=-\infty}^{\infty} \delta\left(a\left(x - \frac{n}{a}\right)\right) = \sum_{n=-\infty}^{\infty} \frac{1}{|a|}\delta\left(x - \frac{n}{a}\right)$$

and for $a = \frac{1}{L}$ with $L > 0$:

$$\text{Ш}\left(\frac{x}{L}\right) = \sum_{n=-\infty}^{\infty} L\delta(x - nL)$$

From which a train of delta functions L distance apart is expressed using a Dirac comb as:

$$\sum_{n=-\infty}^{\infty} \delta(x - nL) = \frac{1}{L}\text{Ш}\left(\frac{x}{L}\right)$$

Using the identity (3.4.6.1), the infinite sum of complex exponentials is also equal to a Dirac comb:

$$\begin{aligned} \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} e^{inx} = \lim_{N\to\infty}\frac{1}{2\pi}\sum_{n=-N}^{N} e^{inx} = \lim_{N\to\infty} D_N(x) = \\ = \lim_{N\to\infty}\frac{\sin\left(\left(N+\frac{1}{2}\right)x\right)}{2\pi\sin\left(\frac{x}{2}\right)} = \sum_{n=-\infty}^{\infty}\delta(x - 2\pi n) = \frac{1}{2\pi}\text{Ш}\left(\frac{x}{2\pi}\right) \end{aligned} \tag{3.4.6.2}$$

Using (3.4.6.2) we can now calculate the Fourier transform:

$$\begin{aligned} F[\text{Ш}(x)](\omega) \equiv \tilde{\text{Ш}}(\omega) = \int_{-\infty}^{\infty} \text{Ш}(x)e^{-i\omega x}\,\mathrm{d}x = \\ = \int_{-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\delta(x-n)e^{-i\omega x}\,\mathrm{d}x = \\ = \sum_{n=-\infty}^{\infty}\int_{-\infty}^{\infty}\delta(x-n)e^{-i\omega x}\,\mathrm{d}x = \\ = \sum_{n=-\infty}^{\infty} e^{-i\omega n} = \\ = 2\pi\sum_{n=-\infty}^{\infty}\delta(\omega - 2\pi n) = \\ = \text{Ш}\left(\frac{\omega}{2\pi}\right) \end{aligned}$$

For the inverse Fourier transform we get (using the previous result):

$$F^{-1}[\text{Ш}(\omega)](x) = F^{-1}\left[\text{Ш}\left(\frac{(2\pi\omega)}{2\pi}\right)\right](x) = F^{-1}[\,F[\text{Ш}(x)](2\pi\omega)\,](x) = F^{-1}\left[F\left[\frac{1}{2\pi}\text{Ш}\left(\frac{x}{2\pi}\right)\right](\omega)\right](x) = \frac{1}{2\pi}\text{Ш}\left(\frac{x}{2\pi}\right)$$

The following Fourier transform is also useful:

$$\begin{aligned} F\left[\sum_{n=-\infty}^{\infty}\delta(x-nL)\right](\omega) = F\left[\frac{1}{L}\text{Ш}\left(\frac{x}{L}\right)\right](\omega) = F[\text{Ш}(x)](L\omega) = \\ = \text{Ш}\left(\frac{L}{2\pi}\omega\right) = \sum_{n=-\infty}^{\infty}\frac{2\pi}{L}\delta\left(x - \frac{2\pi n}{L}\right) \end{aligned}$$

Periodic Summation

The convolution $f(x) * g(x) = \int_{-\infty}^{\infty} f(y)g(x-y)\,\mathrm{d}y$ of a Dirac comb $\mathrm{Ш}(x)$ and an arbitrary function $f(x)$ is called a periodic summation:

$$f(x) * \mathrm{Ш}(x) = \int_{-\infty}^{\infty} f(y)\mathrm{Ш}(x-y)\,\mathrm{d}y = \int_{-\infty}^{\infty} f(y) \sum_{n=-\infty}^{\infty} \delta(x-y-n)\,\mathrm{d}y =$$
$$= \sum_{n=-\infty}^{\infty} f(x-n) = \sum_{n=-\infty}^{\infty} f(x+n)$$

because the result is a periodic function with period 1:

$$(f * \mathrm{Ш})(x+1) = \sum_{n=-\infty}^{\infty} f(x+n+1) = \sum_{m=-\infty}^{\infty} f(x+m) = (f * \mathrm{Ш})(x)$$

Poisson Summation Formula

The Poisson summation formula:

$$\sum_{n=-\infty}^{\infty} f(2\pi n) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \tilde{f}(n) \tag{3.4.6.3}$$

can be derived using a Dirac comb:

$$\sum_{n=-\infty}^{\infty} f(2\pi n) = \int_{-\infty}^{\infty} f(x) \sum_{n=-\infty}^{\infty} \delta(x-2\pi n)\,\mathrm{d}x =$$
$$= \int_{-\infty}^{\infty} f(x) \frac{1}{2\pi} \mathrm{Ш}\left(\frac{x}{2\pi}\right)\,\mathrm{d}x =$$
$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) \cdot F[\mathrm{Ш}(\omega)](x)\,\mathrm{d}x =$$
$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} F[f(x)](\omega) \cdot \mathrm{Ш}(\omega)\,\mathrm{d}\omega =$$
$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{f}(\omega) \cdot \sum_{n=-\infty}^{\infty} \delta(\omega-n)\,\mathrm{d}\omega =$$
$$= \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \tilde{f}(x)$$

An alternative derivation using Fourier series (see next sections):

$$\sum_{n=-\infty}^{\infty} f(x+2\pi n) = g(x) = \sum_{n=-\infty}^{\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} g(y)e^{-iny}\mathrm{d}y\, e^{inx} =$$
$$= \sum_{n=-\infty}^{\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{m=-\infty}^{\infty} f(y+2\pi m)e^{-iny}\mathrm{d}y\, e^{inx} =$$
$$= \sum_{n=-\infty}^{\infty} \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \int_{-\pi}^{\pi} f(y+2\pi m)e^{-in(y+2\pi m)}\mathrm{d}y\, e^{inx} =$$
$$= \sum_{n=-\infty}^{\infty} \frac{1}{2\pi} \int_{-\infty}^{\infty} f(y)e^{-iny}\mathrm{d}y\, e^{inx} =$$
$$= \sum_{n=-\infty}^{\infty} \frac{1}{2\pi} \tilde{f}(n)\, e^{inx}$$

And setting $x = 0$ we get the Poisson summation formula (3.4.6.3).

The last derivation can actually also be done using a Dirac comb function as follows:

$$\begin{aligned}
\sum_{n=-\infty}^{\infty} f(x+2\pi n) = f(x) * \frac{1}{2\pi}\text{Ш}\left(\frac{x}{2\pi}\right) = \\
= F^{-1}\left[\, F\left[f(x) * \frac{1}{2\pi}\text{Ш}\left(\frac{x}{2\pi}\right)\right](\omega)\,\right](x) = \\
= F^{-1}\left[\, F[f(x)](\omega)\; F\left[\frac{1}{2\pi}\text{Ш}\left(\frac{x}{2\pi}\right)\right](\omega)\,\right](x) = \\
= F^{-1}[\, F[f(x)](\omega)\; \text{Ш}(\omega)\,](x) = \\
= \frac{1}{2\pi}\int_{-\infty}^{\infty} F[f(x)](\omega)\text{Ш}(\omega)e^{i\omega x}\mathrm{d}x = \\
= \sum_{n=-\infty}^{\infty} \frac{1}{2\pi} F[f(x)](n)\, e^{inx}
\end{aligned}$$

Fourier Series

Consider a periodic function $f(x)$ with a period L and let us calculate the Fourier transform of it. We define a new function $f_0(x) = f(x)$ in the $[0, L]$ interval and zero otherwise. Then:

$$f(x) = f_0(x) * \frac{1}{L}\text{Ш}\left(\frac{x}{L}\right) = \sum_{n=-\infty}^{\infty} f_0(x+Ln)$$

Apply Fourier transform:

$$\begin{aligned}
F[f(x)](\omega) = F\left[f_0(x) * \frac{1}{L}\text{Ш}\left(\frac{x}{L}\right)\right](\omega) = \\
= F[f_0(x)](\omega)\; F\left[\frac{1}{L}\text{Ш}\left(\frac{x}{L}\right)\right](\omega) = F[f_0(x)](\omega)\; \text{Ш}\left(\frac{L}{2\pi}\omega\right) = \\
= \sum_{n=-\infty}^{\infty} F[f_0(x)]\left(\frac{2\pi n}{L}\right)\frac{2\pi}{L}\delta\left(\omega - \frac{2\pi n}{L}\right) = \\
= \sum_{n=-\infty}^{\infty}\int_0^L f(x)e^{-i2\pi nx/L}\mathrm{d}x\,\frac{2\pi}{L}\delta\left(\omega - \frac{2\pi n}{L}\right) = \\
= 2\pi\sum_{n=-\infty}^{\infty} f_n\delta\left(\omega - \frac{2\pi n}{L}\right)
\end{aligned} \tag{3.4.6.4}$$

where f_n are called Fourier coefficients:

$$f_n = \frac{1}{L}\int_0^L f(x)e^{-i2\pi nx/L}\mathrm{d}x$$

We can see that the Fourier transform is zero for $\omega \neq \frac{2\pi n}{L}$. For $\omega = \frac{2\pi n}{L}$ it is equal to a delta function times a 2π multiple of a Fourier series coefficient. The delta functions structure is given by the period L of the function $f(x)$. All the information that is stored in the answer is inside the f_n coefficients, so those are the only ones that we need to calculate and store.

The function $f(x)$ is calculated from the f_n coefficients by applying the inverse Fourier transform to the final result of

(3.4.6.4) as follows:

$$\begin{aligned} f(x) = F^{-1}[F[f(x)](\omega)](x) = \\ = F^{-1}\left[2\pi \sum_{n=-\infty}^{\infty} f_n \delta\left(\omega - \frac{2\pi n}{L}\right)\right](x) = \\ = \frac{1}{2\pi}\int_{-\infty}^{\infty} 2\pi \sum_{n=-\infty}^{\infty} f_n \delta\left(\omega - \frac{2\pi n}{L}\right) e^{i\omega x} \mathrm{d}\omega = \\ = \sum_{n=-\infty}^{\infty} f_n e^{i2\pi n x/L} \end{aligned} \tag{3.4.6.5}$$

The expansion (3.4.6.5) is called a Fourier series. It is given by the Fourier coefficients f_n. The equation (3.4.6.4) provides the relation between a Fourier transform and a Fourier series.

For example for $f(x) = \sin(x)$, the only nonzero Fourier coefficients for $L = 2\pi$ are $f_{-1} = \frac{i}{2}$ and $f_1 = -\frac{i}{2}$. The Fourier transform then is:

$$\begin{aligned} F[\sin(x)](\omega) = 2\pi \left(f_{-1}\delta(\omega - (-1)) + f_1 \delta(\omega - 1)\right) = \\ = 2\pi\left(\frac{i}{2}\delta(\omega + 1)) - \frac{i}{2}\delta(\omega - 1)\right) = i\pi\delta(\omega + 1) - i\pi\delta(\omega - 1) \end{aligned}$$

For $f(x) = 1$ the only nonzero Fourier coefficient is $f_0 = 1$, the Fourier transform then is:

$$F[1](\omega) = 2\pi f_0 \delta(\omega - 0) = 2\pi\delta(\omega)$$

For $f(x) = e^{3ix}$ the only nonzero Fourier coefficient for $L = 2\pi$ is $f_3 = 1$, the Fourier transform then is:

$$F[e^{3ix}](\omega) = 2\pi f_3 \delta(\omega - 3) = 2\pi\delta(\omega - 3)$$

For $f(x) = \sum_{n=-\infty}^{\infty} \delta(x - 2\pi n)$ the Fourier coefficients for $L = 2\pi$ are all equal to $f_n = \frac{1}{2\pi}$ and the Fourier transform is:

$$F[f(x)](\omega) = 2\pi \sum_{n=-\infty}^{\infty} f_n \delta(\omega - n) = \sum_{n=-\infty}^{\infty} \delta(\omega - n)$$

Note: if we start from (3.4.6.5), for simplicity on an interval $[-\pi, \pi]$:

$$f(x) = \sum_{n=-\infty}^{\infty} f_n e^{inx} \tag{3.4.6.6}$$

To calculate the Fourier coefficients f_n, we can just multiply both sides of (3.4.6.6) by e^{-imx} and integrate:

$$\begin{aligned} \int_{-\pi}^{\pi} f(x) e^{-imx} \mathrm{d}x = \int_{-\pi}^{\pi} \sum_{n=-\infty}^{\infty} f_n e^{inx} e^{-imx} \mathrm{d}x = \\ = \sum_{n=-\infty}^{\infty} f_n \int_{-\pi}^{\pi} e^{i(n-m)x} \mathrm{d}x = \\ = \sum_{n=-\infty}^{\infty} f_n 2\pi\delta_{nm} = \\ = 2\pi f_m \,, \end{aligned}$$

so

$$f_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x) e^{-inx} \mathrm{d}x \tag{3.4.6.7}$$

Convergence of Fourier Series

To see what conditions the function $f(x)$ must satisfy in order for the Fourier series to converge towards it, we can do the following analysis. Substituting (3.4.6.7) into (3.4.6.6) yields:

$$\begin{aligned} f(x) &= \sum_{n=-\infty}^{\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) e^{-iny} \mathrm{d}y \, e^{inx} = \\ &= \lim_{N\to\infty} \int_{-\pi}^{\pi} \frac{1}{2\pi} \sum_{n=-N}^{N} e^{in(x-y)} f(y) \mathrm{d}y = \\ &= \lim_{N\to\infty} \int_{-\pi}^{\pi} D_N(x-y) f(y) \mathrm{d}y \end{aligned}$$

We can now calculate the difference between the Fourier series and the function value:

$$\begin{aligned} &\lim_{N\to\infty} \int_{-\pi}^{\pi} D_N(x-y) f(y) \mathrm{d}y - f(x) = \\ &= \lim_{N\to\infty} \int_{-\pi}^{\pi} D_N(x-y) \left(f(y) - f(x)\right) \mathrm{d}y = \\ &= \lim_{N\to\infty} \int_{-\pi}^{\pi} \frac{\sin\left(\left(N+\frac{1}{2}\right)(x-y)\right)}{2\pi \sin\left(\frac{x-y}{2}\right)} \left(f(y) - f(x)\right) \mathrm{d}y = \\ &= \lim_{N\to\infty} \int_{-\pi}^{\pi} \frac{f(y) - f(x)}{2\pi \sin\left(\frac{x-y}{2}\right)} \sin\left(\left(N+\frac{1}{2}\right)(x-y)\right) \mathrm{d}y = \\ &= \lim_{N\to\infty} \int_{x-\pi}^{x+\pi} \frac{f(x-u) - f(x)}{2\pi \sin\left(\frac{u}{2}\right)} \sin\left(\left(N+\frac{1}{2}\right)u\right) \mathrm{d}u = \\ &= \lim_{N\to\infty} \int_{x-\pi}^{x+\pi} h(u) \sin\left(\left(N+\frac{1}{2}\right)u\right) \mathrm{d}u = 0 \end{aligned}$$

where $h(u)$ is finite and well behaved at the origin $u = 0$:

$$h(u) = \frac{f(x-u) - f(x)}{2\pi \sin\left(\frac{u}{2}\right)} = -\frac{f'(x)}{\pi} + \frac{f''(x)}{2\pi} u + O(u^2)$$

The integral is zero because the more and more oscillating $\sin$ function cancels the contributions of positive and negative parts of the integrand. This can be proven explicitly as follows using the fact that $h(x)$, $h'(x)$ and $\cos(Nx)$ is bounded as $N \to \infty$:

$$\begin{aligned} &\lim_{N\to\infty} \int_a^b h(x) \sin(Nx) \, \mathrm{d}x = \\ &= \lim_{N\to\infty} \frac{1}{N} \left(\left[-h(x)\cos(Nx)\right]_a^b + \int_a^b h'(x) \cos(Nx) \, \mathrm{d}x \right) = 0 \end{aligned}$$

The conditions that we used are that the function $h(u)$ can be integrated, which is satisfied if e.g. $f(x)$ has derivatives. These conditions can be loosened in various ways.

3.5 Fourier Transform of a Periodic Function (e.g. in a Crystal)

The Fourier transform in (3.4.1) requires the function $f(\mathbf{x})$ to be decaying fast enough in order to converge. In an infinite crystal, on the other hand, the function $f(\mathbf{x})$ is typically periodic (and thus not decaying):

$$f(\mathbf{x} + \mathbf{T}(n_1, n_2, n_3)) = f(\mathbf{x})$$

where $\mathbf{T}(\mathbf{n}) = \mathbf{T}(n_1, n_2, n_3) = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$ are the crystal translation vectors. As such, the Fourier transform in (3.4.1) is infinite, but it can be made finite by the following definition:

$$\begin{aligned} F[f(\mathbf{x})] \equiv \tilde{f}(\boldsymbol{\omega}) &= \frac{1}{\Omega_{\text{crystal}}} \int_{\Omega_{\text{crystal}}} f(\mathbf{x}) e^{-i\boldsymbol{\omega}\cdot\mathbf{x}} \,\mathrm{d}^3x = \\ &= \frac{1}{\Omega_{\text{crystal}}} \sum_{\mathbf{n}} \int_{\Omega_{\text{cell}}} f(\mathbf{x}+\mathbf{T}(\mathbf{n})) e^{-i\boldsymbol{\omega}\cdot(\mathbf{x}+\mathbf{T}(\mathbf{n}))} \,\mathrm{d}^3x = \\ &= \frac{1}{\Omega_{\text{crystal}}} \sum_{\mathbf{n}} \int_{\Omega_{\text{cell}}} f(\mathbf{x}) e^{-i\boldsymbol{\omega}\cdot(\mathbf{x}+\mathbf{T}(\mathbf{n}))} \,\mathrm{d}^3x = \\ &= \frac{1}{\Omega_{\text{crystal}}} \sum_{\mathbf{n}} e^{-i\boldsymbol{\omega}\cdot\mathbf{T}(\mathbf{n})} \int_{\Omega_{\text{cell}}} f(\mathbf{x}) e^{-i\boldsymbol{\omega}\cdot\mathbf{x}} \,\mathrm{d}^3x = \\ &= \frac{1}{\Omega_{\text{crystal}}} N_{\text{cell}} \int_{\Omega_{\text{cell}}} f(\mathbf{x}) e^{-i\boldsymbol{\omega}\cdot\mathbf{x}} \,\mathrm{d}^3x = \\ &= \frac{1}{\Omega_{\text{cell}}} \int_{\Omega_{\text{cell}}} f(\mathbf{x}) e^{-i\boldsymbol{\omega}\cdot\mathbf{x}} \,\mathrm{d}^3x \end{aligned} \tag{3.5.1}$$

This assumes that the wave vector $\boldsymbol{\omega} = \mathbf{G}$ is equal to the reciprocal space vectors $\mathbf{G}$, defined by

$$e^{i\mathbf{G}\cdot\mathbf{T}(\mathbf{n})} = 1\,, \tag{3.5.2}$$

because then $\sum_{\mathbf{n}} e^{-i\boldsymbol{\omega}\cdot\mathbf{T}(\mathbf{n})} = \sum_{\mathbf{n}} 1 = N_{\text{cell}}$.

For $\boldsymbol{\omega} \neq \mathbf{G}$, the expression $\frac{1}{\Omega_{\text{crystal}}} \sum_{\mathbf{n}} e^{-i\boldsymbol{\omega}\cdot\mathbf{T}(\mathbf{n})} = 0$ vanishes, because the sum is bounded, and so dividing by the (infinite) crystal volume makes the expression vanish, and so $\tilde{f}(\boldsymbol{\omega}) = 0$. In other words, the only non-zero Fourier components $\tilde{f}(\boldsymbol{\omega})$ of any periodic function $f(\mathbf{x})$ are those with $\boldsymbol{\omega} = \mathbf{G}$. Equivalently said, if the Fourier components of a given function are non-zero for some $\boldsymbol{\omega} \neq \mathbf{G}$, then the function is not periodic.

Summary: the only difference between the crystal Fourier transform (3.5.1) and the usual Fourier transform (3.4.1) is the Ω_{crystal} factor. The Fourier transform (3.5.1) of a periodic function is nonzero only for $\omega = \mathbf{G}$ and is equal to:

$$F[f(\mathbf{x})] \equiv \tilde{f}(\mathbf{G}) = \frac{1}{\Omega_{\text{cell}}} \int_{\Omega_{\text{cell}}} f(\mathbf{x}) e^{-i\mathbf{G}\cdot\mathbf{x}} \,\mathrm{d}^3x \tag{3.5.3}$$

Note: the fact that the sum is bounded follows from:

$$\begin{aligned} \left| \sum_{n=-\infty}^{\infty} e^{ikn} \right| = \left| \lim_{N\to\infty} \sum_{n=-N}^{N} e^{ikn} \right| = \left| \lim_{N\to\infty} \left(1 + 2\sum_{n=1}^{N} \cos kn \right) \right| = \\ = \left| \lim_{N\to\infty} \frac{\cos kN - \cos k(N+1)}{1-\cos k} \right| < \frac{2}{|1-\cos k|} \end{aligned}$$

Because $|\cos kN - \cos k(N+1)| < 2$. So for $k \neq 2\pi$ (i.e. the denominator is non-zero), the sum is bounded (to be precise, the infinite sum does not converge, because it oscillates, but the point is that the partial sum is always bounded). For $k = 2\pi$, the sum is infinite, because $e^{i2\pi n} = 1$.

Since we divided the direct Fourier transform in (3.4.1) by Ω_{crystal} to obtain (3.5.1), we need to multiply the inverse

transform in (3.4.1) by Ω_{crystal}:

$$\begin{aligned} F^{-1}[\tilde{f}(\boldsymbol{\omega})] = f(\mathbf{x}) &= \frac{\Omega_{\text{crystal}}}{(2\pi)^3}\int_{-\infty}^{\infty}\tilde{f}(\boldsymbol{\omega})e^{+i\boldsymbol{\omega}\cdot\mathbf{x}}\,\mathrm{d}^3\omega = \\ &= \frac{\Omega_{\text{cell}}N_{\text{cell}}}{(2\pi)^3}\int_{-\infty}^{\infty}\tilde{f}(\boldsymbol{\omega})e^{+i\boldsymbol{\omega}\cdot\mathbf{x}}\,\mathrm{d}^3\omega = \\ &= \frac{N_{\text{cell}}}{\Omega_{\text{BZ}}}\sum_{\mathbf{G}}\int_{\Omega_{\text{BZ}}}\tilde{f}(\mathbf{G}+\boldsymbol{\omega})e^{+i(\mathbf{G}+\boldsymbol{\omega})\cdot\mathbf{x}}\,\mathrm{d}^3\omega = \\ &= \frac{N_{\text{cell}}}{\Omega_{\text{BZ}}}\sum_{\mathbf{G}}e^{+i\mathbf{G}\cdot\mathbf{x}}\int_{\Omega_{\text{BZ}}}\tilde{f}(\mathbf{G}+\boldsymbol{\omega})e^{+i\boldsymbol{\omega}\cdot\mathbf{x}}\,\mathrm{d}^3\omega = \\ &= \sum_{\mathbf{G}}\tilde{f}(\mathbf{G})e^{+i\mathbf{G}\cdot\mathbf{x}}\int_{\Omega_{\text{BZ}}}\delta(\boldsymbol{\omega})e^{+i\boldsymbol{\omega}\cdot\mathbf{x}}\,d^3\omega = \\ &= \sum_{\mathbf{G}}\tilde{f}(\mathbf{G})e^{+i\mathbf{G}\cdot\mathbf{x}} \end{aligned} \tag{3.5.4}$$

where we used the fact that:

$$\frac{N_{\text{cell}}}{\Omega_{\text{BZ}}}\tilde{f}(\mathbf{G}+\boldsymbol{\omega}) = \tilde{f}(\mathbf{G})\delta(\boldsymbol{\omega})\,.$$

Alternatively, if one is only interested to show that the inverse transformation works, one can directly substitute the direct formula (3.5.3) into (3.5.4) as follows:

$$\begin{aligned} F^{-1}[\tilde{f}(\mathbf{G})] &= \sum_{\mathbf{G}}\tilde{f}(\mathbf{G})e^{+i\mathbf{G}\cdot\mathbf{x}} = \\ &= \sum_{\mathbf{G}}\left(\frac{1}{\Omega_{\text{cell}}}\int_{\Omega_{\text{cell}}}f(\mathbf{x}')e^{-i\mathbf{G}\cdot\mathbf{x}'}\,d^3x'\right)e^{+i\mathbf{G}\cdot\mathbf{x}} = \\ &= \frac{1}{\Omega_{\text{cell}}}\int_{\Omega_{\text{cell}}}f(\mathbf{x}')\sum_{\mathbf{G}}e^{i\mathbf{G}\cdot(\mathbf{x}-\mathbf{x}')}\,d^3x' = \\ &= \frac{1}{\Omega_{\text{cell}}}\int_{\Omega_{\text{cell}}}f(\mathbf{x}')(2\pi)^3\delta\left(\frac{(2\pi)^3}{\Omega_{\text{cell}}}(\mathbf{x}-\mathbf{x}')\right)d^3x' = \\ &= \frac{1}{\Omega_{\text{cell}}}\int_{\Omega_{\text{cell}}}f(\mathbf{x}')(2\pi)^3\frac{\Omega_{\text{cell}}}{(2\pi)^3}\delta(\mathbf{x}-\mathbf{x}')\,d^3x' = \\ &= f(\mathbf{x})\,, \end{aligned}$$

where we used the fact that:

$$\sum_{n=-\infty}^{\infty}e^{inx} = 2\pi\delta(x)\,.$$

Thus we have shown that $F^{-1}[\tilde{f}(\mathbf{G})] = f(\mathbf{x})$.

3.5.1 One Dimension (Fourier Series)

In one dimension with a periodic function $f(x+L) = f(x)$, the volume of a unit cell is $\Omega_{\text{cell}} = L$ and the reciprocal space vectors G are defined using $e^{iGL} = 1$ from which $G_k = \frac{2\pi}{L}k$. The equation (3.5.3) then becomes:

$$F[f(x)] \equiv \tilde{f}(G_k) \equiv c_k = \frac{1}{L}\int_{-\frac{L}{2}}^{\frac{L}{2}}f(x)e^{-iG_kx}\,\mathrm{d}x = \frac{1}{L}\int_{-\frac{L}{2}}^{\frac{L}{2}}f(x)e^{-i(2\pi kx/L)}\,\mathrm{d}x \tag{3.5.1.1}$$

This is exactly the definition of a Fourier series (c_k are the Fourier coefficients). The inverse transform follows from (3.5.4):

$$f(x) = \sum_{k=-\infty}^{\infty} \tilde{f}(G_k) e^{iG_k x} = \sum_{k=-\infty}^{\infty} c_k e^{i(2\pi kx/L)} \tag{3.5.1.2}$$

3.6 Discrete Fourier Transform

In the discrete case, we only have a finite number N of reciprocal points:

$$\begin{aligned} k &= 0, 1, \ldots, N/2-1, -N/2, -N/2+1, \ldots, -1 \quad \text{if } N \text{ is even} \\ k &= 0, 1, \ldots, (N-1)/2, -(N-1)/2, -(N-1)/2+1, \ldots, -1 \quad \text{if } N \text{ is odd} \end{aligned}$$

E.g. for:

$$\begin{aligned} N = 8 \quad &\text{we get} \quad k = 0, 1, 2, 3, -4, -3, -2, -1 \\ N = 9 \quad &\text{we get} \quad k = 0, 1, 2, 3, 4, -4, -3, -2, -1 \end{aligned}$$

The real space function $f(x)$ is sampled at points $x_n = \frac{L}{N} n$ for $n = -N/2, \ldots, N/2-1$ and the equation (3.5.1.1) becomes:

$$\begin{aligned} c_k &= \frac{1}{L} \int_{-\frac{L}{2}}^{\frac{L}{2}} f(x) e^{-i(2\pi kx/L)} \, \mathrm{d}x = \\ &= \lim_{N\to\infty} \frac{1}{L} \sum_{n=-N/2}^{N/2-1} f(x_n) e^{-i(2\pi k x_n/L)} \frac{L}{N} = \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=-N/2}^{N/2-1} f(x_n) e^{-2\pi i \frac{k}{N} n} \end{aligned}$$

The equation (3.5.1.2) becomes:

$$\begin{aligned} f(x_n) &= \sum_{k=-\infty}^{\infty} c_k e^{i(2\pi k x_n/L)} = \\ &= \lim_{N\to\infty} \sum_{k=-N/2}^{N/2-1} c_k e^{i(2\pi k x_n/L)} = \\ &= \lim_{N\to\infty} \sum_{k=-N/2}^{N/2-1} c_k e^{2\pi i \frac{k}{N} n} \end{aligned}$$

Using the fact

$$x_n + L = \frac{L}{N} n + L = \frac{L}{N}(n+N) = x_{n+N} \, ,$$

we can express the periodicity $f(x_n + L) = f(x_n)$ as $f(x_{n+N}) = f(x_n)$. The sums can then be rearranged:

$$\begin{aligned}
c_k &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=-N/2}^{N/2-1} f(x_n) e^{-2\pi i \frac{k}{N} n} = \\
&= \lim_{N\to\infty} \frac{1}{N} \left(\sum_{n=-N/2}^{-1} f(x_n) e^{-2\pi i \frac{k}{N} n} + \sum_{n=0}^{N/2-1} f(x_n) e^{-2\pi i \frac{k}{N} n} \right) = \\
&= \lim_{N\to\infty} \frac{1}{N} \left(\sum_{n=N/2}^{N-1} f(x_{n-N}) e^{-2\pi i \frac{k}{N} (n-N)} + \sum_{n=0}^{N/2-1} f(x_n) e^{-2\pi i \frac{k}{N} n} \right) = \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) e^{-2\pi i \frac{k}{N} n}
\end{aligned}$$

and if we drop the limit and consider a finite N only:

$$\begin{aligned}
f(x_n) &= \sum_{k=-N/2}^{N/2-1} c_k e^{2\pi i \frac{k}{N} n} = \\
&= \left(\sum_{k=-N/2}^{-1} c_k e^{2\pi i \frac{k}{N} n} + \sum_{k=0}^{N/2-1} c_k e^{2\pi i \frac{k}{N} n} \right) = \\
&= \left(\sum_{k=N/2}^{N-1} c_{k-N} e^{2\pi i \frac{(k-N)}{N} n} + \sum_{k=0}^{N/2-1} c_k e^{2\pi i \frac{k}{N} n} \right) = \\
&= \sum_{k=0}^{N-1} c_k e^{2\pi i \frac{k}{N} n}
\end{aligned}$$

Summary, the direct transform:

$$c_k = \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) e^{-2\pi i \frac{k}{N} n} \tag{3.6.1}$$

and inverse transform:

$$f(x_n) = \sum_{k=0}^{N-1} c_k e^{2\pi i \frac{k}{N} n}, \tag{3.6.2}$$

with $x_n = \frac{L}{N} n$. In the limit $N \to \infty$, the equation (3.6.1) becomes (3.5.1.1) and equation (3.6.2) becomes (3.5.1.2) and as we increase N, the discrete Fourier transform numerically converges towards the Fourier series results.

The $\frac{1}{N}$ factor is sometimes moved from the direct to the inverse transform, but then the correspondence with Fourier series is broken (one has to divide and multiply by N appropriately to recover it).

3.7 Fast Fourier Transform (FFT)

We write the discrete Fourier transform (3.6.1) using a notation more commonly used for FFTs:

$$X(k) = \sum_{n=0}^{N-1} x(n) W_N^{kn}$$

where:

$$W_N = e^{-2\pi i/N}$$

Similarly, the inverse discrete Fourier transform (3.6.2) becomes:

$$x(n) = \frac{1}{N}\sum_{k=0}^{N-1} X(k) W_N^{-kn}$$

3.7.1 Decimation In Frequency (DIF)

We start with radix-4:

$$\begin{aligned}
X(k) &= \sum_{n=0}^{N-1} x(n) W_N^{kn} = \\
&= \sum_{n=0}^{\frac{N}{4}-1} x(n) W_N^{kn} + \sum_{n=\frac{N}{4}}^{\frac{2N}{4}-1} x(n) W_N^{kn} + \sum_{n=\frac{2N}{4}}^{\frac{3N}{4}-1} x(n) W_N^{kn} + \sum_{n=\frac{3N}{4}}^{\frac{4N}{4}-1} x(n) W_N^{kn} = \\
&= \sum_{n=0}^{\frac{N}{4}-1} \left[x(n) W_N^{kn} + x\left(n+\frac{N}{4}\right) W_N^{k\left(n+\frac{N}{4}\right)} + x\left(n+\frac{2N}{4}\right) W_N^{k\left(n+\frac{2N}{4}\right)} + x\left(n+\frac{3N}{4}\right) W_N^{k\left(n+\frac{3N}{4}\right)} \right] = \\
&= \sum_{n=0}^{\frac{N}{4}-1} \left[x(n) + x\left(n+\frac{N}{4}\right) W_N^{\frac{kN}{4}} + x\left(n+\frac{2N}{4}\right) W_N^{\frac{2kN}{4}} + x\left(n+\frac{3N}{4}\right) W_N^{\frac{3kN}{4}} \right] W_N^{kn} = \\
&= \sum_{n=0}^{\frac{N}{4}-1} \left[x(n) + x\left(n+\frac{N}{4}\right)(-i)^k + x\left(n+\frac{2N}{4}\right)(-1)^k + x\left(n+\frac{3N}{4}\right) i^k \right] W_N^{kn}
\end{aligned}$$

Now we subdivide the $X(k)$ sequence into 4 subsequences:

$$\begin{aligned}
X(4k) &= \sum_{n=0}^{\frac{N}{4}-1} \left[x(n) + x\left(n+\frac{N}{4}\right)(-i)^{4k} + x\left(n+\frac{2N}{4}\right)(-1)^{4k} + x\left(n+\frac{3N}{4}\right) i^{4k} \right] W_N^{4kn} = \\
&= \sum_{n=0}^{\frac{N}{4}-1} \left[x(n) + x\left(n+\frac{N}{4}\right) + x\left(n+\frac{2N}{4}\right) + x\left(n+\frac{3N}{4}\right) \right] W_{\frac{N}{4}}^{kn}
\end{aligned}$$

Similarly:

$$X(4k+1) = \sum_{n=0}^{\frac{N}{4}-1} \left[x(n) - ix\left(n+\frac{N}{4}\right) - x\left(n+\frac{2N}{4}\right) + ix\left(n+\frac{3N}{4}\right) \right] W_N^n W_{\frac{N}{4}}^{kn}$$

$$X(4k+2) = \sum_{n=0}^{\frac{N}{4}-1} \left[x(n) - x\left(n+\frac{N}{4}\right) + x\left(n+\frac{2N}{4}\right) - x\left(n+\frac{3N}{4}\right) \right] W_N^{2n} W_{\frac{N}{4}}^{kn}$$

$$X(4k+3) = \sum_{n=0}^{\frac{N}{4}-1} \left[x(n) + ix\left(n+\frac{N}{4}\right) - x\left(n+\frac{2N}{4}\right) - ix\left(n+\frac{3N}{4}\right) \right] W_N^{3n} W_{\frac{N}{4}}^{kn}$$

This has a form of a DFT of length $\frac{N}{4}$:

$$X(4k) = \sum_{n=0}^{\frac{N}{4}-1} F_0(n) W_{\frac{N}{4}}^{kn}$$
$$X(4k+1) = \sum_{n=0}^{\frac{N}{4}-1} F_1(n) W_{\frac{N}{4}}^{kn}$$
$$X(4k+2) = \sum_{n=0}^{\frac{N}{4}-1} F_2(n) W_{\frac{N}{4}}^{kn}$$
$$X(4k+3) = \sum_{n=0}^{\frac{N}{4}-1} F_3(n) W_{\frac{N}{4}}^{kn}$$

where

$$\begin{pmatrix} F_0(n) \\ F_1(n) \\ F_2(n) \\ F_3(n) \end{pmatrix} = \begin{pmatrix} x(n) + x\left(n+\frac{N}{4}\right) + x\left(n+\frac{2N}{4}\right) + x\left(n+\frac{3N}{4}\right) \\ x(n) - ix\left(n+\frac{N}{4}\right) - x\left(n+\frac{2N}{4}\right) + ix\left(n+\frac{3N}{4}\right) \\ x(n) - x\left(n+\frac{N}{4}\right) + x\left(n+\frac{2N}{4}\right) - x\left(n+\frac{3N}{4}\right) \\ x(n) + ix\left(n+\frac{N}{4}\right) - x\left(n+\frac{2N}{4}\right) - ix\left(n+\frac{3N}{4}\right) \end{pmatrix} =$$
$$= \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix} \begin{pmatrix} x(n) \\ x\left(n+\frac{N}{4}\right) \\ x\left(n+\frac{2N}{4}\right) \\ x\left(n+\frac{3N}{4}\right) \end{pmatrix}$$

This coefficient matrix for various radix-n schemes can be generated by:

```
>>> from sympy import exp, I, pi, pprint, Matrix
>>> n = 2
>>> Matrix(n, n, lambda i, j: exp(-2*pi*I*i*j/n))
[1  1]
[1 -1]
>>> n = 3
>>> Matrix(n, n, lambda i, j: exp(-2*pi*I*(i*j % n)/n))
[1,              1,              1]
[1, exp(-2*I*pi/3), exp(-4*I*pi/3)]
[1, exp(-4*I*pi/3), exp(-2*I*pi/3)]
>>> n = 4
>>> Matrix(n, n, lambda i, j: exp(-2*pi*I*i*j/n))
[1  1  1  1]
[1 -I -1  I]
[1 -1  1 -1]
[1  I -1 -I]
>>> n = 5
>>> Matrix(n, n, lambda i, j: exp(-2*pi*I*(i*j % n)/n))
[1,              1,              1,              1,              1]
[1, exp(-2*I*pi/5), exp(-4*I*pi/5), exp(-6*I*pi/5), exp(-8*I*pi/5)]
[1, exp(-4*I*pi/5), exp(-8*I*pi/5), exp(-2*I*pi/5), exp(-6*I*pi/5)]
[1, exp(-6*I*pi/5), exp(-2*I*pi/5), exp(-8*I*pi/5), exp(-4*I*pi/5)]
[1, exp(-8*I*pi/5), exp(-6*I*pi/5), exp(-4*I*pi/5), exp(-2*I*pi/5)]
>>> n = 8
>>> Matrix(n, n, lambda i, j: exp(-2*pi*I*(i*j % n)/n))
[1,              1,  1,              1,  1,              1,  1,              1]
[1,   exp(-I*pi/4), -I, exp(-3*I*pi/4), -1, exp(-5*I*pi/4),  I, exp(-7*I*pi/4)]
[1,             -I, -1,              I,  1,             -I, -1,              I]
```

(continues on next page)

(continued from previous page)

```
[1, exp(-3*I*pi/4),  I,   exp(-I*pi/4), -1, exp(-7*I*pi/4), -I, exp(-5*I*pi/4)]
[1,             -1,  1,             -1,  1,             -1,  1,             -1]
[1, exp(-5*I*pi/4), -I, exp(-7*I*pi/4), -1,   exp(-I*pi/4),  I, exp(-3*I*pi/4)]
[1,              I, -1,             -I,  1,              I, -1,             -I]
[1, exp(-7*I*pi/4),  I, exp(-5*I*pi/4), -1, exp(-3*I*pi/4), -I,   exp(-I*pi/4)]
```

One then recursively solves the smaller problems. This approach is used for example in FFTPACK. There are also other approaches how to decompose the DFT, used in various other libraries.

3.8 Laplace Transform

Laplace transform of $f(x)$ is:

$$L[f(x)] = \int_0^\infty f(x)e^{-sx}\,\mathrm{d}x$$

$$L^{-1}[\bar{f}(s)] = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} \bar{f}(s)e^{sx}\,\mathrm{d}s = \sum_{s_0} \mathrm{Res}_{s=s_0}(\bar{f}(s)e^{sx})$$

The contour integration is over the vertical line $\sigma + i\omega$ and σ is chosen large enough so that all residues are to the left of the line (that's because the Laplace transform $\bar{f}(s)$ is only defined for s larger than the residues, so we have to integrate in this range as well). It can be shown that the integral over the left semicircle goes to zero:

$$\begin{aligned}
\left|\int_\Omega e^{sx}g(s)\mathrm{d}s\right| &= \left|\int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} e^{(\sigma+Re^{i\varphi})x}g(\sigma+Re^{i\varphi})iRe^{i\varphi}\mathrm{d}\varphi\right| \leq \\
&\leq R\max_\Omega|g(z)|e^{\sigma x}\int_{\frac{\pi}{2}}^{\frac{3\pi}{2}}\left|e^{xRe^{i\varphi}}\right|\mathrm{d}\varphi = \\
&= R\max_\Omega|g(z)|e^{\sigma x}\int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} e^{xR\cos\varphi}\mathrm{d}\varphi = \\
&= R\max_\Omega|g(z)|e^{\sigma x}\int_0^{\pi} e^{-xR\sin\varphi}\mathrm{d}\varphi = \\
&< \frac{\pi e^{\sigma x}}{x}\max_\Omega|g(z)|
\end{aligned}$$

so the complex integral is equal to the sum of all residues of $\bar{f}(s)e^{sx}$ in the complex plane.

To show that it works:

$$\begin{aligned}
L^{-1}L[f(x)] &= \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty}\left[\int_0^\infty f(x)e^{-sx}\,\mathrm{d}x\right]e^{sx}\,\mathrm{d}s = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty}\left[\int_0^\infty f(x')e^{-sx'}\,\mathrm{d}x'\right]e^{sx}\,\mathrm{d}s = \\
&= \int_0^\infty f(x')\left[\frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} e^{s(x-x')}\,\mathrm{d}s\right]\mathrm{d}x' = \int_0^\infty f(x')\delta(x-x')\,\mathrm{d}x' = f(x)
\end{aligned}$$

where we used:

$$\begin{aligned}
\frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} e^{s(x-x')}\,\mathrm{d}s &= \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} e^{s(x-x')}\,\mathrm{d}s = \frac{1}{2\pi i}\int_{-\infty}^{\infty} e^{(\sigma+i\omega)(x-x')}\,i\mathrm{d}\omega = \\
&= \frac{e^{\sigma(x-x')}}{2\pi}\int_{-\infty}^{\infty} e^{i\omega(x-x')}\,\mathrm{d}\omega = e^{\sigma(x-x')}\delta(x-x') = \delta(x-x')
\end{aligned}$$

and it can be derived from the Fourier transform by transforming a function $U(x)$:

$$U(x) = \begin{cases} f(x)e^{-\sigma x} & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases}$$

and making a substitution $s = \sigma + i\omega$:

$$L[f(x)] \equiv \bar{f}(s) = F[U(x)] \equiv \tilde{U}(\omega) = \int_{-\infty}^{\infty} U(x)e^{-i\omega x}\,\mathrm{d}x = \int_{0}^{\infty} f(x)e^{-\sigma x}e^{-i\omega x}\,\mathrm{d}x = \int_{0}^{\infty} f(x)e^{-sx}\,\mathrm{d}x$$

$$L^{-1}[\bar{f}(s)] \equiv f(x) = U(x)e^{\sigma x} = F^{-1}[\tilde{U}(\omega)]e^{\sigma x} = F^{-1}[\bar{f}(s)]e^{\sigma x} = F^{-1}[\bar{f}(\sigma + i\omega)e^{\sigma x}]$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} \bar{f}(\sigma + i\omega)e^{\sigma x}e^{i\omega x}\,\mathrm{d}\omega = \frac{1}{2\pi i}\int_{\sigma - i\infty}^{\sigma + i\infty} \bar{f}(s)e^{sx}\,\mathrm{d}s = \sum_{s_0} \mathrm{Res}_{s=s_0}(\bar{f}(s)e^{sx})$$

Where the bar ($\bar{f}$) means the Laplace transform and tilde ($\tilde{U}$) means the Fourier transform.

3.9 Hilbert Transform

The Hilbert transform is:

$$H[f(x)](t) \equiv \bar{f}(t) = \mathrm{p.v.}\frac{1}{\pi t} * f(t) = \frac{1}{\pi}\mathrm{p.v.}\int_{-\infty}^{\infty} \frac{f(x)}{t - x}\,\mathrm{d}x$$

By applying the Fourier transform to both sides of the equation, we get:

$$F[H[f(x)](t)](\omega) = F\left[\mathrm{p.v.}\frac{1}{\pi t} * f(t)\right](\omega) = F\left[\mathrm{p.v.}\frac{1}{\pi t}\right](\omega)F[f(t)](\omega) = -i\,\mathrm{sign}(\omega)F[f(t)](\omega)$$

So the Hilbert transform can be calculated using a Fourier transform as:

$$H[f](t) = F^{-1}\left[-i\,\mathrm{sign}(\omega)F[f](\omega)\right](t)$$

The inverse Hilbert transform can then be calculated by inverting:

$$\begin{aligned} F^{-1}\left[-i\,\mathrm{sign}(\omega)F[f(x)](\omega)\right](t) &= H[f](t) = \bar{f}(t) \\ -i\,\mathrm{sign}(\omega)F[f(x)](\omega) &= F[\bar{f}(t)](\omega) \\ F[f(x)](\omega) &= i\,\mathrm{sign}(\omega)F[\bar{f}(t)](\omega) \\ f(x) &= F^{-1}\left[i\,\mathrm{sign}(\omega)F[\bar{f}(t)](\omega)\right](x) \end{aligned}$$

so we get:

$$\begin{aligned} H^{-1}[\bar{f}(t)](x) = f(x) &= F^{-1}\left[i\,\mathrm{sign}(\omega)F[\bar{f}(t)](\omega)\right](x) = \\ &= -F^{-1}\left[-i\,\mathrm{sign}(\omega)F[\bar{f}(t)](\omega)\right](x) = \\ &= -H[\bar{f}(t)](x) = \\ &= -\frac{1}{\pi}\mathrm{p.v.}\int_{-\infty}^{\infty}\frac{\bar{f}(t)}{x - t}\,\mathrm{d}t \end{aligned}$$

From this it also follows:

$$\bar{f}(t) = -H[H[\bar{f}(t)](x)](t)$$

or

$$H[H[f(x)](t)](x) = -f(x)$$

In other words, by applying the Hilbert transform twice, the result is the negative of a function.

3.10 Periodic Functions

A function $f(x)$ is periodic with period T:

$$f(x+T) = f(x)$$

Then you can shift the integration limits by the period T:

$$\int_a^b f(x)\mathrm{d}x = \int_a^b f(x+T)\mathrm{d}x = \int_{a+T}^{b+T} f(x)\mathrm{d}x$$

If you integrate $f(x)$ from 0 to T, you can shift x in $f(x)$ by any constant α:

$$\begin{aligned}
\int_0^T f(x+\alpha)\mathrm{d}x = \\
= \int_\alpha^{T+\alpha} f(x)\mathrm{d}x = \\
= \int_\alpha^0 f(x)\mathrm{d}x + \int_0^T f(x)\mathrm{d}x + \int_T^{T+\alpha} f(x)\mathrm{d}x = \\
= -\int_0^\alpha f(x)\mathrm{d}x + \int_0^T f(x)\mathrm{d}x + \int_0^\alpha f(x)\mathrm{d}x = \\
= \int_0^T f(x)\mathrm{d}x
\end{aligned}$$

3.11 Polar Coordinates

Polar coordinates (radial, azimuth) (r, ϕ) are defined by

$$\begin{aligned}
x &= r\cos\phi \\
y &= r\sin\phi
\end{aligned}$$

3.11.1 Example

When evaluating integrals of the type:

$$l(x,y) = \int_0^{2\pi} \sqrt{(x - r\cos\phi)^2 + (y - r\sin\phi)^2}\,\mathrm{d}\phi$$

we write x and y using polar coordinates:

$$\begin{aligned}
x &= r'\cos\phi' \\
y &= r'\sin\phi'
\end{aligned}$$

and then use the 2π periodicity of $\cos x$:

$$\begin{aligned} l(x,y) &= \int_0^{2\pi} \sqrt{(x - r\cos\phi)^2 + (y - r\sin\phi)^2}\,\mathrm{d}\phi = \\ &= \int_0^{2\pi} \sqrt{x^2 + y^2 + r^2 - 2r(x\cos\phi + y\sin\phi)}\,\mathrm{d}\phi = \\ &= \int_0^{2\pi} \sqrt{r'^2 + r^2 - 2rr'(\cos\phi'\cos\phi + \sin\phi'\sin\phi)}\,\mathrm{d}\phi = \\ &= \int_0^{2\pi} \sqrt{r'^2 + r^2 - 2rr'\cos(\phi - \phi')}\,\mathrm{d}\phi = \\ &= \int_0^{2\pi} \sqrt{r'^2 + r^2 - 2rr'\cos\phi}\,\mathrm{d}\phi = \end{aligned}$$

comparing to:

$$l(0,y) = \int_0^{2\pi} \sqrt{y^2 + r^2 - 2ry\sin\phi}\,\mathrm{d}\phi$$

we can see that because the integral is symmetric, we can just set $x = 0$ and then replace $y \to r'$. The above method does everything algebraically, but you can use this symmetry argument to remember what to do, or even skip the calculation if you are sure that you didn't make a mistake in the "symmetry argument".

3.12 Spherical Coordinates

Spherical coordinates radial (ρ), zenith (θ), azimuth (ϕ):

$$\begin{aligned} x &= \rho\sin\theta\cos\phi \\ y &= \rho\sin\theta\sin\phi \\ z &= \rho\cos\theta \end{aligned} \tag{3.12.1}$$

Note: this meaning of (θ, ϕ) is mostly used in the USA and in many books. In Europe people usually use different symbols, like (ϕ, θ), (ϑ, φ) and others.

The motivation is to first write x and y using polar coordinates:

$$\begin{aligned} x &= \rho_{xy}\cos\phi \\ y &= \rho_{xy}\sin\phi \end{aligned}$$

and then write z and the projection ρ_{xy} of ρ onto the plane $x - y$ using polar coordinates:

$$\begin{aligned} z &= \rho\cos\theta \\ \rho_{xy} &= \rho\sin\theta \end{aligned}$$

so by combining these two we get:

$$\begin{aligned} x &= \rho_{xy}\cos\phi = \rho\sin\theta\cos\phi \\ y &= \rho_{xy}\sin\phi = \rho\sin\theta\sin\phi \\ z &= \rho\cos\theta \end{aligned}$$

3.12.1 Example I

To transform differential operators such as $\frac{\mathrm{d}}{\mathrm{d}x}$ into spherical coordinates, we make use of the chain rule:

$$\frac{\partial}{\partial x} = \frac{\partial\rho}{\partial x}\frac{\partial}{\partial\rho} + \frac{\partial\theta}{\partial x}\frac{\partial}{\partial\theta} + \frac{\partial\phi}{\partial x}\frac{\partial}{\partial\phi}$$

where r, θ and ϕ are functions of x, y, z to be expressed by inverting (3.12.1):

$$\begin{aligned}\rho(x,y,z) &= \sqrt{x^2+y^2+z^2}\\ \theta(x,y,z) &= \arccos\frac{z}{\sqrt{x^2+y^2+z^2}}\\ \phi(x,y,z) &= \arctan\frac{y}{x}\end{aligned}$$

At the end, the derivatives are expressed using ρ, θ, ϕ again. For example

$$\begin{aligned}\frac{\partial\rho}{\partial x} &= \frac{\partial\sqrt{x^2+y^2+z^2}}{\partial x} =\\ &= \frac{x}{\sqrt{x^2+y^2+z^2}} =\\ &= \frac{\rho\sin\theta\cos\phi}{\rho} = \sin\theta\cos\phi\end{aligned}$$

Finally we obtain

$$\begin{aligned}\frac{\partial}{\partial x} &= \sin\theta\cos\phi\frac{\partial}{\partial\rho} + \frac{\cos\theta\cos\phi}{\rho}\frac{\partial}{\partial\theta} - \frac{\sin\phi}{\rho\sin\theta}\frac{\partial}{\partial\phi}\\ \frac{\partial}{\partial y} &= \sin\theta\sin\phi\frac{\partial}{\partial\rho} + \frac{\cos\theta\sin\phi}{\rho}\frac{\partial}{\partial\theta} + \frac{\cos\phi}{\rho\sin\theta}\frac{\partial}{\partial\phi}\\ \frac{\partial}{\partial z} &= \cos\theta\frac{\partial}{\partial\rho} - \frac{\sin\theta}{\rho}\frac{\partial}{\partial\theta}\end{aligned} \tag{3.12.1.1}$$

These expressions can be combined to obtain more complicated objects such as Laplacian (in spherical coordinates). However straightforward this approach is, it is also rather cumbersome; an alternative is discussed in the *Spherical Coordinates* section of differential geometry (where it is shown, that the coefficients in (3.12.1.1) are simply the matrix elements of the inverse Jacobian).

3.12.2 Example II

When evaluating integrals of the type:

$$l(x,y,z) = \int_0^\pi \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi \sqrt{(x-r\sin\theta\cos\phi)^2+(y-r\sin\theta\sin\phi)^2+(z-r\cos\theta)^2}\,\sin\theta$$

we write x and y using polar coordinates:

$$\begin{aligned}x &= \rho_{xy}\cos\phi'\\ y &= \rho_{xy}\sin\phi'\end{aligned}$$

and simplify:

$$\begin{aligned}l(x,y,z) &= \int_0^\pi \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi \sqrt{(x-r\sin\theta\cos\phi)^2+(y-r\sin\theta\sin\phi)^2+(z-r\cos\theta)^2}\,\sin\theta =\\ &= \int_0^\pi \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi \sqrt{x^2+y^2+z^2+r^2-2r(x\sin\theta\cos\phi+y\sin\theta\sin\phi+z\cos\theta)}\,\sin\theta =\\ &= \int_0^\pi \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi \sqrt{\rho_{xy}+z^2+r^2-2r(\rho_{xy}\cos\phi'\sin\theta\cos\phi+\rho_{xy}\sin\phi'\sin\theta\sin\phi+z\cos\theta)}\,\sin\theta =\\ &= \int_0^\pi \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi \sqrt{\rho_{xy}+z^2+r^2-2r(\rho_{xy}\cos(\phi-\phi')\sin\theta+z\cos\theta)}\,\sin\theta =\\ &= \int_0^\pi \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi \sqrt{\rho_{xy}+z^2+r^2-2r(\rho_{xy}\cos\phi\sin\theta+z\cos\theta)}\,\sin\theta\end{aligned}$$

comparing to:

$$l(0,0,z) = \int_0^{\pi} \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi \sqrt{z^2 + r^2 - 2rz\cos\theta}\, \sin\theta =$$

we can see that because the integral is symmetric, we can just set $x = 0$, $y = 0$ and then replace $z \to \rho$.

3.13 Argument function, atan2

Argument function $\arg(z)$ is any φ such that

$$z = re^{i\varphi}$$

Obviously $\arg(z)$ is unique up to any integer multiple of 2π. By taking the principal value of the $\arg(z)$ function, e.g. fixing $\arg(z)$ to the interval $(-\pi, \pi]$ (so that the branch cut is on the negative x-axis, as usual), we get the $\operatorname{Arg}(z)$ function:

$$-\pi < \operatorname{Arg} z \leq \pi$$

then $\arg z = \operatorname{Arg} z + 2\pi n$, where $n = 0, \pm 1, \pm 2, \ldots$. We can then use the following formula to easily calculate $\operatorname{Arg} z$ for any $z = x + iy$ (except $x = y = 0$, i.e. $z = 0$, where it is not defined):

$$\operatorname{Arg}(x+iy) = \begin{cases} \pi & y = 0; x < 0; \\ 2\,\operatorname{atan}\frac{y}{\sqrt{x^2+y^2}+x} & \text{otherwise} \end{cases}$$

Finally we define $\operatorname{atan}2(y, x)$ as:

$$\operatorname{atan}2(y,x) = \operatorname{Arg}(x+iy) = \begin{cases} \pi & y = 0; x < 0; \\ 2\,\operatorname{atan}\frac{y}{\sqrt{x^2+y^2}+x} & \text{otherwise} \end{cases}$$

The angle $\phi = \operatorname{atan}2(y, x)$ is the angle of the point (x, y) on the unit circle (assuming the usual conventions), and it works for all quadrants ($\phi = \operatorname{atan}(\frac{y}{x})$ only works for the first and fourth quadrant, where $\operatorname{atan}(\frac{y}{x}) = \operatorname{atan}2(y, x)$, but in the second and third qudrant, $\operatorname{atan}(\frac{y}{x})$ gives the wrong angles, while $\operatorname{atan}2(y, x)$ gives the correct angles). So in particular:

$$\operatorname{atan}2(0,1) = 2\,\operatorname{atan}\frac{0}{\sqrt{1^2+0^2}+1} = 0$$
$$\operatorname{atan}2(0,-1) = \pi$$
$$\operatorname{atan}2(1,0) = 2\,\operatorname{atan}\frac{1}{\sqrt{0^2+1^2}+0} = 2\,\operatorname{atan}1 = \frac{\pi}{2}$$
$$\operatorname{atan}2(-1,0) = 2\,\operatorname{atan}\frac{-1}{\sqrt{0^2+1^2}+0} = -2\,\operatorname{atan}1 = -\frac{\pi}{2}$$

This convention ($\operatorname{atan}2(y, x)$) is used for example in Python, C or Fortran. Some people might interchange x with y in the definition (i.e. $\operatorname{atan}2(x, y) = \operatorname{Arg}(y + ix)$), but it is not very common.

The following useful relations hold:

$$\sin \operatorname{atan} 2(y,x) = \frac{y}{\sqrt{x^2+y^2}} \qquad \text{except } x=y=0$$
$$\cos \operatorname{atan} 2(y,x) = \frac{x}{\sqrt{x^2+y^2}} \qquad \text{except } x=y=0$$
$$\tan \operatorname{atan} 2(y,x) = \frac{y}{x} \qquad \text{for } x \neq 0$$
$$\operatorname{atan} 2(ky,kx) = \operatorname{atan} 2(y,x) \qquad \text{for } k>0$$
$$\operatorname{atan} 2(\sin x, \cos x) = x + 2\pi \left\lfloor \frac{\pi - x}{2\pi} \right\rfloor$$
$$\operatorname{atan} 2(-y,x) = -\operatorname{atan} 2(y,x) + 2\pi \left\lfloor \frac{\operatorname{atan} 2(y,x)+\pi}{2\pi} \right\rfloor$$
$$\frac{\partial}{\partial y} \operatorname{atan} 2(y,x) = \frac{x}{x^2+y^2}$$
$$\frac{\partial}{\partial x} \operatorname{atan} 2(y,x) = -\frac{y}{x^2+y^2}$$

We now prove them. The following works for all x, y except for $x=y=0$:

$$\sin \operatorname{atan} 2(y,x) = \begin{cases} \sin \pi & y=0; x<0; \\ \sin\left(2 \operatorname{atan} \frac{y}{\sqrt{x^2+y^2}+x}\right) & \text{otherwise} \end{cases} =$$
$$= \begin{cases} 0 & y=0; x<0; \\ \frac{y}{\sqrt{x^2+y^2}} & \text{otherwise} \end{cases} =$$
$$= \begin{cases} \frac{y}{\sqrt{x^2+y^2}} & y=0; x<0; \\ \frac{y}{\sqrt{x^2+y^2}} & \text{otherwise} \end{cases} = \frac{y}{\sqrt{x^2+y^2}}$$
$$\cos \operatorname{atan} 2(y,x) = \begin{cases} \cos \pi & y=0; x<0; \\ \cos\left(2 \operatorname{atan} \frac{y}{\sqrt{x^2+y^2}+x}\right) & \text{otherwise} \end{cases} =$$
$$= \begin{cases} -1 & y=0; x<0; \\ \frac{x}{\sqrt{x^2+y^2}} & \text{otherwise} \end{cases} =$$
$$= \begin{cases} \frac{x}{\sqrt{x^2+y^2}} & y=0; x<0; \\ \frac{x}{\sqrt{x^2+y^2}} & \text{otherwise} \end{cases} = \frac{x}{\sqrt{x^2+y^2}}$$

Tangent is infinite for $\pm\frac{\pi}{2}$, which corresponds to $x=0$, so the following works for all $x \neq 0$:

$$\tan \operatorname{atan} 2(y,x) = \begin{cases} \tan \pi & y=0; x<0; \\ \tan\left(2 \operatorname{atan} \frac{y}{\sqrt{x^2+y^2}+x}\right) & \text{otherwise} \end{cases} =$$
$$= \begin{cases} 0 & y=0; x<0; \\ \frac{y}{x} & \text{otherwise} \end{cases} =$$
$$= \begin{cases} \frac{y}{x} & y=0; x<0; \\ \frac{y}{x} & \text{otherwise} \end{cases} = \frac{y}{x}$$

Finally:

$$\operatorname{atan2}(\sin x, \cos x) = \begin{cases} \pi & x = \pi; \\ 2 \operatorname{atan} \frac{\sin x}{\sqrt{\cos^2 x + \sin^2 x} + \cos x} & \text{otherwise} \end{cases} =$$

$$= \begin{cases} \pi & x = \pi; \\ 2 \operatorname{atan} \frac{\sin x}{1 + \cos x} & \text{otherwise} \end{cases} =$$

$$= \begin{cases} x & x = \pi; \\ 2 \operatorname{atan} \left(\tan \frac{x}{2} \right) & \text{otherwise} \end{cases} = x + 2\pi \left\lfloor \frac{\pi - x}{2\pi} \right\rfloor$$

In the above, we used the following double angle formulas:

$$\sin 2x = \frac{2 \tan x}{1 + \tan^2 x}$$

$$\cos 2x = \frac{1 - \tan^2 x}{1 + \tan^2 x}$$

$$\tan 2x = \frac{2 \tan x}{1 - \tan^2 x}$$

to simplify the following expressions:

$$\sin\left(2\ \operatorname{atan}\frac{y}{\sqrt{x^2+y^2}+x}\right)=\frac{2\tan\operatorname{atan}\frac{y}{\sqrt{x^2+y^2}+x}}{1+\tan^2\operatorname{atan}\frac{y}{\sqrt{x^2+y^2}+x}}=$$
$$=\frac{2\frac{y}{\sqrt{x^2+y^2}+x}}{1+\left(\frac{y}{\sqrt{x^2+y^2}+x}\right)^2}=\frac{2y\left(\sqrt{x^2+y^2}+x\right)}{\left(\sqrt{x^2+y^2}+x\right)^2+y^2}=$$
$$=\frac{y\left(\sqrt{x^2+y^2}+x\right)}{x^2+y^2+x\sqrt{x^2+y^2}}=\frac{y\left(\sqrt{x^2+y^2}+x\right)}{\sqrt{x^2+y^2}\left(\sqrt{x^2+y^2}+x\right)}=$$
$$=\frac{y}{\sqrt{x^2+y^2}}$$
$$\cos\left(2\ \operatorname{atan}\frac{y}{\sqrt{x^2+y^2}+x}\right)=\frac{1-\tan^2\operatorname{atan}\frac{y}{\sqrt{x^2+y^2}+x}}{1+\tan^2\operatorname{atan}\frac{y}{\sqrt{x^2+y^2}+x}}=$$
$$=\frac{1-\left(\frac{y}{\sqrt{x^2+y^2}+x}\right)^2}{1+\left(\frac{y}{\sqrt{x^2+y^2}+x}\right)^2}=\frac{\left(\sqrt{x^2+y^2}+x\right)^2-y^2}{\left(\sqrt{x^2+y^2}+x\right)^2+y^2}=$$
$$=\frac{x\left(\sqrt{x^2+y^2}+x\right)}{x^2+y^2+x\sqrt{x^2+y^2}}=\frac{x\left(\sqrt{x^2+y^2}+x\right)}{\sqrt{x^2+y^2}\left(\sqrt{x^2+y^2}+x\right)}=$$
$$=\frac{x}{\sqrt{x^2+y^2}}$$
$$\tan\left(2\ \operatorname{atan}\frac{y}{\sqrt{x^2+y^2}+x}\right)=\frac{2\tan\operatorname{atan}\frac{y}{\sqrt{x^2+y^2}+x}}{1-\tan^2\operatorname{atan}\frac{y}{\sqrt{x^2+y^2}+x}}=$$
$$=\frac{2\frac{y}{\sqrt{x^2+y^2}+x}}{1-\left(\frac{y}{\sqrt{x^2+y^2}+x}\right)^2}=\frac{2y\left(\sqrt{x^2+y^2}+x\right)}{\left(\sqrt{x^2+y^2}+x\right)^2-y^2}=$$
$$=\frac{y\left(\sqrt{x^2+y^2}+x\right)}{x\left(\sqrt{x^2+y^2}+x\right)}=\frac{y}{x}$$

Finally, for all $k>0$ we get:

$$\operatorname{atan}2(ky,kx)=\operatorname{Arg}(kx+iky)=\begin{cases}\pi & y=0;x<0;\\ 2\ \operatorname{atan}\frac{ky}{\sqrt{(kx)^2+(ky)^2}+kx} & \text{otherwise}\end{cases}=$$
$$=\begin{cases}\pi & y=0;x<0;\\ 2\ \operatorname{atan}\frac{y}{\sqrt{x^2+y^2}+x} & \text{otherwise}\end{cases}=\operatorname{Arg}(x+iy)=\operatorname{atan}2(y,x)$$

The symmetry property can be proven by:

$$\operatorname{atan2}(-y,x) = \begin{cases} \pi & y=0; x<0; \\ 2\,\operatorname{atan}\frac{-y}{\sqrt{x^2+(-y)^2}+x} & \text{otherwise} \end{cases} =$$
$$= \begin{cases} \pi & y=0; x<0; \\ -\left(2\,\operatorname{atan}\frac{y}{\sqrt{x^2+y^2}+x}\right) & \text{otherwise} \end{cases} =$$
$$= -\operatorname{atan2}(y,x) + 2\pi\left\lfloor\frac{\operatorname{atan2}(y,x)+\pi}{2\pi}\right\rfloor$$

To prove the derivatives, we do:

$$\frac{\partial}{\partial y}\operatorname{atan2}(y,x) = 2\frac{\partial}{\partial y}\operatorname{atan}\frac{y}{\sqrt{x^2+y^2}+x} = \frac{x}{x^2+y^2}$$
$$\frac{\partial}{\partial x}\operatorname{atan2}(y,x) = 2\frac{\partial}{\partial x}\operatorname{atan}\frac{y}{\sqrt{x^2+y^2}+x} = -\frac{y}{x^2+y^2}$$

Code:

```
>>> from sympy import atan, sqrt, var
>>> var("x y")
(x, y)
>>> (2*atan(y/(sqrt(x**2+y**2)+x)).diff(y)).simplify()
x/(x**2 + y**2)
>>> (2*atan(y/(sqrt(x**2+y**2)+x)).diff(x)).simplify()
-y/(x**2 + y**2)
```

An example of an application:

$$A\sin x + B\cos x = \sqrt{A^2+B^2}\left(\frac{A}{\sqrt{A^2+B^2}}\sin x + \frac{B}{\sqrt{A^2+B^2}}\cos x\right) =$$
$$= \sqrt{A^2+B^2}\left(\cos\delta\sin x + \sin\delta\cos x\right) = \sqrt{A^2+B^2}\sin(x+\delta) =$$
$$= \sqrt{A^2+B^2}\sin(x+\operatorname{atan2}(B,A))$$

where

$$\delta = \operatorname{atan2}\left(\frac{B}{\sqrt{A^2+B^2}}, \frac{A}{\sqrt{A^2+B^2}}\right) = \operatorname{atan2}(B,A)$$

Another application

$$\operatorname{atan2}(\cos x, -\sin x) = \operatorname{atan2}\left(\sin\left(x+\frac{\pi}{2}\right), \cos\left(x+\frac{\pi}{2}\right)\right) = x+\frac{\pi}{2}$$

3.14 Multiple Argument Formulas

3.14.1 sin(a x)

Systematic way to derive all multiple argument formulas is to use the following relation:

$$\sin(ax) = U_{a-1}(\cos x)\sin x$$

where $U_n(x)$ are the Chebyshev polynomials of the second kind, first few are:

$$
\begin{aligned}
U_{-3}(x) &= -2x \\
U_{-2}(x) &= -1 \\
U_{-1}(x) &= 0 \\
U_{-\frac{1}{2}}(x) &= \frac{1}{\sqrt{2}\sqrt{x+1}} \\
U_0(x) &= 1 \\
U_{\frac{1}{2}}(x) &= \frac{2x+1}{\sqrt{2}\sqrt{x+1}} \\
U_1(x) &= 2x \\
U_2(x) &= 4x^2 - 1 \\
U_3(x) &= 8x^3 - 4x \\
U_4(x) &= 16x^4 - 12x^2 + 1 \\
U_5(x) &= 32x^5 - 32x^3 + 6x \\
U_6(x) &= 64x^6 - 80x^4 + 24x^2 - 1
\end{aligned}
$$

Code:

```
>>> from sympy import chebyshevu, var
>>> var("x")
>>> for i in range(7): print "U_%d(x) = %s" % (i, chebyshevu(i, x))
U_0(x) = 1
U_1(x) = 2*x
U_2(x) = -1 + 4*x**2
U_3(x) = -4*x + 8*x**3
U_4(x) = 1 - 12*x**2 + 16*x**4
U_5(x) = 6*x - 32*x**3 + 32*x**5
U_6(x) = -1 + 24*x**2 - 80*x**4 + 64*x**6
```

One can then use this to calculate:

$$
\begin{aligned}
\sin(-2x) &= U_{-3}(\cos x)\sin x = -2\cos x \sin x \\
\sin(-x) &= U_{-2}(\cos x)\sin x = -\sin x \\
\sin 0 &= U_{-1}(\cos x)\sin x = 0 \\
\sin\frac{x}{2} &= U_{-\frac{1}{2}}(\cos x)\sin x = \frac{\sin x}{\sqrt{2}\sqrt{\cos x+1}} = \frac{\sqrt{1-\cos^2 x}}{\sqrt{2}\sqrt{\cos x+1}} = \frac{\sqrt{1-\cos x}}{\sqrt{2}} \\
\sin x &= U_0(\cos x)\sin x = \sin x \\
\sin\frac{3x}{2} &= U_{\frac{1}{2}}(\cos x)\sin x = \frac{(2\cos x+1)\sin x}{\sqrt{2}\sqrt{\cos x+1}} = \frac{(2\cos x+1)\sqrt{1-\cos^2 x}}{\sqrt{2}\sqrt{\cos x+1}} = \frac{(2\cos x+1)\sqrt{1-\cos x}}{\sqrt{2}} \\
\sin 2x &= U_1(\cos x)\sin x = 2\cos x \sin x \\
\sin 3x &= U_2(\cos x)\sin x = (4\cos^2 x - 1)\sin x
\end{aligned}
$$

Code:

```
>>> from sympy import chebyshevu, var, sin, cos
>>> var("x")
>>> for n in range(1, 7): print "sin(%d*x) = %s" % (n, chebyshevu(n-1, cos(x))*sin(x))
sin(1*x) = sin(x)
sin(2*x) = 2*cos(x)*sin(x)
```

(continues on next page)

(continued from previous page)

```
sin(3*x) = -(1 - 4*cos(x)**2)*sin(x)
sin(4*x) = (-4*cos(x) + 8*cos(x)**3)*sin(x)
sin(5*x) = (1 - 12*cos(x)**2 + 16*cos(x)**4)*sin(x)
sin(6*x) = (6*cos(x) - 32*cos(x)**3 + 32*cos(x)**5)*sin(x)
```

3.14.2 cos(a x)

Similarly as above, we use:

$$\cos(ax) = T_a(\cos x)$$

where $T_n(x)$ are the Chebyshev polynomials of the first kind, first few are:

$$\begin{aligned}
T_0(x) &= 1 \\
T_{\frac{1}{2}}(x) &= \frac{\sqrt{x+1}}{\sqrt{2}} \\
T_1(x) &= x \\
T_{\frac{3}{2}}(x) &= \frac{(2x-1)\sqrt{x+1}}{\sqrt{2}} \\
T_2(x) &= 2x^2 - 1 \\
T_3(x) &= 4x^3 - 3x \\
T_4(x) &= 8x^4 - 8x^2 + 1 \\
T_5(x) &= 16x^5 - 20x^3 + 5x \\
T_6(x) &= 32x^6 - 48x^4 + 18x^2 - 1
\end{aligned}$$

Code:

```
>>> from sympy import chebyshevt, var
>>> var("x")
>>> for i in range(7): print "T_%d(x) = %s" % (i, chebyshevt(i, x))
T_0(x) = 1
T_1(x) = x
T_2(x) = -1 + 2*x**2
T_3(x) = -3*x + 4*x**3
T_4(x) = 1 - 8*x**2 + 8*x**4
T_5(x) = 5*x - 20*x**3 + 16*x**5
T_6(x) = -1 + 18*x**2 - 48*x**4 + 32*x**6
```

One can then use this to calculate:

$$\begin{aligned}
\cos 0 &= T_0(\cos x) = 1 \\
\cos\frac{x}{2} &= T_{\frac{1}{2}}(\cos x) = \frac{\sqrt{1+\cos x}}{\sqrt{2}} \\
\cos x &= T_1(\cos x) = \cos x \\
\cos\frac{3x}{2} &= T_{\frac{3}{2}}(\cos x) = \frac{(2\cos x - 1)\sqrt{1+\cos x}}{\sqrt{2}} \\
\cos 2x &= T_2(\cos x) = 2\cos^2 x - 1 \\
\cos 3x &= T_3(\cos x) = 4\cos^3 x - 3\cos x
\end{aligned}$$

Code:

```
>>> from sympy import chebyshevt, var, cos
>>> var("x")
>>> for n in range(7): print "cos(%d*x) = %s" % (n, chebyshevt(n, cos(x)))
cos(0*x) = 1
cos(1*x) = cos(x)
cos(2*x) = -1 + 2*cos(x)**2
cos(3*x) = -3*cos(x) + 4*cos(x)**3
cos(4*x) = 1 - 8*cos(x)**2 + 8*cos(x)**4
cos(5*x) = 5*cos(x) - 20*cos(x)**3 + 16*cos(x)**5
cos(6*x) = -1 + 18*cos(x)**2 - 48*cos(x)**4 + 32*cos(x)**6
```

3.15 Delta Function

Delta function $\delta(x)$ is defined such that this relation holds:

$$\int f(x)\delta(x-t)\mathrm{d}x = f(t) \tag{3.15.1}$$

No such function exists, but one can find many sequences "converging" to a delta function:

$$\lim_{\alpha\to\infty} \delta_\alpha(x) = \delta(x) \tag{3.15.2}$$

more precisely:

$$\lim_{\alpha\to\infty} \int f(x)\delta_\alpha(x)\mathrm{d}x = \int f(x) \lim_{\alpha\to\infty} \delta_\alpha(x)\mathrm{d}x = f(0) \tag{3.15.3}$$

one example of such a sequence is:

$$\delta_\alpha(x) = \frac{1}{\pi x}\sin(\alpha x)$$

It's clear that (3.15.3) holds for any well behaved function $f(x)$. Some mathematicians like to say that it's incorrect to use such a notation when in fact the integral (3.15.1) doesn't "exist", but we will not follow their approach, because it is not important if something "exists" or not, but rather if it is clear what we mean by our notation: (3.15.1) is a shorthand for (3.15.3) and (3.15.2) gets a mathematically rigorous meaning when you integrate both sides and use (3.15.1) to arrive at (3.15.3). Thus one uses the relations (3.15.1), (3.15.2), (3.15.3) to derive all properties of the delta function.

Let's give an example. Let $\hat{\mathbf{r}}$ be the unit vector in 3D and we can label it using spherical coordinates $\hat{\mathbf{r}} = \hat{\mathbf{r}}(\theta, \phi)$. We can also express it in cartesian coordinates as $\hat{\mathbf{r}}(\theta, \phi) = (\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta)$.

$$f(\hat{\mathbf{r}}') = \int \delta(\hat{\mathbf{r}} - \hat{\mathbf{r}}') f(\hat{\mathbf{r}})\, \mathrm{d}\hat{\mathbf{r}} \tag{3.15.4}$$

Expressing $f(\hat{\mathbf{r}}) = f(\theta, \phi)$ as a function of θ and ϕ we have

$$f(\theta', \phi') = \int \delta(\theta - \theta')\delta(\phi - \phi') f(\theta, \phi)\, \mathrm{d}\theta\mathrm{d}\phi \tag{3.15.5}$$

Expressing (3.15.4) in spherical coordinates we get

$$f(\theta', \phi') = \int \delta(\hat{\mathbf{r}} - \hat{\mathbf{r}}') f(\theta, \phi)\sin\theta\, \mathrm{d}\theta\mathrm{d}\phi$$

and comparing to (3.15.5) we finally get

$$\delta(\hat{\mathbf{r}} - \hat{\mathbf{r}}') = \frac{1}{\sin\theta}\delta(\theta - \theta')\delta(\phi - \phi')$$

In exactly the same manner we get

$$\delta(\mathbf{r}-\mathbf{r}')=\delta(\hat{\mathbf{r}}-\hat{\mathbf{r}}')\frac{\delta(\rho-\rho')}{\rho^2}$$

See also (3.17.4.1) for an example of how to deal with more complex expressions involving the delta function like $\delta^2(x)$.

When integrating over finite interval, this formula is very useful:

$$\int_a^b f(x)\delta(x-t)\mathrm{d}x=f(t)\theta(b-t)\theta(t-a)$$

in other words, the integral vanishes unless $a<t<b$. In the limit $a\to-\infty$ and $b\to\infty$ we get:

$$\begin{aligned}\theta(b-t)&\to\theta(\infty-t)=1\\ \theta(t-a)&\to\theta(t-(-\infty))=\theta(t+\infty)=1\end{aligned}$$

Another integral that converges to a delta function is:

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{i\omega x}\mathrm{d}x=\lim_{L\to\infty}\frac{1}{2\pi}\int_{-L}^{L}e^{i\omega x}\mathrm{d}x=\lim_{L\to\infty}\frac{\sin\omega L}{\pi\omega}=\delta(\omega)$$

3.16 Distributions

Some mathematicians like to use distributions and a mathematical notation for that, which I think is making things less clear, but nevertheless it's important to understand it too, so the notation is explained in this section, but I discourage to use it – I suggest to only use the physical notation as explained below. The math notation below is put into quotation marks, so that it's not confused with the physical notation.

The distribution is a functional and each function $f(x)$ can be identified with a distribution "T_f" that it generates using this definition ($\varphi(x)$ is a test function):

$$\text{"}T_f(\phi(x))\text{"}\equiv\int f(x)\varphi(x)\mathrm{d}x\equiv\text{"}f(\varphi(x))\text{"}\equiv\text{"}(f(x),\varphi(x))\text{"}$$

besides that, one can also define distributions that can't be identified with regular functions, one example is a delta distribution (Dirac delta function):

$$\text{"}\delta(\phi(x))\text{"}\equiv\phi(0)\equiv\int\delta(x)\phi(x)\mathrm{d}x$$

The last integral is not used in mathematics, in physics on the other hand, the first expressions ("$\delta(\phi(x))$") is not used, so $\delta(x)$ always means that you have to integrate it, as explained in the previous section, so it behaves like a regular function (except that such a function doesn't exist and the precise mathematical meaning is only after you integrate it, or through the identification above with distributions).

One then defines common operations via acting on the generating function, then observes the pattern and defines it for all distributions. For example differentiation:

$$\text{"}\frac{\mathrm{d}}{\mathrm{d}x}T_f(\varphi)\text{"}=\text{"}T_{f'}(\varphi)\text{"}=\int f'\varphi\mathrm{d}x=-\int f\varphi'\mathrm{d}x=\text{"}-T_f(\varphi')\text{"}$$

so:

$$\text{"}\frac{\mathrm{d}}{\mathrm{d}x}T(\varphi)\text{"}=\text{"}-T(\varphi')\text{"}$$

Multiplication:

$$"gT_f(\varphi)" = "T_{gf}(\varphi)" = \int gf\varphi \mathrm{d}x = "T_f(g\varphi)"$$

so:

$$"gT(\varphi)" = "T(g\varphi)"$$

Fourier transform:

$$"FT_f(\varphi)" = "T_{Ff}(\varphi)" = \int F(f)\varphi \mathrm{d}x =$$
$$= \int \left[\int e^{-ikx} f(k)\mathrm{d}k\right] \varphi(x)\mathrm{d}x = \int f(k) \left[\int e^{-ikx}\varphi(x)\mathrm{d}x\right] \mathrm{d}k = \int f(x) \left[\int e^{-ikx}\varphi(k)\mathrm{d}k\right] \mathrm{d}x =$$
$$= \int fF(\varphi)\mathrm{d}x = "T_f(F\varphi)"$$

so:

$$"FT(\varphi)" = "T(F\varphi)"$$

But as you can see, the notation is just making things more complex, since it's enough to just work with the integrals and forget about the rest. One can then even omit the integrals, with the understanding that they are implicit.

Some more examples:

$$\int \delta(x - x_0)\varphi(x)\mathrm{d}x = \int \delta(x)\varphi(x + x_0)\mathrm{d}x = \varphi(x_0) \equiv "\delta(\varphi(x + x_0))"$$

Proof of $\delta(-x) = \delta(x)$:

$$\int_{-\infty}^{\infty} \delta(-x)\varphi(x)\mathrm{d}x = -\int_{\infty}^{-\infty} \delta(y)\varphi(-y)\mathrm{d}y = \int_{-\infty}^{\infty} \delta(x)\varphi(-x)\mathrm{d}x \equiv "\delta(\varphi(-x))" = \varphi(0) = "\delta(\phi(x))" \equiv \int_{-\infty}^{\infty} \delta(x)\varphi(x)\mathrm{d}x$$

Proof of $x\delta(x) = 0$:

$$\int x\delta(x)\varphi(x)\mathrm{d}x = "\delta(x\varphi(x))" = 0 \cdot \varphi(0) = 0$$

Proof of $\delta(cx) = \frac{\delta(x)}{|c|}$:

$$\int \delta(cx)\varphi(x)\mathrm{d}x = \frac{1}{|c|}\int \delta(x)\varphi\left(\frac{x}{c}\right)\mathrm{d}x = "\delta\left(\frac{\varphi\left(\frac{x}{c}\right)}{|c|}\right)" = \frac{\delta(0)}{|c|} = "\frac{\delta(\varphi(x))}{|c|}" = \int \frac{\delta(x)}{|c|}\varphi(x)\mathrm{d}x$$

To prove that $\lim_{L\to\infty} \frac{1}{\pi x}\sin(Lx) = \delta(x)$ we do the following calculation:

$$\left[\int_{-\infty}^{\infty} \lim_{L\to\infty} \frac{1}{\pi x}\sin(Lx)\varphi(x)\mathrm{d}x\right] - \varphi(0) =$$
$$= \lim_{L\to\infty} \int_{-\infty}^{\infty} \frac{1}{\pi x}\sin(Lx)(\varphi(x) - \varphi(0))\mathrm{d}x =$$
$$= \lim_{L\to\infty} \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\varphi(x) - \varphi(0)}{x}\sin(Lx)\mathrm{d}x =$$
$$= \lim_{L\to\infty} \frac{1}{\pi}\int_{-\infty}^{\infty} h(x)\sin(Lx)\mathrm{d}x = 0$$

where the function $h(x) = \frac{\varphi(x)-\varphi(0)}{x}$ is bounded and $h(0) = \lim_{x\to 0} \frac{\varphi(x)-\varphi(0)}{x} = \varphi'(0)$ is finite since the test function $\varphi(x)$ is infinitely differentiable. From the Riemann–Lebesgue lemma, the integral then converges towards zero as $L \to \infty$.

3.17 Variations and Functional Derivatives

Variations and functional derivatives are generalization of differentials and partial derivatives to functionals. It is important to master this subject just like regular differentials/derivatives in calculus.

3.17.1 Functions of One Variable

Let's first review differentials and derivatives of functions of one variable. We will use an approach that directly generalizes to multivariable functions and functionals. The differential $\mathrm{d}f$ is defined as:

$$\mathrm{d}f \equiv \lim_{\varepsilon\to 0} \frac{f(x+\varepsilon h)-f(x)}{\varepsilon} = ah$$

Last equality follows from the fact, that the limit is a linear function of h:

$$\lim_{\varepsilon\to 0} \frac{f(x+\varepsilon h)-f(x)}{\varepsilon} = \lim_{\eta\to 0} \frac{f(x+\eta)-f(x)}{\left(\frac{\eta}{h}\right)} = \left(\lim_{\eta\to 0} \frac{f(x+\eta)-f(x)}{\eta}\right) h$$

Where we used the substitution $\eta = \varepsilon h$. We define the derivative $\frac{\mathrm{d}f}{\mathrm{d}x}$ as:

$$\frac{\mathrm{d}f}{\mathrm{d}x} = a$$

To get a formula for $\frac{\mathrm{d}f}{\mathrm{d}x}$, we set $h = 1$ and get:

$$\frac{\mathrm{d}f}{\mathrm{d}x} = a = a\cdot 1 = \lim_{\varepsilon\to 0} \frac{f(x+\varepsilon\cdot 1)-f(x)}{\varepsilon} = \lim_{\varepsilon\to 0} \frac{f(x+\varepsilon)-f(x)}{\varepsilon}$$

Using the formulas above we get an equivalent expression for the differential:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\varepsilon} f(x+\varepsilon h)\bigg|_{\varepsilon=0} &= \lim_{\eta\to 0} \frac{f(x+(\varepsilon+\eta)h)-f(x+\varepsilon h)}{\eta}\bigg|_{\varepsilon=0} = \\ &= \lim_{\eta\to 0} \frac{f(x+\eta h)-f(x)}{\eta} = \\ &= \lim_{\varepsilon\to 0} \frac{f(x+\varepsilon h)-f(x)}{\varepsilon}\end{aligned}$$

So we get a general formula (the analogy of which we will use later):

$$\mathrm{d}f \equiv \frac{\mathrm{d}}{\mathrm{d}\varepsilon} f(x+\varepsilon h)\bigg|_{\varepsilon=0} = \lim_{\varepsilon\to 0} \frac{f(x+\varepsilon h)-f(x)}{\varepsilon} = ah$$

The variable x can be treated as a function (a very simple one):

$$x = g(x)$$

So we define $\mathrm{d}x$ as:

$$\mathrm{d}x \equiv \mathrm{d}g = \frac{\mathrm{d}g}{\mathrm{d}x} h = h$$

As such, $\mathrm{d}x$ can have two meanings: either $\mathrm{d}x = h = x - x_0$ (a finite change in the variable x) or a differential (if x depends on another variable, thanks to the chain rule everything will work). With this understanding, for all calculations, we only need the following two formulas — the definition of the differential (using a limit):

$$\mathrm{d}f = \lim_{\varepsilon\to 0} \frac{f(x+\varepsilon \mathrm{d}x)-f(x)}{\varepsilon}$$

and the definition of the derivative (using the differential):

$$\mathrm{d}f = \frac{\mathrm{d}f}{\mathrm{d}x}\mathrm{d}x$$

where $\mathrm{d}x$ is either a differential or a finite change in the variable x.

If for example $x = p(y)$ is a function of y then in the above $\mathrm{d}x$ is a differential and we get:

$$\mathrm{d}f = \frac{\mathrm{d}f}{\mathrm{d}x}\mathrm{d}x = \frac{\mathrm{d}f}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}y}\mathrm{d}y$$

Thanks to the chain rule, this can also be written as:

$$\mathrm{d}f = \frac{\mathrm{d}f}{\mathrm{d}y}\mathrm{d}y$$

and so the notation is consistent.

3.17.2 Functions of several variables

Let's have $\mathbf{x} = (x_1, x_2, \ldots, x_N)$. The function $f(\mathbf{x})$ assigns a number to each $\mathbf{x}$. We define a differential of f in the direction of $\mathbf{h}$ as:

$$\mathrm{d}f \equiv \frac{\mathrm{d}}{\mathrm{d}\varepsilon} f(\mathbf{x} + \varepsilon\mathbf{h})\bigg|_{\varepsilon=0} = \lim_{\varepsilon\to 0} \frac{f(\mathbf{x} + \varepsilon\mathbf{h}) - f(\mathbf{x})}{\varepsilon} = \mathbf{a}\cdot\mathbf{h}$$

The last equality follows from the fact, that $\frac{\mathrm{d}}{\mathrm{d}\varepsilon} f(\mathbf{x} + \varepsilon\mathbf{h})\big|_{\varepsilon=0}$ is a linear function of $\mathbf{h}$. We define the partial derivative $\frac{\partial f}{\partial x_i}$ of f with respect to x_i as the i-th component of the vector $\mathbf{a}$:

$$\mathbf{a} \equiv \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_N}\right) \equiv \nabla f$$

This also gives a formula for computing $\frac{\partial f}{\partial x_i}$: we set $h_j = \delta_{ij} h_i$ and

$$\frac{\partial f}{\partial x_i} = a_i = \mathbf{a}\cdot\mathbf{h} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} f(\mathbf{x} + \varepsilon(0, 0, \ldots, 1, \ldots, 0))\bigg|_{\varepsilon=0} =$$
$$= \lim_{\varepsilon\to 0} \frac{f(x_1, x_2, \ldots, x_i + \varepsilon, \ldots, x_N) - f(x_1, x_2, \ldots, x_i, \ldots, x_N)}{\varepsilon}$$

The usual way to define partial derivatives is to use the last formula as the definition, but here this formula is a consequence of our definition in terms of the components of $\mathbf{a}$. Every variable can be treated as a function (very simple one):

$$x_i = g(x_1, \ldots, x_N) = \delta_{ij} x_j$$

and so we define

$$\mathrm{d}x_i \equiv \mathrm{d}g = \mathrm{d}(\delta_{ij} x_j) = h_i$$

and thus we write $h_i = \mathrm{d}x_i$ and $\mathbf{h} = \mathrm{d}\mathbf{x}$ and

$$\mathrm{d}f = \frac{\mathrm{d}f}{\mathrm{d}x_i}\mathrm{d}x_i = (\nabla f)\cdot\mathrm{d}\mathbf{x}$$

So $\mathrm{d}\mathbf{x}$ has two meanings — it's either $\mathbf{h} = \mathbf{x} - \mathbf{x}_0$ (a finite change in the independent variable $\mathbf{x}$) or a differential, depending on the context. The above is a detailed explanation why things are defined the way they are and what

the exact meaning is. With this understanding, the only things that are actually needed for any calculations are the following – the definition of a differential:

$$\mathrm{d}f = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} f(\mathbf{x} + \varepsilon \mathrm{d}\mathbf{x})\bigg|_{\varepsilon=0}$$

Only a regular derivative (defined in the previous section) is needed for this definition. The definition of a partial derivative (and a gradient):

$$\mathrm{d}f = \frac{\mathrm{d}f}{\mathrm{d}x_i}\mathrm{d}x_i = (\nabla f)\cdot \mathrm{d}\mathbf{x}$$

And finally the understanding that $\mathrm{d}\mathbf{x}$ means either $\mathbf{h} = \mathbf{x} - \mathbf{x}_0$ or a differential depending on the context. That's all there is to it.

3.17.3 Functionals

Let's now define functional derivatives and variations. Functional $F[f]$ assigns a number to each function $f(x)$. The variation is defined as

$$\delta F[f] \equiv \frac{\mathrm{d}}{\mathrm{d}\varepsilon} F[f + \varepsilon h]\bigg|_{\varepsilon=0} = \lim_{\epsilon\to 0}\frac{F[f+\epsilon h] - F[f]}{\epsilon} = \int a(x)h(x)\mathrm{d}x$$

We define $\frac{\delta F}{\delta f(x)}$ as

$$a(x) \equiv \frac{\delta F}{\delta f(x)}$$

This also gives a formula for computing $\frac{\delta F}{\delta f(x)}$: we set $h(y) = \delta(x-y)$ and

$$\begin{aligned}\frac{\delta F}{\delta f(x)} = a(x) = \int a(y)\delta(x-y)\mathrm{d}y = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} F[f(y) + \varepsilon\delta(x-y)]\bigg|_{\varepsilon=0} = \\ = \lim_{\varepsilon\to 0}\frac{F[f(y)+\varepsilon\delta(x-y)] - F[f(y)]}{\varepsilon}\end{aligned} \tag{3.17.3.1}$$

Sometimes the functional derivative is defined using the last formula, here this formula just follows from our definition. Every function can be treated as a functional (although a very simple one):

$$f(x) = G[f] = \int f(y)\delta(x-y)\mathrm{d}y$$

and so we define

$$\delta f \equiv \delta G[f] = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} G[f(x) + \varepsilon h(x)]\bigg|_{\varepsilon=0} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(f(x) + \varepsilon h(x))\bigg|_{\varepsilon=0} = h(x)$$

thus we write $h = \delta f$ and

$$\delta F[f] = \int \frac{\delta F}{\delta f(x)}\delta f(x)\mathrm{d}x$$

so δf have two meanings — it's either $h(x) = f(x) - f_0(x)$ (a finite change in the function f) or a variation of a functional, depending on the context. It is completely analogous to $\mathrm{d}\mathbf{x}$. Let's summarize the only formulas needed in actual calculations – the definition of a variation (using a regular derivative):

$$\delta F[f] = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} F[f + \varepsilon\delta f]\bigg|_{\varepsilon=0} \tag{3.17.3.2}$$

the definition of the functional derivative:

$$\delta F[f] = \int \frac{\delta F}{\delta f(x)} \delta f(x) \mathrm{d}x$$

and the understanding that δf means either $h(x) = f(x) - f_0(x)$ or a variation. The last equation is the best way to calculate functional derivative — apply δ variation, until you get the integral into the form $\int (\ ) \delta f(x) \mathrm{d}x$ and then you read off the functional derivative from the expression in the parentheses.

The correspondence between the finite and infinite dimensional case can be summarized using a functional $F[n]$, function $n(\mathbf{x})$ of continuous parameter $\mathbf{x}$ (which can be a scalar or a vector) and its discretized version $n_i \equiv n(\mathbf{x}_i)$, together with a function $F(n)$:

$$\begin{aligned} i &\iff \mathbf{x} \\ n_i &\iff n(\mathbf{x}) \\ F(n_i) &\iff F[n] \\ \mathrm{d}F = 0 &\iff \delta F = 0 \\ \frac{\partial F}{\partial n_i} = 0 &\iff \frac{\delta F}{\delta n(\mathbf{x})} = 0 \end{aligned}$$

In other words, the basic difference is that the continuous parameter $\mathbf{x}$ has been replaced with a discrete parameter i. Then the function $n(\mathbf{x})$ becomes a vector of values n_i, variation becomes a differential and functional derivative becomes a partial derivative. To minimize a functional, one must search for zero functional derivative, while in the discrete case one searches for zero partial derivatives (gradient).

We now extend the δ-variation notation to any any function g which contains the function $f(x)$ being varied, you just need to replace f by $f + \epsilon \delta f$ and apply $\frac{\mathrm{d}}{\mathrm{d}\epsilon}$ to the whole g, for example (here $g = \partial_\mu \phi$ and $f = \phi$):

$$\delta \partial_\mu \phi = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \partial_\mu (\phi + \varepsilon \delta \phi) \bigg|_{\varepsilon=0} = \partial_\mu \frac{\mathrm{d}}{\mathrm{d}\varepsilon} (\phi + \varepsilon \delta \phi) \bigg|_{\varepsilon=0} = \partial_\mu \delta \phi$$

As such, the F in (3.17.3.2) can be either a functional or any expression that contains the function f. This notation allows us a very convenient computation, as shown in the following examples.

First, when computing a variation of some integral, we can interchange δ and $\int$:

$$F[f] = \int K(x) f(x) \mathrm{d}x$$

$$\delta F = \delta \int K(x) f(x) \mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \int K(x)(f + \varepsilon h) \mathrm{d}x \bigg|_{\varepsilon=0} = \int \frac{\mathrm{d}}{\mathrm{d}\varepsilon} (K(x)(f + \varepsilon h)) \mathrm{d}x \bigg|_{\varepsilon=0} =$$

$$= \int \delta(K(x) f(x)) \mathrm{d}x$$

In the expression $\delta(K(x)f(x))$ we must understand from the context if we are treating it as a functional of f or K. In our case it's a functional of f, so we have $\delta(Kf) = K\delta f$.

The second very important note is when taking variation of expression like:

$$\begin{aligned}
\delta \int f(t_1)f(t_2)\mathrm{d}t_1\mathrm{d}t_2 = \\
= \int \delta(f(t_1)f(t_2))\mathrm{d}t_1\mathrm{d}t_2 = \\
= \int \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(f(t_1)+\varepsilon\delta f(t_1))(f(t_2)+\varepsilon\delta f(t_2))\Big|_{\varepsilon=0}\mathrm{d}t_1\mathrm{d}t_2 = \\
= \int (\delta f(t_1))f(t_2)+f(t_1)(\delta f(t_2))\mathrm{d}t_1\mathrm{d}t_2 = \\
= \int (\delta f(t_1))f(t_2)+f(t_2)(\delta f(t_1))\mathrm{d}t_1\mathrm{d}t_2 = \\
= 2\int f(t_2)\delta f(t_1)\mathrm{d}t_1\mathrm{d}t_2
\end{aligned}$$

then when f is replaced by $f+\epsilon\delta f$, one has to keep track of the independent variable, so $f(t_1)$ gets replaced by $f(t_1)+\epsilon\delta f(t_1)$ and $f(t_2)$ gets replaced by $f(t_2)+\epsilon\delta f(t_2)$. Thus the two variations $\delta f(t_1)$ and $\delta f(t_2)$ are different (independent). If there is only one indepenent variable, one can simply write δf as it is clear what the independent variable is. This is analogous to using differentials, e.g. $\mathrm{d}(f(x)f(y)) = (\mathrm{d}f(x))f(y)+f(x)\mathrm{d}f(y) = f'(x)\mathrm{d}xf(y)+f(x)f'(y)\mathrm{d}y$, where one has to keep track of the independent variable as well for each $\mathrm{d}f$.

Another useful formula is differentiation of a functional $F[\psi(\theta)]$ where the function $\psi(\theta)$ depends on a parameter θ:

$$\frac{\mathrm{d}F[\psi(\theta)]}{\mathrm{d}\theta} = \frac{\mathrm{d}}{\mathrm{d}\epsilon}F[\psi(\theta+\epsilon)]\Big|_{\epsilon=0} = \frac{\mathrm{d}}{\mathrm{d}\epsilon}F\left[\psi(\theta)+\epsilon\frac{d\psi(\theta)}{d\theta}+O(\epsilon^2)\right]\Big|_{\epsilon=0} = \frac{\mathrm{d}}{\mathrm{d}\epsilon}F\left[\psi(\theta)+\epsilon\frac{d\psi(\theta)}{d\theta}\right]\Big|_{\epsilon=0} = \int \frac{\delta F[\psi]}{\delta\psi}\frac{\mathrm{d}\psi(\theta)}{\mathrm{d}\theta}\mathrm{d}x$$

where we used the definition of a variation and a functional derivative with $\delta\psi = \frac{\mathrm{d}\psi(\theta)}{\mathrm{d}\theta}$:

$$\delta F = \frac{\mathrm{d}}{\mathrm{d}\epsilon}F[\psi+\epsilon\delta\psi]\Big|_{\epsilon=0} = \int \frac{\delta F}{\delta\psi}\delta\psi\mathrm{d}x$$

3.17.4 Examples

Some of these examples show how to use the delta function definition of the functional derivative in equation (3.17.3.1). However, the simplest way is to calculate variation first and then read off the functional derivative from the result, as explained above.

$$\frac{\delta}{\delta f(t)}\int \mathrm{d}t' f(t')g(t') = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\int \mathrm{d}t'(f(t')+\varepsilon\delta(t-t'))g(t')\Big|_{\varepsilon=0} = g(t)$$

$$\delta\int \mathrm{d}t f(t)g(t) = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\int \mathrm{d}t(f(t)+\varepsilon\delta f(t))g(t)\Big|_{\varepsilon=0} = \int g(t)(\delta f(t))\mathrm{d}t$$

$$\frac{\delta f(t')}{\delta f(t)} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(f(t')+\varepsilon\delta(t-t'))\Big|_{\varepsilon=0} = \delta(t-t')$$

$$\frac{\delta f(t_1)f(t_2)}{\delta f(t)} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(f(t_1)+\varepsilon\delta(t-t_1))(f(t_2)+\varepsilon\delta(t-t_2))\Big|_{\varepsilon=0} = \delta(t-t_1)f(t_2)+f(t_1)\delta(t-t_2)$$

The next example shows that when taking variation of an expression containing the function f of different independent variables, one has to keep track of these variables in the variations:

$$\delta(f(t_1)f(t_2)) = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(f(t_1)+\varepsilon\delta f(t_1))(f(t_2)+\varepsilon\delta f(t_2))\Big|_{\varepsilon=0} = (\delta f(t_1))f(t_2)+f(t_1)(\delta f(t_2))$$

$$\frac{\delta}{\delta f(t)}\tfrac{1}{2}\int \mathrm{d}t_1\mathrm{d}t_2 K(t_1,t_2)f(t_1)f(t_2) = \tfrac{1}{2}\int \mathrm{d}t_1\mathrm{d}t_2 K(t_1,t_2)\frac{\delta f(t_1)f(t_2)}{\delta f(t)} =$$
$$= \tfrac{1}{2}\left(\int \mathrm{d}t_1 K(t_1,t)f(t_1) + \int \mathrm{d}t_2 K(t,t_2)f(t_2)\right) = \int \mathrm{d}t_2 K(t,t_2)f(t_2)$$

$$\delta\tfrac{1}{2}\int \mathrm{d}t_1\mathrm{d}t_2 K(t_1,t_2)f(t_1)f(t_2) = \tfrac{1}{2}\int \mathrm{d}t_1\mathrm{d}t_2 K(t_1,t_2)(\delta f(t_1)f(t_2)) =$$
$$= \tfrac{1}{2}\int \mathrm{d}t_1\mathrm{d}t_2 K(t_1,t_2)((\delta f(t_1))f(t_2) + f(t_1)(\delta f(t_2)) =$$
$$= \int \mathrm{d}t_1\mathrm{d}t_2 K(t_1,t_2)f(t_2)(\delta f(t_1))$$

The last equality follows from $K(t_1,t_2) = K(t_2,t_1)$ (any antisymmetrical part of a K would not contribute to the symmetrical integration).

Another example is the derivation of Euler-Lagrange equations for the Lagrangian density $\mathcal{L} = \mathcal{L}(\eta_\rho, \partial_\nu\eta_\rho, x^\nu)$:

$$0 = \delta S = \delta\int \mathcal{L}\,\mathrm{d}^4x^\mu = \int \delta\mathcal{L}\,\mathrm{d}^4x^\mu = \int \frac{\partial\mathcal{L}}{\partial\eta_\rho}\delta\eta_\rho + \frac{\partial\mathcal{L}}{\partial(\partial_\nu\eta_\rho)}\delta(\partial_\nu\eta_\rho)\,\mathrm{d}^4x^\mu =$$
$$= \int \frac{\partial\mathcal{L}}{\partial\eta_\rho}\delta\eta_\rho + \frac{\partial\mathcal{L}}{\partial(\partial_\nu\eta_\rho)}\partial_\nu(\delta\eta_\rho)\,\mathrm{d}^4x^\mu =$$
$$= \int \frac{\partial\mathcal{L}}{\partial\eta_\rho}\delta\eta_\rho - \partial_\nu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\nu\eta_\rho)}\right)\delta\eta_\rho\,\mathrm{d}^4x^\mu + \int \partial_\nu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\nu\eta_\rho)}\delta\eta_\rho\right)\mathrm{d}^4x^\mu =$$
$$= \int \left[\frac{\partial\mathcal{L}}{\partial\eta_\rho} - \partial_\nu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\nu\eta_\rho)}\right)\right]\delta\eta_\rho\,\mathrm{d}^4x^\mu$$

We can also write it using a functional derivative $\frac{\delta S}{\delta\eta_\rho}$ as:

$$\frac{\delta S}{\delta\eta_\rho} = \frac{\partial\mathcal{L}}{\partial\eta_\rho} - \partial_\nu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\nu\eta_\rho)}\right)$$

Another example:

$$\frac{\delta}{\delta f(t)}\int f^3(x)\mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\int (f(x)+\varepsilon\delta(x-t))^3\mathrm{d}x\Big|_{\varepsilon=0} =$$
$$= \int 3(f(x)+\varepsilon\delta(x-t))^2\delta(x-t)\mathrm{d}x\Big|_{\varepsilon=0} = \int 3f^2(x)\delta(x-t)\mathrm{d}x = 3f^2(t)$$

One might think that the above calculation is incorrect, because $\delta^2(x-t)$ is undefined. In case of such problems the above notation automatically implies working with some sequence $\delta_\alpha(x) \to \delta(x)$ (for example $\delta_\alpha(x) = \frac{1}{\pi x}\sin(\alpha x)$) and taking the limit $\alpha\to\infty$:

$$\frac{\delta}{\delta f(t)}\int f^3(x)\mathrm{d}x = \lim_{\alpha\to\infty}\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\int (f(x)+\varepsilon\delta_\alpha(x-t))^3\mathrm{d}x\Big|_{\varepsilon=0} =$$
$$= \lim_{\alpha\to\infty}\int 3(f(x)+\varepsilon\delta_\alpha(x-t))^2\delta_\alpha(x-t)\mathrm{d}x\Big|_{\varepsilon=0} = \lim_{\alpha\to\infty}\int 3f^2(x)\delta_\alpha(x-t)\mathrm{d}x =$$
$$= \int 3f^2(x)\lim_{\alpha\to\infty}\delta_\alpha(x-t)\mathrm{d}x = \int 3f^2(x)\delta(x-t)\mathrm{d}x = 3f^2(t) \qquad (3.17.4.1)$$

As you can see, we got the same result, with the same rigor, but using an obfuscating notation. That's why such obvious manipulations with δ_α are tacitly implied. However, the best method is to first calculate the variation:

$$\delta\int f^3(x)\mathrm{d}x = \int \delta f^3(x)\mathrm{d}x = \int 3f^2(x)\delta f(x)\mathrm{d}x$$

and immediately read off the functional derivative:

$$\frac{\delta}{\delta f(t)} \int f^3(x) \mathrm{d}x = 3f^2(t)$$

Another example with a metric as a function of coordinates $g_{\mu\nu} = g_{\mu\nu}(x^\mu)$:

$$\delta g_{\mu\nu} = \delta g_{\mu\nu}(x^\mu) = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} g_{\mu\nu}(x^\mu + \varepsilon(\delta x^\mu))\bigg|_{\varepsilon=0} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(x^\sigma + \varepsilon(\delta x^\sigma))\bigg|_{\varepsilon=0} \partial_\sigma g_{\mu\nu} = (\delta x^\sigma)\partial_\sigma g_{\mu\nu}$$

And an example of varying with respect to a metric:

$$\begin{aligned}
\delta\sqrt{|\det g_{\mu\nu}|} = \sqrt{|\det g_{\mu\nu}|}\,\delta \log\sqrt{|\det g_{\mu\nu}|} = \tfrac{1}{2}\sqrt{|\det g_{\mu\nu}|}\,\delta\log|\det g_{\mu\nu}| = \\
= \tfrac{1}{2}\sqrt{|\det g_{\mu\nu}|}\,\delta\,\mathrm{Tr}\log g_{\mu\nu} = \tfrac{1}{2}\sqrt{|\det g_{\mu\nu}|}\;\mathrm{Tr}\,\delta\log g_{\mu\nu} = \\
= \tfrac{1}{2}\sqrt{|\det g_{\mu\nu}|}\;\mathrm{Tr}\,g^{\mu\nu}\delta g_{\mu\nu} = \tfrac{1}{2}\sqrt{|\det g_{\mu\nu}|}\,g^{\mu\nu}\delta g_{\mu\nu} = \\
= -\tfrac{1}{2}\sqrt{|\det g_{\mu\nu}|}\,g_{\mu\nu}\delta g^{\mu\nu}
\end{aligned}$$

Another example (varying energy functional):

$$\begin{aligned}
E[\rho] = 4\pi\int \frac{a\rho(r)}{b + r_s(r)} r^2 \mathrm{d}r \\
r_s(r) = \left(\frac{3}{4\pi(-\rho)}\right)^{\frac{1}{3}} \\
\frac{\mathrm{d}r_s}{\mathrm{d}\rho} = \frac{1}{3}\left(\frac{3}{4\pi(-\rho)}\right)^{-\frac{2}{3}} \frac{3}{4\pi\rho^2} = -\frac{1}{3\rho}\left(\frac{3}{4\pi(-\rho)}\right)^{\frac{1}{3}} = -\frac{r_s}{3\rho} \\
\delta E[\rho] = 4\pi\delta\int \frac{a\rho}{b + r_s} r^2\mathrm{d}r = \\
= 4\pi\int\left(\frac{a\delta\rho}{b+r_s} - \frac{a\rho}{(b+r_s)^2}\delta r_s\right) r^2\mathrm{d}r = \\
= 4\pi\int\left(\frac{a\delta\rho}{b+r_s} - \frac{a\rho}{(b+r_s)^2}\left(-\frac{r_s}{3\rho}\right)\delta\rho\right) r^2\mathrm{d}r = \\
= 4\pi\int\left(\frac{a}{b+r_s} + \frac{1}{3}\frac{a r_s}{(b+r_s)^2}\right)(\delta\rho) r^2\mathrm{d}r \\
\frac{\delta E[\rho]}{\delta\rho} = 4\pi r^2\left(\frac{a}{b+r_s} + \frac{1}{3}\frac{a r_s}{(b+r_s)^2}\right)
\end{aligned}$$

Another example (Hartree energy):

$$E[n] = \tfrac{1}{2}\int \frac{n(\mathbf{r}')n(\mathbf{r}'')}{|\mathbf{r}' - \mathbf{r}''|}\mathrm{d}^3r'\mathrm{d}^3r''$$

we calculate the variation first:

$$\begin{aligned}
\delta E[n] = \tfrac{1}{2}\delta\int\frac{n(\mathbf{r}')n(\mathbf{r}'')}{|\mathbf{r}'-\mathbf{r}''|}\mathrm{d}^3r'\mathrm{d}^3r'' = \\
= \tfrac{1}{2}\int\frac{(\delta n(\mathbf{r}'))n(\mathbf{r}'') + n(\mathbf{r}')(\delta n(\mathbf{r}''))}{|\mathbf{r}'-\mathbf{r}''|}\mathrm{d}^3r'\mathrm{d}^3r'' = \\
= \int\frac{n(\mathbf{r}')}{|\mathbf{r}'-\mathbf{r}''|}(\delta n(\mathbf{r}''))\mathrm{d}^3r'\mathrm{d}^3r'' = \\
= \int\frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}(\delta n(\mathbf{r}))\mathrm{d}^3r'\mathrm{d}^3r
\end{aligned}$$

so the functional derivative is:

$$\frac{\delta E[n]}{\delta n(\mathbf{r})} = \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \mathrm{d}^3 r'$$

Another example (functional with gradients):

$$F[n] = \int h(n) \frac{|\nabla n|^2}{n} \mathrm{d}^3 r$$

the variation is:

$$\begin{aligned}
\delta F[n] = \int \delta \left(h(n) \frac{|\nabla n|^2}{n} \right) \mathrm{d}^3 r = \\
= \int \frac{\mathrm{d}h}{\mathrm{d}n} \frac{|\nabla n|^2}{n} \delta n + h(n) \left(\frac{2n \nabla n \cdot \nabla \delta n - |\nabla n|^2 \delta n}{n^2} \right) \mathrm{d}^3 r = \\
= \int \left(\frac{\mathrm{d}h}{\mathrm{d}n} - \frac{h(n)}{n} \right) \frac{|\nabla n|^2}{n} \delta n + 2h(n) \frac{\nabla n \cdot \nabla \delta n}{n} \mathrm{d}^3 r = \\
= \int \left(\frac{\mathrm{d}h}{\mathrm{d}n} - \frac{h(n)}{n} \right) \frac{|\nabla n|^2}{n} \delta n - 2\nabla \cdot \left(h(n) \frac{\nabla n}{n} \right) \delta n \, \mathrm{d}^3 r = \\
= \int \left(\frac{\mathrm{d}h}{\mathrm{d}n} - \frac{h(n)}{n} \right) \frac{|\nabla n|^2}{n} \delta n - 2 \left(\frac{\mathrm{d}h}{\mathrm{d}n} \frac{|\nabla n|^2}{n} + h(n) \frac{\nabla^2 n}{n} - h(n) \frac{|\nabla n|^2}{n^2} \right) \delta n \, \mathrm{d}^3 r = \\
= \int \left[\left(\frac{h(n)}{n} - \frac{\mathrm{d}h}{\mathrm{d}n} \right) \frac{|\nabla n|^2}{n} - 2h(n) \frac{\nabla^2 n}{n} \right] \delta n \, \mathrm{d}^3 r
\end{aligned}$$

from which we read off the functional derivative:

$$\frac{\delta F[n]}{\delta n(\mathbf{r})} = \left(\frac{h(n)}{n} - \frac{\mathrm{d}h}{\mathrm{d}n} \right) \frac{|\nabla n|^2}{n} - 2h(n) \frac{\nabla^2 n}{n}$$

3.18 Dirac Notation

The Dirac notation allows a very compact and powerful way of writing equations that describe a function expansion into a basis, both discrete (e.g. a Fourier series expansion) and continuous (e.g. a Fourier transform) and related things. The notation is designed so that it is very easy to remember and it just guides you to write the correct equation.

Let's have a function $f(x)$. We define

$$\begin{aligned}
\langle x|f\rangle &\equiv f(x) \\
\langle x'|f\rangle &\equiv f(x') \\
\langle x'|x\rangle &\equiv \delta(x' - x) \\
\int |x\rangle \langle x| \, \mathrm{d}x &\equiv \mathbb{1}
\end{aligned}$$

The following equation

$$f(x') = \int \delta(x' - x) f(x) \mathrm{d}x$$

then becomes

$$\langle x'|f\rangle = \int \langle x'|x\rangle \langle x|f\rangle \, \mathrm{d}x$$

and thus we can interpret $|f\rangle$ as a vector, $|x\rangle$ as a basis and $\langle x|f\rangle$ as the coefficients in the basis expansion:

$$|f\rangle = \mathbb{1}\,|f\rangle = \int |x\rangle\,\langle x|\,\mathrm{d}x\,|f\rangle = \int |x\rangle\,\langle x|f\rangle\,\mathrm{d}x$$

That's all there is to it. Take the above rules as the operational definition of the Dirac notation. It's like with the delta function - written alone it doesn't have any meaning, but there are clear and non-ambiguous rules to convert any expression with δ to an expression which even mathematicians understand (i.e. integrating, applying test functions and using other relations to get rid of all δ symbols in the expression – but the result is usually much more complicated than the original formula). It's the same with the ket $|f\rangle$: written alone it doesn't have any meaning, but you can always use the above rules to get an expression that make sense to everyone (i.e. attaching any bra to the left and rewriting all brackets $\langle a|b\rangle$ with their equivalent expressions) – but it will be more complex and harder to remember and – that is important – less general.

Now, let's look at the spherical harmonics:

$$Y_{lm}(\hat{\mathbf{r}}) \equiv \langle \hat{\mathbf{r}}|lm\rangle$$

on the unit sphere, we have

$$\int |\hat{\mathbf{r}}\rangle\,\langle\hat{\mathbf{r}}|\,\mathrm{d}\hat{\mathbf{r}} = \int |\hat{\mathbf{r}}\rangle\,\langle\hat{\mathbf{r}}|\,\mathrm{d}\Omega = \mathbb{1}$$

$$\delta(\hat{\mathbf{r}} - \hat{\mathbf{r}}') = \langle\hat{\mathbf{r}}|\hat{\mathbf{r}}'\rangle$$

thus

$$\int_0^{2\pi}\int_0^{\pi} Y_{lm}(\theta,\phi)\,Y^*_{l'm'}(\theta,\phi)\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi = \int \langle l'm'|\hat{\mathbf{r}}\rangle\,\langle\hat{\mathbf{r}}|lm\rangle\,\mathrm{d}\Omega = \langle l'm'|lm\rangle$$

and from (3.30.1) we get

$$\langle l'm'|lm\rangle = \delta_{mm'}\delta_{ll'}$$

now

$$\sum_{lm} Y_{lm}(\theta,\phi)Y^*_{lm}(\theta',\phi') = \sum_{lm}\langle\hat{\mathbf{r}}|lm\rangle\,\langle lm|\hat{\mathbf{r}}'\rangle$$

from (3.30.3) we get

$$\sum_{lm}\langle\hat{\mathbf{r}}|lm\rangle\,\langle lm|\hat{\mathbf{r}}'\rangle = \langle\hat{\mathbf{r}}|\hat{\mathbf{r}}'\rangle$$

so we have

$$\sum_{lm}|lm\rangle\,\langle lm| = \mathbb{1}$$

so $|lm\rangle$ forms an orthonormal basis. Any function defined on the sphere $f(\hat{\mathbf{r}})$ can be written using this basis:

$$f(\hat{\mathbf{r}}) = \langle\hat{\mathbf{r}}|f\rangle = \sum_{lm}\langle\hat{\mathbf{r}}|lm\rangle\,\langle lm|f\rangle = \sum_{lm} Y_{lm}(\hat{\mathbf{r}})f_{lm}$$

where

$$f_{lm} = \langle lm|f\rangle = \int \langle lm|\hat{\mathbf{r}}\rangle\,\langle\hat{\mathbf{r}}|f\rangle\,\mathrm{d}\Omega = \int Y^*_{lm}(\hat{\mathbf{r}})f(\hat{\mathbf{r}})\mathrm{d}\Omega$$

If we have a function $f(\mathbf{r})$ in 3D, we can write it as a function of ρ and $\hat{\mathbf{r}}$ and expand only with respect to the variable $\hat{\mathbf{r}}$:

$$f(\mathbf{r}) = f(\rho\hat{\mathbf{r}}) \equiv g(\rho,\hat{\mathbf{r}}) = \sum_{lm} Y_{lm}(\hat{\mathbf{r}})g_{lm}(\rho)$$

In Dirac notation we are doing the following: we decompose the space into the angular and radial part

$$|\mathbf{r}\rangle = |\hat{\mathbf{r}}\rangle \otimes |\rho\rangle \equiv |\hat{\mathbf{r}}\rangle\,|\rho\rangle$$

and write

$$f(\mathbf{r}) = \langle \mathbf{r}|f\rangle = \langle \hat{\mathbf{r}}|\,\langle \rho|f\rangle = \sum_{lm} Y_{lm}(\hat{\mathbf{r}})\,\langle lm|\,\langle \rho|f\rangle$$

where

$$\langle lm|\,\langle \rho|f\rangle = \int \langle lm|\hat{\mathbf{r}}\rangle\,\langle \hat{\mathbf{r}}|\,\langle \rho|f\rangle\,\mathrm{d}\Omega = \int Y^*_{lm}(\hat{\mathbf{r}}) f(\mathbf{r})\mathrm{d}\Omega$$

Let's calculate $\langle \rho|\rho'\rangle$

$$\langle \mathbf{r}|\mathbf{r}'\rangle = \langle \hat{\mathbf{r}}|\,\langle \rho|\rho'\rangle\,|\hat{\mathbf{r}}'\rangle = \langle \hat{\mathbf{r}}|\hat{\mathbf{r}}'\rangle\,\langle \rho|\rho'\rangle$$

so

$$\langle \rho|\rho'\rangle = \frac{\langle \mathbf{r}|\mathbf{r}'\rangle}{\langle \hat{\mathbf{r}}|\hat{\mathbf{r}}'\rangle} = \frac{\delta(\rho-\rho')}{\rho^2}$$

We must stress that $|lm\rangle$ only acts in the $|\hat{\mathbf{r}}\rangle$ space (not the $|\rho\rangle$ space) which means that

$$\langle \mathbf{r}|lm\rangle = \langle \hat{\mathbf{r}}|\,\langle \rho|lm\rangle = \langle \hat{\mathbf{r}}|lm\rangle\,\langle \rho| = Y_{lm}(\hat{\mathbf{r}})\,\langle \rho|$$

and $V\,|lm\rangle$ leaves $V\,|\rho\rangle$ intact. Similarly,

$$\sum_{lm} |lm\rangle\,\langle lm| = \mathbb{1}$$

is a unity in the $|\hat{\mathbf{r}}\rangle$ space only (i.e. on the unit sphere).

Let's rewrite the equation (3.30.4):

$$\sum_{m} \langle \hat{\mathbf{r}}|lm\rangle\,\langle lm|\hat{\mathbf{r}}'\rangle = \frac{4\pi}{2l+1}\,\langle \hat{\mathbf{r}}\cdot\hat{\mathbf{r}}'|P_l\rangle$$

Using the completeness relation (3.29.1):

$$\sum_{l} \frac{2l+1}{2}\,\langle x'|P_l\rangle\,\langle P_l|x\rangle = \langle x'|x\rangle$$

$$\sum_{l} |P_l\rangle\,\frac{2l+1}{2}\,\langle P_l| = \mathbb{1}$$

we can now derive a very important formula true for every function $f(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}')$:

$$f(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}') = \langle \hat{\mathbf{r}}\cdot\hat{\mathbf{r}}'|f\rangle = \sum_{l} \langle \hat{\mathbf{r}}\cdot\hat{\mathbf{r}}'|P_l\rangle\,\frac{2l+1}{2}\,\langle P_l|f\rangle = \sum_{lm} \langle \hat{\mathbf{r}}|lm\rangle\,\langle lm|\hat{\mathbf{r}}'\rangle\,\frac{(2l+1)^2}{8\pi}\,\langle P_l|f\rangle =$$

$$= \sum_{lm} \langle \hat{\mathbf{r}}|lm\rangle\, f_l\,\langle lm|\hat{\mathbf{r}}'\rangle$$

where

$$f_l = \frac{(2l+1)^2}{8\pi}\,\langle P_l|f\rangle = \frac{(2l+1)^2}{8\pi}\int_{-1}^{1} \langle P_l|x\rangle\,\langle x|f\rangle\,\mathrm{d}x = \frac{(2l+1)^2}{8\pi}\int_{-1}^{1} P_l(x) f(x)\mathrm{d}x$$

or written explicitly

$$f(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}') = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} Y_{lm}(\hat{\mathbf{r}}) f_l Y^*_{lm}(\hat{\mathbf{r}}') \tag{3.18.1}$$

3.19 Homogeneous Functions (Euler's Theorem)

A function of several variables $f(x_1, x_2, \dots) \equiv f(x_i)$ is homogeneous of degree k if

$$f(\lambda x_i) = \lambda^k f(x_i)$$

By differentiating with respect to λ:

$$x_i \frac{\partial f(\lambda x_i)}{\partial x_i} = k\lambda^{k-1} f(x_i)$$

and setting $\lambda = 1$ we get the so called Euler equation:

$$x_i \frac{\partial f(x_i)}{\partial x_i} = k f(x_i)$$

in 3D this can also be written as:

$$\mathbf{x} \cdot \nabla f(\mathbf{x}) = k f(\mathbf{x})$$

3.19.1 Example 1

The function $f(x, y, z) = \frac{xy}{z}$ is homogeneous of degree 1, because:

$$f(\lambda x, \lambda y, \lambda z) = \frac{\lambda x \lambda y}{\lambda z} = \lambda \frac{xy}{z} = \lambda f(x, y, z)$$

and the Euler equation is:

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} + z\frac{\partial f}{\partial z} = f$$

or

$$x\frac{y}{z} + y\frac{x}{z} + z\left(-\frac{xy}{z^2}\right) = \frac{xy}{z}$$

Which is true.

3.19.2 Example 2

The function $V(r) = -\frac{Z}{r}$ is homogeneous of degree -1, because:

$$V(\lambda r) = -\frac{Z}{\lambda r} = \lambda^{-1} V(r)$$

and the Euler equation is:

$$r\frac{\mathrm{d}V}{\mathrm{d}r} = -V$$

or

$$r\frac{Z}{r^2} = -\left(-\frac{Z}{r}\right)$$

Which is true.

3.20 Green Functions

Green functions are an excellent tool for working with a solution to any ODE or PDE. In this text we explain how it works and then show how one can calculate them using FEM.

3.20.1 Introduction

Let's put any ODE or PDE in the form:

$$Lu(x) = f(x) \tag{3.20.1.1}$$

Here L is a differential operator and x can have any dimension, e.g. 1D (ODE), 2D, 3D or more (PDE). Then we can express the solution as

$$u(x) = L^{-1}f(x) = \int G(x, x')f(x')\mathrm{d}x' \tag{3.20.1.2}$$

where $G(x, x')$ is a Green function, that needs to satisfy the equation:

$$LG(x, x') = \delta(x - x') \tag{3.20.1.3}$$

Remember, that L acts on x only, so we can check, that (3.20.1.2) indeed solves the PDE (3.20.1.1):

$$Lu(x) = L\int G(x, x')f(x')\mathrm{d}x' = \int LG(x, x')f(x')\mathrm{d}x' = \int \delta(x - x')f(x')\mathrm{d}x' = f(x)$$

3.20.2 Boundary Conditions

The equation (3.20.1.3) doesn't determine the Green function uniquely, because one can add to it any solution of the homogeneous equation $Lu(x) = 0$. We can use this freedom to solve (3.20.1.3) for any boundary condition. So we prescribe a boundary condition and find the Green function (by solving (3.20.1.3)) that satisfies the boundary condition. It can be shown, that $u(x)$ determined from (3.20.1.2) then also needs to satisfy the same boundary condition.

3.20.3 Symmetry

We write the equation for Green functions at two different points x_1 and x_2:

$$LG(x, x_1) = \delta(x - x_1)$$
$$LG(x, x_2) = \delta(x - x_2)$$

and multiply the first equation by $G(x, x_2)$, second by $G(x, x_1)$:

$$G(x, x_2)LG(x, x_1) = \delta(x - x_1)G(x, x_2)$$
$$G(x, x_1)LG(x, x_2) = \delta(x - x_2)G(x, x_1)$$

substract them and integrate over x:

$$G(x, x_2)LG(x, x_1) - G(x, x_1)LG(x, x_2) = \delta(x - x_1)G(x, x_2) - \delta(x - x_2)G(x, x_1)$$
$$\int (G(x, x_2)LG(x, x_1) - G(x, x_1)LG(x, x_2))\,\mathrm{d}x = \int (\delta(x - x_1)G(x, x_2) - \delta(x - x_2)G(x, x_1))\,\mathrm{d}x$$
$$\int (G(x, x_2)LG(x, x_1) - G(x, x_1)LG(x, x_2))\,\mathrm{d}x = G(x_1, x_2) - G(x_2, x_1)$$

Assuming that the operator L is Hermitean, we get:

$$\int ((LG(x,x_2))G(x,x_1) - G(x,x_1)LG(x,x_2))\,\mathrm{d}x = G(x_1,x_2) - G(x_2,x_1)$$
$$0 = G(x_1,x_2) - G(x_2,x_1)$$

So the Green function is symmetric for Hermitean operators L.

3.20.4 Examples

Poisson Equation in 1D

Poisson equation:

$$-\frac{\mathrm{d}^2}{\mathrm{d}x^2}u(x) = f(x)$$

We calculate the Green function using the Fourier transform:

$$-\frac{\partial^2}{\partial x^2}G(x,x') = \delta(x-x')$$
$$-(ik)^2\tilde{G}(k,x') = \frac{e^{ikx'}}{\sqrt{2\pi}}$$
$$\tilde{G}(k,x') = \frac{e^{ikx'}}{\sqrt{2\pi}\,k^2}$$
$$G(x,x') = -\tfrac{1}{2}(x-x')\operatorname{sign}(x-x') = -\tfrac{1}{2}(x-x')(2H(x-x')-1) = H(x-x')(x'-x) + \tfrac{1}{2}(x-x')$$

Check:

$$\frac{\partial}{\partial x}G(x,x') = \delta(x-x')(x'-x) + H(x-x')(-1) + \tfrac{1}{2} = -H(x-x') + \tfrac{1}{2}$$
$$\frac{\partial^2}{\partial x^2}G(x,x') = -\delta(x-x')$$

Then:

$$u(x) = \int G(x,x')f(x')\mathrm{d}x' = \int \left(H(x-x')(x'-x) + \tfrac{1}{2}(x-x')\right) f(x')\mathrm{d}x'$$

The green function can also be written using $x_< = \min(x,x')$ and $x_> = \max(x,x')$:

$$G(x,x') = H(x-x')(x'-x) + \tfrac{1}{2}(x-x') = \tfrac{1}{2}(x_< - x_>)$$

Radial Poisson Equation

Let's write $r_>$ and $r_<$ using the Heaviside step function:

$$r_> = \max(r,r') = \begin{cases} r & \text{for } r > r' \\ r' & \text{for } r < r' \end{cases} = H(r-r')r + H(r'-r)r' =$$
$$= H(r-r')r + (1-H(r-r'))r' = H(r-r')(r-r') + r'$$

and:

$$r_< = \min(r,r') = \begin{cases} r' & \text{for } r > r' \\ r & \text{for } r < r' \end{cases} = H(r-r')r' + H(r'-r)r =$$
$$= H(r-r')r' + (1-H(r-r'))r = H(r-r')(r'-r) + r$$

Then we can differentiate:

$$\frac{\partial}{\partial r} r_> = \delta(r-r')(r-r') + H(r-r') = H(r-r')$$
$$\frac{\partial^2}{\partial r^2} r_> = \delta(r-r')$$
$$\frac{\partial}{\partial r} r_< = \delta(r-r')(r'-r) - H(r-r') + 1 = 1 - H(r-r') = H(r'-r)$$
$$\frac{\partial^2}{\partial r^2} r_< = -\delta(r-r')$$

Given:

$$u(r) = \int_0^\infty \frac{f(r')r'^2}{r_>} \mathrm{d}r' \tag{3.20.4.1}$$

The Green function is

$$G(r,r') = \frac{r'^2}{r_>}$$

Let's differentiate:

$$\frac{\partial}{\partial r} G(r,r') = -\frac{r'^2}{r_>^2} \frac{\partial}{\partial r} r_> = -\frac{r'^2}{r_>^2} H(r-r') = -\frac{r'^2}{r^2} H(r-r')$$

and

$$\frac{\partial^2}{\partial r^2} G(r,r') = +\frac{2}{r}\frac{r'^2}{r^2} H(r-r') - \frac{r'^2}{r^2}\delta(r-r')$$

So we get:

$$-\frac{\partial^2}{\partial r^2} G(r,r') - \frac{2}{r}\frac{\partial}{\partial r} G(r,r') = -\frac{2}{r}\frac{r'^2}{r^2} H(r-r') + \frac{r'^2}{r^2}\delta(r-r') + \frac{2}{r}\frac{r'^2}{r^2} H(r-r') = \frac{r'^2}{r^2}\delta(r-r') = \delta(r-r')$$

So $u(r)$ from (3.20.4.1) is a solution to the radial Poisson equation:

$$-\frac{\mathrm{d}^2}{\mathrm{d}r^2} u(r) - \frac{2}{r}\frac{\mathrm{d}}{\mathrm{d}r} u(r) = f(r)$$

Helmholtz Equation in 1D

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}x^2} + 1\right) u(x) = f(x)$$
$$\left(\frac{\mathrm{d}^2}{\mathrm{d}x^2} + 1\right) G(x,x') = \delta(x-x')$$

with boundary conditions $u(0) = u(\frac{\pi}{2}) = 0$. We use the Fourier transform:

$$\left((ik)^2 + 1\right) \tilde{G}(k,x') = \frac{e^{ikx'}}{\sqrt{2\pi}}$$
$$\tilde{G}(k,x') = \frac{e^{ikx'}}{\sqrt{2\pi}(1-k^2)}$$
$$G(x,x') = \tfrac{1}{2}\operatorname{sign}(x-x')\sin(x-x')$$

Check:

$$\frac{\partial}{\partial x}G(x,x') = \delta(x-x')\sin(x-x') + \tfrac{1}{2}\operatorname{sign}(x-x')\cos(x-x') = \\ = \tfrac{1}{2}\operatorname{sign}(x-x')\cos(x-x')$$

$$\frac{\partial^2}{\partial x^2}G(x,x') = \delta(x-x')\cos(x-x') - \tfrac{1}{2}\operatorname{sign}(x-x')\sin(x-x') = \\ = -\tfrac{1}{2}\operatorname{sign}(x-x')\sin(x-x') + \delta(x-x')$$

$$\frac{\partial^2}{\partial x^2}G(x,x') + \frac{\partial}{\partial x}G(x,x') = \delta(x-x')$$

The general solution of the homogeneous equation is:

$$u(x) = C_1\sin x + C_2\cos x$$

so the general Green function is:

$$G(x,x') = \tfrac{1}{2}\operatorname{sign}(x-x')\sin(x-x') + C_1\sin(x+x') + C_2\cos(x+x')$$

Satisfying the boundary conditions (for all $x' \neq 0$):

$$G(0,x') = G(\frac{\pi}{2},x') = 0$$

we get:

$$C_1 = -\tfrac{1}{2}$$
$$C_2 = 0$$

and:

$$G(x,x') = \tfrac{1}{2}\operatorname{sign}(x-x')\sin(x-x') - \tfrac{1}{2}\sin(x+x') = \\ = -H(x'-x)\sin x\cos x' - H(x-x')\cos x\sin x' = \\ = \begin{cases} -\sin x\cos x' & x < x' \\ -\cos x\sin x' & x > x' \end{cases} = -\sin x_<\cos x_>$$

and

$$u(x) = \int G(x,x')f(x')\mathrm{d}x' = -\cos x\int_0^x f(x')\sin x'\mathrm{d}x' - \sin x\int_x^{\frac{\pi}{2}} f(x')\cos x'\mathrm{d}x'$$

To show that this really works, let's take for example $f(x) = 3\sin 2x$. Then

$$u(x) = -\cos x\int_0^x 3\sin 2x'\sin x'\mathrm{d}x' - \sin x\int_x^{\frac{\pi}{2}} 3\sin 2x'\cos x'\mathrm{d}x'$$

We can use SymPy to evaluate the integrals:

```
In [1]: u = -cos(x)*integrate(3*sin(2*y)*sin(y), (y, 0, x)) - \
    sin(x)*integrate(3*sin(2*y)*cos(y), (y, x, pi/2))

In [2]: u
Out[2]:
-(cos(x)*sin(2*x) - 2*cos(2*x)*sin(x))*cos(x) - (sin(x)*sin(2*x)
    + 2*cos(x)*cos(2*x))*sin(x)
```

(continues on next page)

(continued from previous page)

```
In [3]: simplify(u)
Out[3]:
     2                   2
- cos (x)*sin(2*x) - sin (x)*sin(2*x)

In [4]: trigsimp(_)
Out[4]: -sin(2*x)
```

And we get

$$u(x) = -\sin 2x$$

We can easily check, that $u'' + u = 3 \sin 2x$:

```
>>> u = -sin(2*x)
>>> u.diff(x, 2) + u
3*sin(2*x)
```

and since $f(x) = 3 \sin 2x$, we have verified, that $u(x) = -\sin 2x$ is the correct solution.

Poisson Equation in 2D

Let $\mathbf{x} = (x, y)$ and we want to solve:

$$\nabla^2 u(\mathbf{x}) = f(\mathbf{x})$$

So we have:

$$\nabla^2 G(\mathbf{x}, \mathbf{x}') = \delta(\mathbf{x} - \mathbf{x}')$$

The solution is:

$$G(\mathbf{x}, \mathbf{x}') = \frac{1}{2\pi} \log |\mathbf{x} - \mathbf{x}'| = \frac{1}{4\pi} \log |\mathbf{x} - \mathbf{x}'|^2 = \frac{1}{4\pi} \log((x - x')^2 + (y - y')^2)$$

Poisson Equation in 3D

$$\nabla^2 u(x) = f(x)$$
$$\nabla^2 G(x, x') = \delta(x - x')$$

with boundary condition $G(x) = 0$ at infinity. Then:

$$G(x, x') = -\frac{1}{4\pi} \frac{1}{|x - x'|}$$

and

$$u(x) = -\frac{1}{4\pi} \int \frac{f(x')}{|x - x'|} \mathrm{d}x'$$

Helmholtz Equation in 3D

$$(\nabla^2 + k^2)u(x) = f(x)$$
$$(\nabla^2 + k^2)G(x, x') = \delta(x - x')$$

with boundary condition $G(x) = 0$ at infinity. Then:

$$G(x, x') = -\frac{1}{4\pi}\frac{e^{ik|x-x'|}}{|x - x'|}$$
$$u(x) = -\frac{1}{4\pi}\int \frac{f(x')e^{ik|x-x'|}}{|x - x'|}\mathrm{d}x'$$

Finite Element Method

Let's show it on the Laplace equation. We want to solve:

$$\nabla^2 G(x, x') = \delta(x - x')$$

We will treat x' as a parameter, so we define $g_{x'}(x) \equiv G(x, x')$:

$$\nabla^2 g_{x'}(x) = \delta(x - x')$$

We set $g_{x'}(x) = 0$ on the boundary and we get:

$$-\int \nabla g_{x'}(x) \cdot \nabla v(x)\mathrm{d}x = \int v(x)\delta(x - x')\mathrm{d}x$$
$$-\int \nabla g_{x'}(x) \cdot \nabla v(x)\mathrm{d}x = v(x')$$

So we choose x' and then solve for $g_{x'}(x)$ using FEM and we get the Green function $G(x, x')$ for all x and one particular x'. We can then evaluate the integral (3.20.1.2) numerically – one would have to use FEM for all x' that are needed in the integral, so that is not efficient, but it should work. One will then be able to play with Green functions and be able to calculate them numerically for any boundary condition (which is not possible analytically).

3.21 Binomial Coefficients

For n and k integers, the binomial coefficients are defined by:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{k!}$$

For r real, one just uses the second formula as a definition:

$$\binom{r}{k} = \frac{r(r-1)\cdots(r-k+1)}{k!}$$

Example I:

$$\binom{-n}{k} = \frac{(-n)(-n-1)\cdots(-n-k+1)}{k!} = (-1)^k\frac{n(n+1)\cdots(n+k-1)}{k!} = (-1)^k\binom{n+k-1}{k}$$

Example II:

$$\binom{k-\frac{1}{2}}{k} = \frac{(k-\frac{1}{2})(k-\frac{1}{2}-1)\cdots(k-\frac{1}{2}-k+1)}{k!} = \frac{(k-\frac{1}{2})(k-\frac{1}{2}-1)\cdots\frac{1}{2}}{k!} =$$
$$= \frac{(2k-1)(2k-3)\cdots 1}{2^k k!} = \frac{(2k-1)!!}{2^k k!} = \frac{(2k)!}{(2^k k!)^2} = \frac{1}{4^k}\binom{2k}{k}$$

The binomial formula is for n integer:

$$(x+y)^n = \sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k}$$

and for r real and $|x| < |y|$ this can be generalized to:

$$(x+y)^r = \sum_{k=0}^{\infty}\binom{r}{k}x^k y^{r-k}$$

Example: (for $|x| < 1$)

$$(1+x)^{-n} = \sum_{k=0}^{\infty}\binom{-n}{k}x^k = \sum_{k=0}^{\infty}(-1)^k\binom{n+k-1}{k}x^k = \sum_{k=0}^{\infty}\binom{n+k-1}{k}(-x)^k$$

so:

$$(1-x)^{-n} = \sum_{k=0}^{\infty}\binom{n+k-1}{k}x^k$$
$$(1-x)^{-1} = \sum_{k=0}^{\infty}\binom{1+k-1}{k}x^k = \sum_{k=0}^{\infty}\binom{k}{k}x^k = \sum_{k=0}^{\infty}x^k$$
$$(1-x)^{-\frac{1}{2}} = \sum_{k=0}^{\infty}\binom{\frac{1}{2}+k-1}{k}x^k = \sum_{k=0}^{\infty}\binom{k-\frac{1}{2}}{k}x^k = \sum_{k=0}^{\infty}\frac{1}{4^k}\binom{2k}{k}x^k$$

Another example:

$$\begin{aligned}
\frac{1}{\sqrt{1-2xt+t^2}} = \left(1-(2xt-t^2)\right)^{-\frac{1}{2}} &= \sum_{n=0}^{\infty}\frac{1}{4^n}\binom{2n}{n}(2xt-t^2)^n = \\
&= \sum_{n=0}^{\infty}\frac{1}{4^n}\binom{2n}{n}\sum_{k=0}^{n}\binom{n}{k}(-t^2)^k(2xt)^{n-k} = \\
&= \sum_{n=0}^{\infty}\frac{1}{4^n}\binom{2n}{n}\sum_{k=0}^{n}\binom{n}{k}(-1)^k t^{2k}(2x)^{n-k}t^{n-k} = \\
&= \sum_{n=0}^{\infty}\frac{1}{4^n}\binom{2n}{n}\sum_{k=0}^{n}\binom{n}{k}(-1)^k t^{n+k}(2x)^{n-k} = \\
&= \sum_{n=0}^{\infty}\sum_{k=0}^{\lfloor\frac{n}{2}\rfloor}\frac{1}{4^{n-k}}\binom{2n-2k}{n-k}\binom{n-k}{k}(-1)^k t^n(2x)^{n-2k} = \\
&= \sum_{n=0}^{\infty}\left(\frac{1}{2^n}\sum_{k=0}^{\lfloor\frac{n}{2}\rfloor}(-1)^k\binom{n}{k}\binom{2n-2k}{n}x^{n-2k}\right)t^n = \\
&= \sum_{n=0}^{\infty}\left(\frac{1}{2^n}\sum_{k=0}^{\lfloor\frac{n}{2}\rfloor}(-1)^k\binom{n}{k}\frac{1}{n!}\frac{\mathrm{d}^n}{\mathrm{d}x^n}x^{2n-2k}\right)t^n = \\
&= \sum_{n=0}^{\infty}\left(\frac{1}{2^n}\sum_{k=0}^{n}(-1)^k\binom{n}{k}\frac{1}{n!}\frac{\mathrm{d}^n}{\mathrm{d}x^n}x^{2n-2k}\right)t^n = \\
&= \sum_{n=0}^{\infty}\left(\frac{1}{2^n n!}\frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^2-1)^n\right)t^n = \\
&= \sum_{n=0}^{\infty}P_n(x)t^n
\end{aligned}$$

where we used (3.22.2) and

$$\binom{2n-2k}{n-k}\binom{n-k}{k} = \binom{2n-2k}{n}\binom{n}{k}$$

The $P_n(x)$ are *Legendre Polynomials*.

3.22 Double Sums

When evaluating double sums, one can use triangular summation to reorder them:

$$\sum_{n=0}^{\infty}\sum_{k=0}^{\infty}a_{nk} = \sum_{n=0}^{\infty}\sum_{k=0}^{n}a_{n-k,k} \tag{3.22.1}$$

Also it holds

$$\sum_{n=0}^{\infty}\sum_{k=0}^{n}a_{nk} = \sum_{n=0}^{\infty}\sum_{k=0}^{\lfloor\frac{n}{2}\rfloor}a_{n-k,k} \tag{3.22.2}$$

3.23 Triangle Inequality

Triangle inequality (condition) means that none of the three quantities a, b, c is greater than the sum of the other two:

$$\begin{aligned} a+b &\geq c \\ b+c &\geq a \\ c+a &\geq b \end{aligned} \tag{3.23.1}$$

This is equivalent to just one equation:

$$|a-b| \leq c \leq a+b \tag{3.23.2}$$

we can do any permutation of the symbols, i.e. the above equation is equivalent to any of these:

$$\begin{aligned} |b-c| &\leq a \leq b+c \\ |c-a| &\leq b \leq c+a \end{aligned}$$

So instead of stating the three inequalities (3.23.1) it is more convenient to just write (3.23.2), using any permutation that we like.

To show, that (3.23.1) implies (3.23.2) we rewrite (3.23.1):

$$\begin{aligned} a+b &\geq c \\ c &\geq a-b \\ c &\geq b-a \end{aligned}$$

so

$$\begin{aligned} a+b &\geq c \\ c &\geq |a-b| \end{aligned}$$

and we get (3.23.2). To show, that (3.23.2) implies (3.23.1) we rewrite (3.23.2) for $a \geq b$ first:

$$\begin{aligned} a &\geq b \\ |a-b| \leq c &\leq a+b \end{aligned}$$

so:

$$\begin{aligned} a &\geq b \\ a-b \leq c &\leq a+b \end{aligned}$$

rearranging:

$$\begin{aligned} a+b &\geq c \\ b+c &\geq a \\ a &\geq b \end{aligned}$$

since c is positive, if $a \geq b$ then also $c+a \geq b$ and we get (3.23.1). Finally, for $a < b$:

$$\begin{aligned} a &< b \\ |a-b| \leq c &\leq a+b \end{aligned}$$

so:

$$\begin{aligned} a &< b \\ -(a-b) \leq c &\leq a+b \end{aligned}$$

rearranging:

$$\begin{aligned} a+b &\geq c \\ b &> a \\ c+a &\geq b \end{aligned}$$

since c is positive, if $b > a$ then also $b+c \geq a$ and we get (3.23.1).

3.24 Gamma Function

The Gamma function $\Gamma(x)$ is defined by the following properties for $x > 0$:

$$\Gamma(1) = 1 \tag{3.24.1}$$

$$\Gamma(x+1) = x\Gamma(x) \tag{3.24.2}$$

$$\log\Gamma(x) \quad \text{is convex} \tag{3.24.3}$$

It can be shown that this determines the function uniquely for $x > 0$ (this is called the Bohr-Mollerup theorem) and then it can be extended analytically to the whole complex plane.

The most common formula for $\Gamma(z)$ that satisfies (3.24.1), (3.24.2) and (3.24.3) is:

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}\mathrm{d}t \tag{3.24.4}$$

It satisfies (3.24.1) because:

$$\Gamma(1) = \int_0^\infty t^{1-1}e^{-t}\mathrm{d}t = \int_0^\infty e^{-t}\mathrm{d}t = [-e^{-t}]_0^\infty = 1$$

It satisfies (3.24.2) by integrating by parts:

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}\mathrm{d}t = (z-1)\int_0^\infty t^{z-2}e^{-t}\mathrm{d}t - [t^{z-1}e^{-t}]_0^\infty = (z-1)\Gamma(z-1)$$

Finally it satisfies (3.24.3) by verifying the convex condition directly ($x, y > 0$ and $0 \le \lambda \le 1$):

$$\begin{aligned}
\log\Gamma(\lambda x + (1-\lambda)y) &= \log\int_0^\infty t^{\lambda x+(1-\lambda)y-1}e^{-t}\mathrm{d}t = \\
&= \log\int_0^\infty (t^{x-1}e^{-t})^\lambda (t^{y-1}e^{-t})^{1-\lambda}\mathrm{d}t \le \\
&\le \log\left(\left(\int_0^\infty t^{x-1}e^{-t}\mathrm{d}t\right)^\lambda \left(\int_0^\infty t^{y-1}e^{-t}\mathrm{d}t\right)^{1-\lambda}\right) = \\
&= \lambda\log\Gamma(x) + (1-\lambda)\log\Gamma(y)
\end{aligned}$$

And thus (3.24.4) uniquely determines the Gamma function. We can use (3.24.4) to calculate $\Gamma(\frac{1}{2})$:

$$\begin{aligned}
\Gamma(\tfrac{1}{2}) &= \int_0^\infty t^{\frac{1}{2}-1}e^{-t}\mathrm{d}t = \int_0^\infty \frac{e^{-t}}{\sqrt{t}}\mathrm{d}t = \int_0^\infty \frac{e^{-x^2}}{x}2x\,\mathrm{d}x = 2\int_0^\infty e^{-x^2}\mathrm{d}x = \\
&= \int_{-\infty}^\infty e^{-x^2}\mathrm{d}x = \sqrt{\int_{-\infty}^\infty e^{-x^2}\mathrm{d}x\int_{-\infty}^\infty e^{-y^2}\mathrm{d}y} = \sqrt{2\pi\int_0^\infty e^{-r^2}r\mathrm{d}r} = \\
&= \sqrt{2\pi\int_0^\infty e^{-u}\tfrac{1}{2}\mathrm{d}u} = \sqrt{\pi}
\end{aligned}$$

From this and the definition of the Gamma function we get for integer n:

$$\begin{aligned}
\Gamma(n+1) = n\Gamma(n) = n(n-1)\Gamma(n-1) = n(n-1)(n-2)\cdots 2\cdot 1\cdot\Gamma(1) = \\
= n(n-1)(n-2)\cdots 1 = n!
\end{aligned} \tag{3.24.5}$$

and

$$\begin{aligned}
\Gamma(n+\tfrac{1}{2}) = (n-\tfrac{1}{2})\Gamma(n-\tfrac{1}{2}) = (n-\tfrac{1}{2})(n-1-\tfrac{1}{2})\Gamma(n-1-\tfrac{1}{2}) = (n-\tfrac{1}{2})(n-1-\tfrac{1}{2})\cdots\tfrac{1}{2}\Gamma(\tfrac{1}{2}) = \\
= \frac{2n-1}{2}\frac{2n-3}{2}\frac{2n-5}{2}\cdots\frac{1}{2}\Gamma(\tfrac{1}{2}) = \frac{(2n-1)!!}{2^n}\Gamma(\tfrac{1}{2}) = \frac{(2n-1)!!}{2^n}\sqrt{\pi}
\end{aligned} \tag{3.24.6}$$

3.25 Incomplete Gamma Function

The upper incomplete gamma function is defined by:

$$\Gamma(z,x) = \int_x^\infty t^{z-1}e^{-t}\mathrm{d}t$$

Integrating by parts we get:

$$\Gamma(z+1,x) = \int_x^\infty t^z e^{-t}\mathrm{d}t = z\int_x^\infty t^{z-1}e^{-t}\mathrm{d}t - [t^z e^{-t}]_x^\infty = z\Gamma(z,x) + x^z e^{-x}$$

Some special values are:

$$\Gamma(z,0) = \int_0^\infty t^{z-1}e^{-t}\mathrm{d}t = \Gamma(z)$$
$$\Gamma(1,x) = \int_x^\infty e^{-t}\mathrm{d}t = -[e^{-t}]_x^\infty = e^{-x}$$
$$\Gamma(\tfrac{1}{2},x) = \int_x^\infty t^{-\frac{1}{2}}e^{-t}\mathrm{d}t = 2\int_{\sqrt{x}}^\infty e^{-s^2}\mathrm{d}s = \sqrt{\pi}\mathrm{erfc}(\sqrt{x})$$

For integer n we get:

$$\begin{aligned}\Gamma(n+1,x) &= n\Gamma(n,x) + x^n e^{-x} = n(n-1)\Gamma(n-1,x) + (nx^{n-1}+x^n)e^{-x} = \\ &= n(n-1)(n-2)\Gamma(n-2,x) + (n(n-1)x^{n-2} + nx^{n-1} + x^n)e^{-x} = \\ &= n(n-1)(n-2)\cdots 2\cdot 1\cdot\Gamma(1,x) + (n(n-1)\cdots 2x^1 + \cdots + n(n-1)x^{n-2} + nx^{n-1} + x^n)e^{-x} = \\ &= n!e^{-x} + (n(n-1)\cdots 2x^1 + \cdots + n(n-1)x^{n-2} + nx^{n-1} + x^n)e^{-x} = \\ &= n!e^{-x}\sum_{\nu=0}^{n}\frac{x^\nu}{\nu!}\end{aligned}$$

and

$$\begin{aligned}\Gamma(n+\tfrac{1}{2},x) &= (n-\tfrac{1}{2})\Gamma(n-\tfrac{1}{2},x) + x^{n-\frac{1}{2}}e^{-x} = (n-\tfrac{1}{2})(n-1-\tfrac{1}{2})\Gamma(n-1-\tfrac{1}{2},x) + ((n-\tfrac{1}{2})x^{n-1-\frac{1}{2}} + x^{n-\frac{1}{2}}))e^{-x} = \\ &= (n-\tfrac{1}{2})(n-1-\tfrac{1}{2})\cdots\tfrac{1}{2}\Gamma(\tfrac{1}{2},x) + ((n-\tfrac{1}{2})(n-1-\tfrac{1}{2})\cdots(1+\tfrac{1}{2})x^{\frac{1}{2}} + \cdots + (n-\tfrac{1}{2})x^{n-1-\frac{1}{2}} + x^{n-\frac{1}{2}}))e^{-x} = \\ &= \frac{(2n-1)!!}{2^n}\Gamma(\tfrac{1}{2},x) + \frac{(2n-1)!!}{2^n}e^{-x}\sum_{\nu=1}^{n}\frac{2^\nu x^{\nu-\frac{1}{2}}}{(2\nu-1)!!} = \\ &= \frac{(2n-1)!!}{2^n}\left(\sqrt{\pi}\mathrm{erfc}(\sqrt{x}) + e^{-x}\sum_{\nu=1}^{n}\frac{2^\nu x^{\nu-\frac{1}{2}}}{(2\nu-1)!!}\right)\end{aligned}$$

The lower incomplete gamma function is defined by:

$$\gamma(z,x) = \int_0^x t^{z-1}e^{-t}\mathrm{d}t = \Gamma(z) - \Gamma(z,x)$$

and as such all expressions can be easily derived using the gamma and upper incomplete gamma functions. The recursion relation is then:

$$\gamma(z+1,x) = \Gamma(z+1) - \Gamma(z+1,x) = z\Gamma(z) - (z\Gamma(z,x) + x^z e^{-x}) = z\gamma(z,x) - x^z e^{-x}$$

Some special values are:

$$\gamma(z,0) = \Gamma(z) - \Gamma(z,0) = \Gamma(z) - \Gamma(z) = 0$$
$$\gamma(1,x) = \Gamma(1) - \Gamma(1,x) = 1 - e^{-x}$$
$$\gamma(\tfrac{1}{2},x) = \Gamma(\tfrac{1}{2}) - \Gamma(\tfrac{1}{2},x) = \sqrt{\pi} - \sqrt{\pi}\mathrm{erfc}(\sqrt{x}) = \sqrt{\pi}\mathrm{erf}(\sqrt{x})$$

By repeated application of the recursion formula we get:

$$\gamma(z,x) = \frac{1}{z}\gamma(z+1,x) + e^{-x}\frac{x^z}{z} = \frac{1}{z(z+1)}\gamma(z+2,x) + e^{-x}\left(\frac{x^z}{z} + \frac{x^{z+1}}{z(z+1)}\right) =$$
$$= \frac{1}{z(z+1)\cdots(z+n)}\gamma(z+n+1,x) + e^{-x}\sum_{k=0}^{n}\frac{x^{z+k}}{z(z+1)\cdots(z+k)} =$$
$$= \frac{\Gamma(z)}{\Gamma(z+n+1)}\gamma(z+n+1,x) + x^z\Gamma(z)e^{-x}\sum_{k=0}^{n}\frac{x^k}{\Gamma(z+k+1)} =$$
$$= x^z\Gamma(z)e^{-x}\sum_{k=0}^{\infty}\frac{x^k}{\Gamma(z+k+1)} \tag{3.25.1}$$

where we used:

$$\lim_{z\to\infty}\frac{\gamma(z,x)}{\Gamma(z)} = 0$$

which can be proven by the following inequality which uses the fact that the function $f(t) = t^{z-1}e^{-t}$ is an increasing function for $t < z-1$, so as long as $x < z-1$ we get:

$$\gamma(z,x) = \int_0^x t^{z-1}e^{-t}\mathrm{d}t = \int_0^x f(t)\mathrm{d}t <$$
$$< \int_0^x f(x)\mathrm{d}t = xf(x) =$$
$$= \frac{x}{z-1-x}\int_x^{z-1} f(x)\mathrm{d}t <$$
$$< \frac{x}{z-1-x}\int_x^{z-1} f(t)\mathrm{d}t <$$
$$< \frac{x}{z-1-x}\int_0^{\infty} f(t)\mathrm{d}t =$$
$$= \frac{x}{z-1-x}\Gamma(z)$$

Using (3.25.1) we can now write $\gamma(z,x)$ using the Kummer confluent hypergeometric function ${}_1F_1(a,b,z)$ as follows:

$$\gamma(z,x) = x^z\Gamma(z)e^{-x}\sum_{k=0}^{\infty}\frac{x^k}{\Gamma(z+k+1)} = x^z z^{-1}e^{-x}\,{}_1F_1(1,z+1,x) = x^z z^{-1}\,{}_1F_1(z,z+1,-x)$$

3.25.1 Example

Consider the class of integrals:

$$F_m(t) = \int_0^1 u^{2m}e^{-tu^2}\mathrm{d}u, \qquad (t>0; m=0,1,2,\dots)$$

We write them using the lower incomplete gamma function as:

$$F_m(t) = \int_0^t \left(\frac{v}{t}\right)^m e^{-v}\left(\frac{v}{t}\right)^{-\frac{1}{2}}\frac{\mathrm{d}v}{2t} = \frac{1}{2t^{m+\frac{1}{2}}}\int_0^t v^{m-\frac{1}{2}}e^{-v}\mathrm{d}v = \frac{\gamma(m+\frac{1}{2},t)}{2t^{m+\frac{1}{2}}}$$

We can also write it using the confluent hypergeometric function as follows:

$$F_m(t) = \frac{\gamma(m+\frac{1}{2},t)}{2t^{m+\frac{1}{2}}} = \frac{t^{m+\frac{1}{2}}(m+\frac{1}{2})^{-1}}{2t^{m+\frac{1}{2}}}\,{}_1F_1(m+\tfrac{1}{2},m+\tfrac{3}{2},-t) = \frac{{}_1F_1(m+\frac{1}{2},m+\frac{3}{2},-t)}{2m+1}$$

For $m = 0$ we get:

$$F_0(t) = \frac{\gamma(\frac{1}{2}, t)}{2t^{\frac{1}{2}}} = \frac{1}{2}\sqrt{\frac{\pi}{t}}\text{erf}(\sqrt{t})$$

Using the recursion relation we get:

$$F_{m+1}(t) = \frac{\gamma(m+\frac{1}{2}+1, t)}{2t^{m+\frac{1}{2}+1}} = \frac{(m+\frac{1}{2})\gamma(m+\frac{1}{2}, t) - t^{m+\frac{1}{2}}e^{-t}}{2t^{m+\frac{1}{2}}t} = \frac{(m+\frac{1}{2})}{t}F_m(t) - \frac{e^{-t}}{2t} = \\ = \frac{(2m+1)F_m(t) - e^{-t}}{2t}$$

By expressing $F_m(t)$ from the equation we obtain the inverse relation:

$$F_m(t) = \frac{2tF_{m+1}(t) + e^{-t}}{2m+1}$$

From (3.25.1) we get:

$$F_m(t) = \frac{\gamma(m+\frac{1}{2}, t)}{2t^{m+\frac{1}{2}}} = \frac{1}{2t^{m+\frac{1}{2}}}t^{m+\frac{1}{2}}\Gamma(m+\frac{1}{2})e^{-t}\sum_{k=0}^{\infty}\frac{t^k}{\Gamma(m+\frac{1}{2}+k+1)} = \\ = \frac{1}{2}\Gamma(m+\frac{1}{2})e^{-t}\sum_{k=0}^{\infty}\frac{t^k}{\Gamma(m+k+\frac{3}{2})} = \\ = \frac{1}{2}\frac{(2m-1)!!\sqrt{\pi}}{2^m}e^{-t}\sum_{k=0}^{\infty}\frac{t^k}{\frac{(2m+2k+1)!!\sqrt{\pi}}{2^{m+k+1}}} = \\ = e^{-t}\sum_{k=0}^{\infty}\frac{(2m-1)!!(2t)^k}{(2m+2k+1)!!} = \\ = e^{-t}\sum_{k=0}^{\infty}\frac{(2t)^k}{(2m+1)(2m+3)\cdots(2m+2k+1)}$$

3.26 Factorial

The factorial $n!$ is defined as

$$n! = n(n-1)(n-2)\cdots 3\cdot 2\cdot 1$$

By (3.24.5) it can be written using the Gamma function as:

$$n! = \Gamma(n+1)$$

3.27 Double Factorial

The double factorial $n!!$ is defined as:

$$n!! = \begin{cases} n(n-2)(n-4)(n-6)\cdots 5\cdot 3\cdot 1 & \text{for odd } n = 2k+1 \\ n(n-2)(n-4)(n-6)\cdots 6\cdot 4\cdot 2 & \text{for even } n = 2k \end{cases}$$

One can rewrite double factorial using a factorial as:

$$(2k)!! = 2 \cdot 4 \cdot 6 \cdots (2k) = 2^k (1 \cdot 2 \cdot 3 \cdots k) = 2^k k!$$
$$(2k-1)!! = 1 \cdot 3 \cdot 5 \cdots (2k-1) = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdots (2k)}{2 \cdot 4 \cdot 6 \cdots (2k)} = \frac{(2k)!}{(2k)!!} = \frac{(2k)!}{2^k k!}$$

For odd n it can be written using the Gamma function, see (3.24.6):

$$(2k-1)!! = \frac{1}{\sqrt{\pi}} 2^k \Gamma\left(k + \tfrac{1}{2}\right)$$

3.27.1 Example

$$A(n) = \frac{1 \cdot 3 \cdot 5 \cdot \cdots \cdot (2n-1)}{1 \cdot 2 \cdot 3 \cdot \cdots \cdot n} = \frac{(2n-1)!!}{n!} = \frac{(2n)!}{2^n (n!)^2} = \frac{1}{2^n} \binom{2n}{n}$$
$$B(n) = \frac{1 \cdot 3 \cdot 5 \cdot \cdots \cdot (2n-1)}{2 \cdot 4 \cdot \cdot 6 \cdot \cdots \cdot 2n} = \frac{(2n-1)!!}{(2n)!!} = \frac{(2n)!}{(2^n n!)^2} = \frac{1}{4^n} \binom{2n}{n}$$

3.28 Fermi-Dirac Integral

The Fermi-Dirac integral (sometimes just called a Fermi integral) is defined as:

$$I_\alpha(\mu) = \int_0^\infty \frac{\epsilon^\alpha \mathrm{d}\epsilon}{e^{\epsilon - \mu} + 1}$$

Examples:

$$I_{\frac{1}{2}}(\mu) = \int_0^\infty \frac{\sqrt{\epsilon}\mathrm{d}\epsilon}{e^{\epsilon - \mu} + 1}$$
$$I_{\frac{3}{2}}(\mu) = \int_0^\infty \frac{\epsilon^{\frac{3}{2}}\mathrm{d}\epsilon}{e^{\epsilon - \mu} + 1}$$

The Fermi-Dirac integral can also be written using the polylogarithm, see *The Series pFq* for details.

3.29 Legendre Polynomials

Legendre polynomials $P_l(x)$ defined by the Rodrigues' formula

$$P_l(x) = \frac{1}{2^l l!} \frac{\mathrm{d}^l}{\mathrm{d}x^l} [(x^2 - 1)^l]$$

they also obey the completeness relation

$$\sum_{l=0}^{\infty} \frac{2l+1}{2} P_l(x') P_l(x) = \delta(x - x') \tag{3.29.1}$$

and orthogonality relation:

$$\int_{-1}^{1} P_k(x) P_l(x) \mathrm{d}x = \frac{2\delta_{kl}}{2k+1}$$

Two Legendre polynomials can be expanded in a series:

$$P_k(x)P_l(x) = \sum_{m=|k-l|}^{k+l} \begin{pmatrix} k & l & m \\ 0 & 0 & 0 \end{pmatrix}^2 (2m+1)P_m(x)$$

This was first proven by *[Adams]*, where he shows:

$$P_k(x)P_l(x) = \sum_{m=|k-l|}^{k+l} \frac{A(s-k)A(s-l)A(s-m)}{A(s)} \frac{2m+1}{2s+1} P_m(x)$$

where $s = \frac{k+l+m}{2}$ and

$$A(n) = \frac{1 \cdot 3 \cdot 5 \cdot \cdots \cdot (2n-1)}{1 \cdot 2 \cdot 3 \cdot \cdots \cdot n} = \frac{(2n)!}{2^n (n!)^2} = \frac{1}{2^n} \binom{2n}{n}$$

The coefficient in the expansion can then be written using a $3j$ symbol as:

$$\begin{aligned}
\frac{A(s-k)A(s-l)A(s-m)}{A(s)} \frac{1}{2s+1} &= \\
= \frac{\frac{1}{2^{s-k}}\binom{2s-2k}{s-k}\frac{1}{2^{s-l}}\binom{2s-2l}{s-l}\frac{1}{2^{s-m}}\binom{2s-2m}{s-m}}{\frac{1}{2^s}\binom{2s}{s}} \frac{1}{2s+1} &= \\
= \frac{2^s}{2^{s-k+s-l+s-m}} \frac{\binom{2s-2k}{s-k}\binom{2s-2l}{s-l}\binom{2s-2m}{s-m}}{\binom{2s}{s}} \frac{1}{2s+1} &= \\
= \frac{\binom{2s-2k}{s-k}\binom{2s-2l}{s-l}\binom{2s-2m}{s-m}}{\binom{2s}{s}} \frac{1}{2s+1} &= \\
= \frac{(2s-2k)!}{((s-k)!)^2} \frac{(2s-2l)!}{((s-l)!)^2} \frac{(2s-2m)!}{((s-m)!)^2} \frac{(s!)^2}{(2s)!} \frac{1}{2s+1} &= \\
= \frac{(2s-2k)!(2s-2l)!(2s-2m)!}{(2s+1)!} \left[\frac{s!}{(s-k)!(s-l)!(s-m)!}\right]^2 &= \\
= \begin{pmatrix} k & l & m \\ 0 & 0 & 0 \end{pmatrix}^2 &
\end{aligned}$$

So we will be just using the $3j$ symbol form from now on. We can now calculate the integral of three Legendre polynomials:

$$\begin{aligned}
\int_{-1}^{1} P_k(x)P_l(x)P_m(x)\mathrm{d}x &= \\
= \int_{-1}^{1} \sum_{n=|k-l|}^{k+l} \begin{pmatrix} k & l & n \\ 0 & 0 & 0 \end{pmatrix}^2 (2n+1)P_n(x)P_m(x)\mathrm{d}x &= \\
= \sum_{n=|k-l|}^{k+l} \begin{pmatrix} k & l & n \\ 0 & 0 & 0 \end{pmatrix}^2 (2n+1)\frac{2\delta_{nm}}{2n+1} &= \\
= 2\begin{pmatrix} k & l & m \\ 0 & 0 & 0 \end{pmatrix}^2 &
\end{aligned} \tag{3.29.2}$$

This is consistent with the series expansion:

$$\begin{aligned}
P_k(x)P_l(x) = \sum_{m=|k-l|}^{k+l} \frac{2m+1}{2} \int_{-1}^{1} P_k(x)P_l(x)P_m(x)\mathrm{d}x \; P_m(x) &= \\
= \sum_{m=|k-l|}^{k+l} \begin{pmatrix} k & l & m \\ 0 & 0 & 0 \end{pmatrix}^2 (2m+1)P_m(x) &
\end{aligned}$$

Any function $f(x)$ (where $-1 \le x \le 1$) can be expanded as:

$$f(x) = \sum_{l=0}^{\infty} f_l P_l(x)$$

$$f_l = \frac{(2l+1)}{2} \int_{-1}^{1} f(x) P_l(x) \mathrm{d}x$$

For the following choice of $f(x)$ we get (for $|t| \le 1$):

$$f(x) = \frac{1}{\sqrt{1-2xt+t^2}}$$

$$f_l = \frac{(2l+1)}{2} \int_{-1}^{1} \frac{P_l(x)}{\sqrt{1-2xt+t^2}} \mathrm{d}x = \frac{(2l+1)}{2} \int_{|1+t|}^{|1-t|} \frac{P_l\left(\frac{1-R^2+t^2}{2t}\right)}{R} \left(-\frac{R}{t}\right) \mathrm{d}R =$$

$$= \frac{(2l+1)}{2t} \int_{|1-t|}^{|1+t|} P_l\left(\frac{1-R^2+t^2}{2t}\right) \mathrm{d}R = \frac{(2l+1)}{2t} \int_{1-t}^{1+t} P_l\left(\frac{1-R^2+t^2}{2t}\right) \mathrm{d}R =$$

$$= t^l$$

Code:

```
>>> from sympy import var, legendre, integrate
>>> var("l R t")
(l, R, t)
>>> f = (2*l+1) / (2*t) * integrate(legendre(l, (1-R**2+t**2) / (2*t)),
...           (R, 1-t, 1+t))
>>> for _l in range(20): print _l, f.subs(l, _l).doit().simplify()
...
0 1
1 t
2 t**2
3 t**3
4 t**4
5 t**5
6 t**6
7 t**7
8 t**8
9 t**9
10 t**10
11 t**11
12 t**12
13 t**13
14 t**14
15 t**15
16 t**16
17 t**17
18 t**18
19 t**19
```

So the Legendre polynomials are the coefficients of the following expansion for $|t| \le 1$:

$$\frac{1}{\sqrt{1-2xt+t^2}} = \sum_{l=0}^{\infty} P_l(x) t^l$$

Note that for $|t| > 1$ we get:

$$\frac{1}{\sqrt{1-2xt+t^2}} = \frac{1}{|t|} \frac{1}{\sqrt{1-2x\frac{1}{t}+\left(\frac{1}{t}\right)^2}} = \frac{1}{|t|} \sum_{l=0}^{\infty} P_l(x) \left(\frac{1}{t}\right)^l = \operatorname{sign} t \sum_{l=0}^{\infty} P_l(x) t^{-l-1}$$

3.29.1 Example I

Very important is the following multipole expansion:

$$\begin{aligned}\frac{1}{|\mathbf{r}-\mathbf{r}'|}=\frac{1}{\sqrt{(\mathbf{r}-\mathbf{r}')^2}}=\frac{1}{\sqrt{r^2-2\mathbf{r}\cdot\mathbf{r}'+r'^2}}=\frac{1}{r_>\sqrt{1-2\left(\frac{r_<}{r_>}\right)\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}'+\left(\frac{r_<}{r_>}\right)^2}}=\\ =\frac{1}{r_>}\sum_{l=0}^{\infty}\left(\frac{r_<}{r_>}\right)^l P_l(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}')=\sum_{l=0}^{\infty}\frac{r_<^l}{r_>^{l+1}}P_l(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}')\end{aligned} \tag{3.29.1.1}$$

Where $r_>=\max(r,r')$ and $r_<=\min(r,r')$. Assuming $r>r'$, we get for the first few terms:

$$\begin{aligned}\frac{1}{|\mathbf{r}-\mathbf{r}'|}=\frac{1}{r}\left(P_0(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}')+P_1(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}')\frac{r'}{r}+P_2(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}')\left(\frac{r'}{r}\right)^2+O\left(\frac{r'^3}{r^3}\right)\right)=\\ =\frac{1}{r}\left(1+\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}'\frac{r'}{r}+\tfrac{1}{2}\left(3(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}')^2-1\right)\left(\frac{r'}{r}\right)^2+O\left(\frac{r'^3}{r^3}\right)\right)=\\ =\frac{1}{r}+\frac{\mathbf{r}\cdot\mathbf{r}'}{r^3}+\frac{3(\mathbf{r}\cdot\mathbf{r}')^2-r^2r'^2}{2r^5}+O\left(\frac{r'^3}{r^4}\right)\end{aligned}$$

3.29.2 Example II

Let's find the expansion of

$$f(x)=\frac{e^{-\alpha\sqrt{1-2xt+t^2}}}{\sqrt{1-2xt+t^2}}$$

for $|t|\leq 1$. We get:

$$\begin{aligned}f_l=\frac{(2l+1)}{2}\int_{-1}^{1}\frac{P_l(x)e^{-\alpha\sqrt{1-2xt+t^2}}}{\sqrt{1-2xt+t^2}}\mathrm{d}x=\frac{(2l+1)}{2}\int_{|1+t|}^{|1-t|}\frac{P_l\left(\frac{1-R^2+t^2}{2t}\right)e^{-\alpha R}}{R}\left(-\frac{R}{t}\right)\mathrm{d}R=\\ =\frac{(2l+1)}{2t}\int_{|1-t|}^{|1+t|}P_l\left(\frac{1-R^2+t^2}{2t}\right)e^{-\alpha R}\mathrm{d}R=\frac{(2l+1)}{2t}\int_{1-t}^{1+t}P_l\left(\frac{1-R^2+t^2}{2t}\right)e^{-\alpha R}\mathrm{d}R\end{aligned}$$

Here is the result for the first few l:

$$f_0=\frac{\left(e^{2\alpha t}-1\right)e^{-\alpha t-\alpha}}{2\alpha t}$$

$$f_1=\frac{3}{2}\frac{\left(\alpha^2te^{2\alpha t}+\alpha^2t+\alpha te^{2\alpha t}+\alpha t-\alpha e^{2\alpha t}+\alpha-e^{2\alpha t}+1\right)e^{-\alpha}}{\alpha^3t^2}$$

$$f_2=\frac{5}{2}\frac{\left(\alpha^4t^2e^{2\alpha t}-\alpha^4t^2+3\alpha^3t^2e^{2\alpha t}-3\alpha^3t^2-3\alpha^3te^{2\alpha t}-3\alpha^3t-\right.}{\alpha^5t^3}$$

$$X=3\alpha^2e^{2\alpha t}-3\alpha^2-9\alpha te^{2\alpha t}-9\alpha t+9\alpha e^{2\alpha t}-9\alpha+9e^{2\alpha t}-9$$

Expanding in t up to $\mathcal{O}\left(t^7\right)$ we get:

$$f_l = e^{-\alpha} g_l$$

$$g_0 = 1 + \frac{1}{6}\alpha^2 t^2 + \frac{1}{120}\alpha^4 t^4 + \frac{1}{5040}\alpha^6 t^6 + \mathcal{O}\left(t^7\right)$$

$$g_1 = t + \alpha t + \frac{1}{10}\alpha^2 t^3 + \frac{1}{10}\alpha^3 t^3 + \frac{1}{280}\alpha^4 t^5 + \frac{1}{280}\alpha^5 t^5 + \mathcal{O}\left(t^7\right)$$

$$g_2 = t^2 + \alpha t^2 + \frac{1}{3}\alpha^2 t^2 + \frac{1}{14}\alpha^2 t^4 + \frac{1}{14}\alpha^3 t^4 + \frac{1}{42}\alpha^4 t^4 + \frac{1}{504}\alpha^4 t^6 + \frac{1}{504}\alpha^5 t^6 + \frac{1}{1512}\alpha^6 t^6 + \mathcal{O}\left(t^7\right)$$

$$g_3 = t^3 + \alpha t^3 + \frac{2}{5}\alpha^2 t^3 + \frac{1}{18}\alpha^2 t^5 + \frac{1}{15}\alpha^3 t^3 + \frac{1}{18}\alpha^3 t^5 + \frac{1}{45}\alpha^4 t^5 + \frac{1}{270}\alpha^5 t^5 + \mathcal{O}\left(t^7\right)$$

$$g_4 = t^4 + \alpha t^4 + \frac{3}{7}\alpha^2 t^4 + \frac{1}{22}\alpha^2 t^6 + \frac{2}{21}\alpha^3 t^4 + \frac{1}{22}\alpha^3 t^6 + \frac{1}{105}\alpha^4 t^4 + \frac{3}{154}\alpha^4 t^6 + \frac{1}{231}\alpha^5 t^6 + \frac{1}{2310}\alpha^6 t^6 + \mathcal{O}\left(t^7\right)$$

Code:

```
>>> from sympy import var, legendre, integrate, exp, latex, cse
>>> var("l R t alpha")
(l, R, t, alpha)
>>>
>>> f = (2*l+1) / (2*t) * integrate(legendre(l, (1-R**2+t**2) / (2*t)) \
...         * exp(-alpha*R),
...         (R, 1-t, 1+t))
>>>
>>> for _l in range(3):
...     print "f_%d & =" %_l, latex(f.subs(l, _l).doit().simplify()), "\\\\"
...
f_0 & = \frac{\left(e^{2 \alpha t} -1\right) e^{- \alpha t - \alpha}}{2 \alpha t} \\
f_1 & = \frac{3}{2} \frac{\left(\alpha^{2} t e^{2 \alpha t} + \alpha^{2} t + \alpha t␣
↪e^{2 \alpha t} + \alpha t - \alpha e^{2 \alpha t} + \alpha - e^{2 \alpha t} +␣
↪1\right) e^{- \alpha t - \alpha}}{\alpha^{3} t^{2}} \\
f_2 & = \frac{5}{2} \frac{\left(\alpha^{4} t^{2} e^{2 \alpha t} - \alpha^{4} t^{2} +␣
↪3 \alpha^{3} t^{2} e^{2 \alpha t} - 3 \alpha^{3} t^{2} - 3 \alpha^{3} t e^{2 \alpha␣
↪t} - 3 \alpha^{3} t + 3 \alpha^{2} t^{2} e^{2 \alpha t} - 3 \alpha^{2} t^{2} - 9␣
↪\alpha^{2} t e^{2 \alpha t} - 9 \alpha^{2} t + 3 \alpha^{2} e^{2 \alpha t} - 3␣
↪\alpha^{2} - 9 \alpha t e^{2 \alpha t} - 9 \alpha t + 9 \alpha e^{2 \alpha t} - 9␣
↪\alpha + 9 e^{2 \alpha t} -9\right) e^{- \alpha t - \alpha}}{\alpha^{5} t^{3}} \\
>>> for _l in range(5):
...     result = f.subs(l, _l).doit().simplify() / exp(-alpha)
...     print "g_%d & =" %_l, latex(result.series(t, 0, 7)), "\\\\"
...
g_0 & = 1 + \frac{1}{6} \alpha^{2} t^{2} + \frac{1}{120} \alpha^{4} t^{4} + \frac{1}
↪{5040} \alpha^{6} t^{6} + \operatorname{\mathcal{O}}\left(t^{7}\right) \\
g_1 & = t + \alpha t + \frac{1}{10} \alpha^{2} t^{3} + \frac{1}{10} \alpha^{3} t^{3}␣
↪+ \frac{1}{280} \alpha^{4} t^{5} + \frac{1}{280} \alpha^{5} t^{5} + \operatorname
↪{\mathcal{O}}\left(t^{7}\right) \\
g_2 & = t^{2} + \alpha t^{2} + \frac{1}{3} \alpha^{2} t^{2} + \frac{1}{14} \alpha^{2}␣
↪t^{4} + \frac{1}{14} \alpha^{3} t^{4} + \frac{1}{42} \alpha^{4} t^{4} + \frac{1}
↪{504} \alpha^{4} t^{6} + \frac{1}{504} \alpha^{5} t^{6} + \frac{1}{1512} \alpha^{6}␣
↪t^{6} + \operatorname{\mathcal{O}}\left(t^{7}\right) \\
g_3 & = t^{3} + \alpha t^{3} + \frac{2}{5} \alpha^{2} t^{3} + \frac{1}{18} \alpha^{2}␣
↪t^{5} + \frac{1}{15} \alpha^{3} t^{3} + \frac{1}{18} \alpha^{3} t^{5} + \frac{1}{45}
↪ \alpha^{4} t^{5} + \frac{1}{270} \alpha^{5} t^{5} + \operatorname{\mathcal{O}}
↪\left(t^{7}\right) \\
g_4 & = t^{4} + \alpha t^{4} + \frac{3}{7} \alpha^{2} t^{4} + \frac{1}{22} \alpha^{2}␣
↪t^{6} + \frac{2}{21} \alpha^{3} t^{4} + \frac{1}{22} \alpha^{3} t^{6} + \frac{1}
↪{105} \alpha^{4} t^{4} + \frac{3}{154} \alpha^{4} t^{6} + \frac{1}{231} \alpha^{5}␣
↪t^{6} + \frac{1}{2310} \alpha^{6} t^{6} + \operatorname{\mathcal{O}}\left(t^{7}
↪\right) \\
```

(continues on next page)

(continued from previous page)

3.29.3 Example III

$$\frac{e^{-\frac{|\mathbf{r}-\mathbf{r}'|}{D}}}{|\mathbf{r}-\mathbf{r}'|} = \frac{e^{\frac{-r_>\sqrt{1-2\left(\frac{r_<}{r_>}\right)\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}'+\left(\frac{r_<}{r_>}\right)^2}}{D}}}{r_>\sqrt{1-2\left(\frac{r_<}{r_>}\right)\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}'+\left(\frac{r_<}{r_>}\right)^2}} = \frac{1}{r_>}\frac{e^{-\alpha\sqrt{1-2xt+t^2}}}{\sqrt{1-2xt+t^2}}$$

where:

$$\begin{aligned}\alpha &= \frac{r_>}{D}\\ x &= \hat{\mathbf{r}}\cdot\hat{\mathbf{r}}'\\ t &= \frac{r_<}{r_>}\end{aligned}$$

3.29.4 Example IV

$$V(|\mathbf{r}_1-\mathbf{r}_2|) = \frac{e^{-\frac{|\mathbf{r}_1-\mathbf{r}_2|}{D}}}{|\mathbf{r}_1-\mathbf{r}_2|}$$

The potential V is a function of r_1, r_2 and $\cos\theta$ only:

$$V(|\mathbf{r}_1-\mathbf{r}_2|) = V\left(\sqrt{r_1^2-2\mathbf{r_1}\cdot\mathbf{r_2}+r_2^2}\right) = V\left(\sqrt{r_1^2-2r_1r_2\cos\theta+r_2^2}\right) = V(r_1,r_2,\cos\theta)$$

So we expand in the $\cos\theta$ variable using the Legendre expansion:

$$V(|\mathbf{r}_1-\mathbf{r}_2|) = V(r_1,r_2,\cos\theta) = \sum_{l=0}^{\infty} V_l(r_1,r_2)P_l(\cos\theta)$$

where $V_l(r_1,r_2)$ only depends on r_1 and r_2:

$$\begin{aligned}V_l(r_1,r_2) &= \frac{2l+1}{2}\int_{-1}^{1} V(|\mathbf{r}_1-\mathbf{r}_2|)P_l(\cos\theta)\mathrm{d}(\cos\theta) =\\ &= \frac{2l+1}{2}\int_{-1}^{1}\frac{e^{-\frac{|\mathbf{r}_1-\mathbf{r}_2|}{D}}}{|\mathbf{r}_1-\mathbf{r}_2|}P_l(\cos\theta)\mathrm{d}(\cos\theta) =\\ &= \frac{2l+1}{2r_1r_2}\int_{|r_1-r_2|}^{r_1+r_2} e^{-\frac{r}{D}}P_l\left(\frac{r_1^2-r^2+r_2^2}{2r_1r_2}\right)\mathrm{d}r\end{aligned}$$

In the limit $D\to\infty$ we get:

$$V_l(r_1,r_2) \to \frac{r_<^l}{r_>^{l+1}}$$

In general, the $V_l(r_1,r_2)$ expressions are complicated. For the first few l we get:

$$V_0(r_1,r_2) = \frac{D}{2r_1r_2}\left(e^{-\frac{|r_1-r_2|}{D}} - e^{-\frac{r_1+r_2}{D}}\right)$$

$$V_1(r_1,r_2) = \frac{3}{2}\frac{D\left(-D^2e^{2\frac{r_2}{D}}+D^2-Dr_1e^{2\frac{r_2}{D}}+Dr_1+Dr_2e^{2\frac{r_2}{D}}+Dr_2+r_1r_2e^{2\frac{r_2}{D}}+r_1r_2\right)e^{-\frac{r_1}{D}-\frac{r_2}{D}}}{r_1^2r_2^2}$$

In $V_1(r_1,r_2)$ we assume $r_1\geq r_2$.

3.30 Spherical Harmonics

Are defined for $m \geq 0$ by

$$Y_{lm}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta)\, e^{im\phi}$$

where P_l^m are associated Legendre polynomials defined by

$$P_l^m(x) = (-1)^m (1-x^2)^{m/2} \frac{\mathrm{d}^m}{\mathrm{d}x^m} P_l(x)$$

and P_l are Legendre polynomials. For $m < 0$ they are defined by:

$$Y_{lm}(\Omega) = (-1)^m Y^*_{l,-m}(\Omega)$$

Sometimes the spherical harmonics are written as:

$$Y_{lm}(\theta, \phi) = \Theta_{lm}(\theta)\Phi_m(\phi)$$

where:

$$\Phi_m(\phi) = \frac{1}{\sqrt{2\pi}} e^{im\phi}$$
$$\Theta_{lm}(\theta) = \begin{cases} \sqrt{\frac{2l+1}{2} \frac{(l-m)!}{(l+m)!}}\, P_l^m(\cos\theta) & \text{for } m \geq 0 \\ (-1)^m \Theta_{l,-m}(\theta) & \text{for } m < 0 \end{cases}$$

The spherical harmonics are ortonormal:

$$\int Y_{lm}\, Y^*_{l'm'}\, \mathrm{d}\Omega = \int_0^{2\pi} \int_0^{\pi} Y_{lm}(\theta, \phi)\, Y^*_{l'm'}(\theta, \phi) \sin\theta\, \mathrm{d}\theta\, \mathrm{d}\phi = \delta_{mm'}\delta_{ll'} \tag{3.30.1}$$

and complete (both in the l-subspace and the whole space):

$$\sum_{m=-l}^{l} |Y_{lm}(\theta, \phi)|^2 = \frac{2l+1}{4\pi} \tag{3.30.2}$$

$$\sum_{l=0}^{\infty} \sum_{m=-l}^{l} Y_{lm}(\theta, \phi) Y^*_{lm}(\theta', \phi') = \frac{1}{\sin\theta} \delta(\theta - \theta')\delta(\phi - \phi') = \delta(\hat{\mathbf{r}} - \hat{\mathbf{r}}') \tag{3.30.3}$$

The relation (3.30.2) is a special case of an addition theorem for spherical harmonics

$$\sum_{m=-l}^{l} Y_{lm}(\theta, \phi) Y^*_{lm}(\theta', \phi') = \frac{2l+1}{4\pi} P_l(\cos\gamma) \tag{3.30.4}$$

where γ is the angle between the unit vectors given by $\hat{\mathbf{r}} = (\theta, \phi)$ and $\hat{\mathbf{r}}' = (\theta', \phi')$:

$$\cos\gamma = \cos\theta\cos\theta' + \sin\theta\sin\theta'\cos(\phi - \phi') = \hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'$$

Relations between complex conjugates is:

$$Y^*_{lm}(\Omega) = (-1)^m Y_{l,-m}(\Omega)$$
$$(-1)^m Y^*_{l,-m}(\Omega) = Y_{lm}(\Omega)$$

3.30.1 Examples

$$\int_{-1}^{1} P_k(x)\mathrm{d}x = \int_{-1}^{1} P_k(x)P_0(x)\mathrm{d}x = 2\delta_{k0}$$
$$\int Y_{k0}(\Omega)\mathrm{d}\Omega = \int Y_{k0}(\Omega)\sqrt{4\pi}Y_{00}(\Omega)\mathrm{d}\Omega = \sqrt{4\pi}\delta_{k0}$$

3.31 Gaunt Coefficients

We use the Wigner-Eckart theorem:

$$\langle jm|T^k_q|j'm'\rangle = (-1)^{j-m}\begin{pmatrix} j & k & j' \\ -m & q & m' \end{pmatrix}(j||T^k||j')$$

Where:

$$T^k_q = Y_{kq}$$

In order to calculate the reduced matrix element $(j||T^k||j')$, we evaluate the W-E theorem for $m = q = m' = 0$:

$$\langle j0|T^k_0|j'0\rangle = (-1)^{j}\begin{pmatrix} j & k & j' \\ 0 & 0 & 0 \end{pmatrix}(j||T^k||j')$$

and also evaluate the left hand side explicitly:

$$\langle j0|T^k_0|j'0\rangle = \langle j0|Y_{k0}|j'0\rangle = \int Y^*_{j0}(\Omega)Y_{k0}(\Omega)Y_{j'0}(\Omega)\mathrm{d}\Omega =$$
$$= \sqrt{\frac{(2j+1)(2k+1)(2j'+1)}{4\pi}}\frac{1}{4\pi}\int P_j(\cos\theta)P_k(\cos\theta)P_{j'}(\cos\theta)\sin\theta\mathrm{d}\theta\mathrm{d}\phi =$$
$$= \sqrt{\frac{(2j+1)(2k+1)(2j'+1)}{4\pi}}\frac{1}{2}\int_{-1}^{1} P_j(x)P_k(x)P_{j'}(x)\mathrm{d}x =$$
$$= \sqrt{\frac{(2j+1)(2k+1)(2j'+1)}{4\pi}}\begin{pmatrix} j & k & j' \\ 0 & 0 & 0 \end{pmatrix}^2$$

where we used (3.29.2). Comparing these two results, we get:

$$(j||T^k||j') = (-1)^{-j}\sqrt{\frac{(2j+1)(2k+1)(2j'+1)}{4\pi}}\begin{pmatrix} j & k & j' \\ 0 & 0 & 0 \end{pmatrix}$$

and finally:

$$\int Y^*_{jm}(\Omega)Y_{kq}(\Omega)Y_{j'm'}(\Omega)\mathrm{d}\Omega =$$
$$= \langle jm|T^k_q|j'm'\rangle = (-1)^{j-m}\begin{pmatrix} j & k & j' \\ -m & q & m' \end{pmatrix}(j||T^k||j') =$$
$$= (-1)^{j-m}\begin{pmatrix} j & k & j' \\ -m & q & m' \end{pmatrix}(-1)^{-j}\sqrt{\frac{(2j+1)(2k+1)(2j'+1)}{4\pi}}\begin{pmatrix} j & k & j' \\ 0 & 0 & 0 \end{pmatrix} =$$
$$= (-1)^{-m}\sqrt{\frac{(2j+1)(2k+1)(2j'+1)}{4\pi}}\begin{pmatrix} j & k & j' \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} j & k & j' \\ -m & q & m' \end{pmatrix}$$

In order to evaluate other integrals of spherical harmonics, we just use the above result, for example:

$$\int Y_{l_1 m_1}(\Omega) Y_{l_2 m_2}(\Omega) Y_{l_3 m_3}(\Omega) \mathrm{d}\Omega =$$
$$= (-1)^{m_1} \int Y^*_{l_1 -m_1}(\Omega) Y_{l_2 m_2}(\Omega) Y_{l_3 m_3}(\Omega) \mathrm{d}\Omega =$$
$$= (-1)^{m_1} (-1)^{-(-m_1)} \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_3 \\ -(-m_1) & m_2 & m_3 \end{pmatrix} =$$
$$= \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$$

This is the most symmetric relation. It was first obtained by *[Gaunt]* (equation (9), p. 194, where he expanded the $3j$ symbols, so his formula is more complex but equivalent to the above).

It is useful to incorporate the selection rule $m_1 + m_2 + m_3 = 0$ of the $3j$ symbols into the formula and we get:

$$c^k(l, m, l', m') = \sqrt{\frac{4\pi}{2k+1}} \int Y^*_{lm}(\Omega) Y_{k,m-m'}(\Omega) Y_{l'm'}(\Omega) \mathrm{d}\Omega =$$
$$= (-1)^{-m} \sqrt{\frac{4\pi}{2k+1}} \sqrt{\frac{(2l+1)(2k+1)(2l'+1)}{4\pi}} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l & k & l' \\ -m & m-m' & m' \end{pmatrix} =$$
$$= (-1)^{-m} \sqrt{(2l+1)(2l'+1)} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l & k & l' \\ -m & m-m' & m' \end{pmatrix}$$

From the other selection rules of the $3j$ symbols it follows, that the $c^k(l, m, l', m')$ coefficients are nonzero only when:

$$|l - l'| \le k \le l + l'$$
$$l + l' + k = \text{even integer}$$

3.31.1 Example I

$$c^0(l, m, l', m') = \sqrt{4\pi} \int Y^*_{lm}(\Omega) Y_{00}(\Omega) Y_{l'm'}(\Omega) \mathrm{d}\Omega = \delta_{ll'} \delta_{mm'}$$

3.31.2 Example II

$$\sum_{m=-l}^{l} c^k(l, m, l, m) = \sum_m \sqrt{\frac{4\pi}{2k+1}} \int Y^*_{lm}(\Omega) Y_{k0}(\Omega) Y_{lm}(\Omega) \mathrm{d}\Omega =$$
$$= \sqrt{\frac{4\pi}{2k+1}} \int \sum_m |Y_{lm}(\Omega)|^2 Y_{k0}(\Omega) \mathrm{d}\Omega =$$
$$= \sqrt{\frac{4\pi}{2k+1}} \frac{2l+1}{4\pi} \int Y_{k0}(\Omega) \mathrm{d}\Omega =$$
$$= \sqrt{\frac{4\pi}{2k+1}} \frac{2l+1}{4\pi} \sqrt{4\pi} \delta_{k0} =$$
$$= (2l+1) \delta_{k0}$$

3.31.3 Example III

$$
\begin{aligned}
c^k(l,m,l',m') = \sqrt{\frac{4\pi}{2k+1}} \int Y^*_{lm}(\Omega) Y_{k,m-m'}(\Omega) Y_{l'm'}(\Omega) \mathrm{d}\Omega = \\
= \sqrt{\frac{4\pi}{2k+1}} \int \Theta_{lm} \Phi^*_m \Theta_{k,m-m'} \Phi_{m-m'} \Theta_{l'm'} \Phi_{m'} \sin\theta \mathrm{d}\theta \mathrm{d}\phi = \\
= \sqrt{\frac{4\pi}{2k+1}} \int_0^\pi \Theta_{lm} \Theta_{k,m-m'} \Theta_{l'm'} \sin\theta \mathrm{d}\theta \int_0^{2\pi} \Phi^*_m \Phi_{m-m'} \Phi_{m'} \mathrm{d}\phi = \\
= \sqrt{\frac{4\pi}{2k+1}} \int_0^\pi \Theta_{lm} \Theta_{k,m-m'} \Theta_{l'm'} \sin\theta \mathrm{d}\theta \left(\frac{1}{\sqrt{2\pi}}\right)^3 \int_0^{2\pi} e^{-im\phi} e^{i(m-m')\phi} e^{im'\phi} \mathrm{d}\phi = \\
= \sqrt{\frac{4\pi}{2k+1}} \int_0^\pi \Theta_{lm} \Theta_{k,m-m'} \Theta_{l'm'} \sin\theta \mathrm{d}\theta \left(\frac{1}{\sqrt{2\pi}}\right)^3 \int_0^{2\pi} \mathrm{d}\phi = \\
= \sqrt{\frac{2}{2k+1}} \int_0^\pi \Theta_{lm} \Theta_{k,m-m'} \Theta_{l'm'} \sin\theta \mathrm{d}\theta
\end{aligned}
$$

3.31.4 Example IV

$$
\begin{aligned}
c^k(l,-m,l',-m') = \\
= (-1)^m \sqrt{(2l+1)(2l'+1)} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l & k & l' \\ m & -m+m' & -m' \end{pmatrix} = \\
= (-1)^m (-1)^{l+k+l'} \sqrt{(2l+1)(2l'+1)} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l & k & l' \\ -m & m-m' & m' \end{pmatrix} = \\
= (-1)^{-m} \sqrt{(2l+1)(2l'+1)} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l & k & l' \\ -m & m-m' & m' \end{pmatrix} = \\
c^k(l,m,l',m')
\end{aligned}
$$

Where we used the fact, that $l+k+l'$ is an even integer and $(-1)^m = (-1)^{-m}$. c^k is not symmetric in lm and $l'm'$:

$$
\begin{aligned}
c^k(l',m',l,m) \\
= (-1)^{-m'} \sqrt{(2l'+1)(2l+1)} \begin{pmatrix} l' & k & l \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l' & k & l \\ -m' & m'-m & m \end{pmatrix} = \\
= (-1)^{-m'} \sqrt{(2l+1)(2l'+1)} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l & k & l' \\ m & m'-m & -m' \end{pmatrix} = \\
= (-1)^{-m'} \sqrt{(2l+1)(2l'+1)} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l & k & l' \\ -m & m-m' & m' \end{pmatrix} = \\
= (-1)^{m-m'} (-1)^{-m} \sqrt{(2l+1)(2l'+1)} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l & k & l' \\ -m & m-m' & m' \end{pmatrix} = \\
= (-1)^{m-m'} c^k(l,m,l',m')
\end{aligned}
$$

Few other identities:

$$
c^k(l,0,l',0) = \sqrt{(2l+1)(2l'+1)} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2
$$

$$
\begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 = \frac{c^k(l,0,l',0)}{\sqrt{(2l+1)(2l'+1)}} = \frac{c^{l'}(l,0,k,0)}{\sqrt{(2l+1)(2k+1)}} = \frac{c^l(l',0,k,0)}{\sqrt{(2l'+1)(2k+1)}}
$$

$$
c^k(l,0,l',0) = c^k(l',0,l,0)
$$

3.31.5 Example V

$$\sum_{m'} \left(c^k(l,m,l',m')\right)^2 =$$

$$= \sum_{m'} (2l+1)(2l'+1) \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 \begin{pmatrix} l & k & l' \\ -m & m-m' & m' \end{pmatrix}^2 =$$

$$= (2l+1)(2l'+1) \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 \sum_{m'} \begin{pmatrix} l & k & l' \\ -m & m-m' & m' \end{pmatrix}^2 =$$

$$= (2l+1)(2l'+1) \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 \frac{1}{2l+1} =$$

$$= (2l'+1) \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 =$$

$$= \sqrt{\frac{2l'+1}{2l+1}} c^k(l',0,l,0)$$

3.31.6 Example VI

$$\sum_{m'} \sum_{q} \int Y_{l'm'}(\Omega) Y^*_{l'm'}(\Omega') Y_{kq}(\Omega) Y^*_{kq}(\Omega') Y_{lm}(\Omega') \mathrm{d}\Omega' =$$

$$= \int \frac{2l'+1}{4\pi} P_{l'}(\hat{\mathbf{x}} \cdot \hat{\mathbf{x}}') \frac{2k+1}{4\pi} P_k(\hat{\mathbf{x}} \cdot \hat{\mathbf{x}}') Y_{lm}(\Omega') \mathrm{d}\Omega' =$$

$$= \int \frac{2l'+1}{4\pi} \frac{2k+1}{4\pi} \sum_{\lambda=|l'-k|}^{\lambda=l'+k} \sqrt{\frac{2\lambda+1}{2l'+1}} c^k(l',0,\lambda,0) \frac{4\pi}{2\lambda+1} \sum_{\mu=-\lambda}^{\lambda} Y^*_{\lambda\mu}(\Omega') Y_{\lambda\mu}(\Omega) Y_{lm}(\Omega') \mathrm{d}\Omega' = \tag{3.31.6.1}$$

$$= \frac{2l'+1}{4\pi} \frac{2k+1}{4\pi} \sum_{\lambda=|l'-k|}^{\lambda=l'+k} \sqrt{\frac{2\lambda+1}{2l'+1}} c^k(l',0,\lambda,0) \frac{4\pi}{2\lambda+1} \sum_{\mu=-\lambda}^{\lambda} Y_{\lambda\mu}(\Omega) \delta_{\lambda l} \delta_{\mu m} =$$

$$= \frac{2k+1}{4\pi} \sqrt{\frac{2l'+1}{2l+1}} c^k(l',0,l,0) Y_{lm}(\Omega)$$

Where we used the following identities:

$$\sum_{m'} Y_{l'm'}(\Omega) Y^*_{l'm'}(\Omega') = \frac{2l'+1}{4\pi} P_{l'}(\hat{\mathbf{x}} \cdot \hat{\mathbf{x}}')$$

$$\sum_{q} Y_{kq}(\Omega) Y^*_{kq}(\Omega') = \frac{2k+1}{4\pi} P_k(\hat{\mathbf{x}} \cdot \hat{\mathbf{x}}')$$

$$P_k(\hat{\mathbf{x}} \cdot \hat{\mathbf{x}}') P_{l'}(\hat{\mathbf{x}} \cdot \hat{\mathbf{x}}') = \sum_{\lambda=|l'-k|}^{l'+k} \begin{pmatrix} k & l' & \lambda \\ 0 & 0 & 0 \end{pmatrix}^2 (2\lambda+1) P_\lambda(\hat{\mathbf{x}} \cdot \hat{\mathbf{x}}') =$$

$$= \sum_{\lambda=|l'-k|}^{\lambda=l'+k} \sqrt{\frac{2\lambda+1}{2l'+1}} c^k(l',0,\lambda,0) P_\lambda(\hat{\mathbf{x}} \cdot \hat{\mathbf{x}}') =$$

$$= \sum_{\lambda=|l'-k|}^{\lambda=l'+k} \sqrt{\frac{2\lambda+1}{2l'+1}} c^k(l',0,\lambda,0) \frac{4\pi}{2\lambda+1} \sum_{\mu=-\lambda}^{\lambda} Y^*_{\lambda\mu}(\Omega') Y_{\lambda\mu}(\Omega)$$

Note: using the integral of 3 spherical harmonics directly in (3.31.6.1):

$$\sum_{m'}\sum_{q}\int Y_{l'm'}(\Omega)Y^*_{l'm'}(\Omega')Y_{kq}(\Omega)Y^*_{kq}(\Omega')Y_{lm}(\Omega')\mathrm{d}\Omega' =$$
$$= \sum_{m'} Y_{l'm'}(\Omega)Y_{k,m-m'}(\Omega)\sqrt{\frac{4\pi}{2k+1}}c^k(l,m,l',m')$$

doesn't straightforwardly lead to the final result, as it is not obvious how to simplify things further.

3.32 Wigner 3j Symbols

Relation between the Wigner $3j$ symbols and Clebsch-Gordan coefficients:

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \frac{(-1)^{j_1-j_2-m_3}}{\sqrt{2j_3+1}}(j_1m_1j_2m_2|j_3-m_3)$$
$$(j_1m_1j_2m_2|j_3m_3) = (-1)^{j_1-j_2+m_3}\sqrt{2j_3+1}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{pmatrix}$$

They are nonzero only when:

$$m_1+m_2+m_3=0$$
$$j_1+j_2+j_3 = \text{integer (or even integer if } m_1=m_2=m_3=0)$$
$$|m_i| \le j_i$$
$$|j_1-j_2| \le j_3 \le j_1+j_2$$

They have lots of symmetries. The $3j$ symbol is invariant for an even permutation of columns:

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} =$$
$$= \begin{pmatrix} j_2 & j_3 & j_1 \\ m_2 & m_3 & m_1 \end{pmatrix} =$$
$$= \begin{pmatrix} j_3 & j_1 & j_2 \\ m_3 & m_1 & m_2 \end{pmatrix}$$

For an odd permutation of columns it changes sign if $j_1+j_2+j_3$ is an odd integer:

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} =$$
$$= (-1)^{j_1+j_2+j_3}\begin{pmatrix} j_2 & j_1 & j_3 \\ m_2 & m_1 & m_3 \end{pmatrix} =$$
$$= (-1)^{j_1+j_2+j_3}\begin{pmatrix} j_1 & j_3 & j_2 \\ m_1 & m_3 & m_2 \end{pmatrix} =$$
$$= (-1)^{j_1+j_2+j_3}\begin{pmatrix} j_3 & j_2 & j_1 \\ m_3 & m_2 & m_1 \end{pmatrix}$$

and the same if you change the sign of the second row:

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} =$$
$$= (-1)^{j_1+j_2+j_3}\begin{pmatrix} j_1 & j_2 & j_3 \\ -m_1 & -m_2 & -m_3 \end{pmatrix}$$

Orthogonality relations:

$$\sum_{m_1 m_2} \begin{pmatrix} j_1 & j_2 & j \\ m_1 & m_2 & m \end{pmatrix} \begin{pmatrix} j_1 & j_2 & j' \\ m_1 & m_2 & m' \end{pmatrix} = \frac{\delta_{jj'}\delta_{mm'}}{2j+1}$$

As a special case, we get:

$$\sum_{m'} \begin{pmatrix} l & k & l' \\ -m & m-m' & m' \end{pmatrix}^2 = \frac{1}{2l+1} \tag{3.32.1}$$

Here is a script to check that the equation (3.32.1) works:

```
from sympy import S
from sympy.physics.wigner import wigner_3j

def doit(l, k, lp, m):
    s = 0
    for mp in range(-lp, lp+1):
        s += wigner_3j(l, k, lp, -m, m-mp, mp)**2
    print "%2d %2d %2d %2d  " % (l, k, lp, m), s, " ", S(1)/(2*l+1)

k = 4
lp = 3
print " l  k  lp m:  lhs   rhs"
for l in range(1, 6):
    for m in range(-l, l+1):
        doit(l, k, lp, m)
```

it prints:

```
l  k  lp m:  lhs   rhs
1  4  3 -1   1/3   1/3
1  4  3  0   1/3   1/3
1  4  3  1   1/3   1/3
2  4  3 -2   1/5   1/5
2  4  3 -1   1/5   1/5
2  4  3  0   1/5   1/5
2  4  3  1   1/5   1/5
2  4  3  2   1/5   1/5
3  4  3 -3   1/7   1/7
3  4  3 -2   1/7   1/7
3  4  3 -1   1/7   1/7
3  4  3  0   1/7   1/7
3  4  3  1   1/7   1/7
3  4  3  2   1/7   1/7
3  4  3  3   1/7   1/7
4  4  3 -4   1/9   1/9
4  4  3 -3   1/9   1/9
4  4  3 -2   1/9   1/9
4  4  3 -1   1/9   1/9
4  4  3  0   1/9   1/9
4  4  3  1   1/9   1/9
4  4  3  2   1/9   1/9
4  4  3  3   1/9   1/9
4  4  3  4   1/9   1/9
5  4  3 -5   1/11   1/11
5  4  3 -4   1/11   1/11
5  4  3 -3   1/11   1/11
```

(continues on next page)

(continued from previous page)

```
5   4   3  -2    1/11    1/11
5   4   3  -1    1/11    1/11
5   4   3   0    1/11    1/11
5   4   3   1    1/11    1/11
5   4   3   2    1/11    1/11
5   4   3   3    1/11    1/11
5   4   3   4    1/11    1/11
5   4   3   5    1/11    1/11
```

Values of the $3j$ coefficients for a few special cases (use the symmetries above to obtain values for permuted symbols):

$$\begin{pmatrix} k & l & m \\ 0 & 0 & 0 \end{pmatrix} = (-1)^s \sqrt{\frac{(2s-2k)!(2s-2l)!(2s-2m)!}{(2s+1)!}} \frac{s!}{(s-k)!(s-l)!(s-m)!} \qquad \text{for } 2s = k+l+m \text{ even}$$

$$\begin{pmatrix} k & l & m \\ 0 & 0 & 0 \end{pmatrix} = 0 \qquad \text{for } 2s = k+l+m \text{ odd}$$

$$\begin{pmatrix} j+\frac{1}{2} & j & \frac{1}{2} \\ m & -m-\frac{1}{2} & \frac{1}{2} \end{pmatrix} = (-1)^{j-m-\frac{1}{2}} \sqrt{\frac{j-m+\frac{1}{2}}{(2j+1)(2j+2)}}$$

$$\begin{pmatrix} j+1 & j & 1 \\ m & -m-1 & 1 \end{pmatrix} = (-1)^{j-m-1} \sqrt{\frac{(j-m)(j-m+1)}{(2j+1)(2j+2)(2j+3)}}$$

$$\begin{pmatrix} j+1 & j & 1 \\ m & -m & 0 \end{pmatrix} = (-1)^{j-m-1} \sqrt{\frac{2(j+m+1)(j-m+1)}{(2j+1)(2j+2)(2j+3)}}$$

3.32.1 Examples

$$\begin{pmatrix} j_3-\frac{1}{2} & \frac{1}{2} & j_3 \\ m_3-\frac{1}{2} & \frac{1}{2} & -m_3 \end{pmatrix} = \begin{pmatrix} j_3 & j_3-\frac{1}{2} & \frac{1}{2} \\ -m_3 & m_3-\frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} j+\frac{1}{2} & j & \frac{1}{2} \\ m & -m-\frac{1}{2} & \frac{1}{2} \end{pmatrix}\Bigg|_{j=j_3-\frac{1}{2};m=-m_3} =$$

$$= (-1)^{j_3-\frac{1}{2}+m_3-\frac{1}{2}} \sqrt{\frac{j_3-\frac{1}{2}+m_3+\frac{1}{2}}{(2j_3-1+1)(2j_3-1+2)}} = (-1)^{j_3+m_3-1} \sqrt{\frac{j_3+m_3}{2j_3(2j_3+1)}}$$

$$\begin{pmatrix} j_3-\frac{1}{2} & \frac{1}{2} & j_3 \\ m_3+\frac{1}{2} & -\frac{1}{2} & -m_3 \end{pmatrix} = (-1)^{j_3-\frac{1}{2}+\frac{1}{2}+j_3} \begin{pmatrix} j_3 & j_3-\frac{1}{2} & \frac{1}{2} \\ m_3 & -m_3-\frac{1}{2} & \frac{1}{2} \end{pmatrix} = (-1)^{2j_3} \begin{pmatrix} j+\frac{1}{2} & j & \frac{1}{2} \\ m & -m-\frac{1}{2} & \frac{1}{2} \end{pmatrix}\Bigg|_{j=j_3-\frac{1}{2};m=m_3} =$$

$$= (-1)^{2j_3}(-1)^{j_3-\frac{1}{2}-m_3-\frac{1}{2}} \sqrt{\frac{j_3-\frac{1}{2}-m_3+\frac{1}{2}}{(2j_3-1+1)(2j_3-1+2)}} = (-1)^{2j_3}(-1)^{j_3-m_3-1} \sqrt{\frac{j_3-m_3}{2j_3(2j_3+1)}}$$

$$\begin{pmatrix} j_3+\frac{1}{2} & \frac{1}{2} & j_3 \\ m_3-\frac{1}{2} & \frac{1}{2} & -m_3 \end{pmatrix} = (-1)^{j_3+\frac{1}{2}+\frac{1}{2}+j_3} \begin{pmatrix} j_3+\frac{1}{2} & j_3 & \frac{1}{2} \\ m_3-\frac{1}{2} & -m_3 & \frac{1}{2} \end{pmatrix} = (-1)^{2j_3+1} \begin{pmatrix} j+\frac{1}{2} & j & \frac{1}{2} \\ m & -m-\frac{1}{2} & \frac{1}{2} \end{pmatrix}\Bigg|_{j=j_3;m=m_3-\frac{1}{2}} =$$

$$= (-1)^{2j_3+1}(-1)^{j_3-m_3+\frac{1}{2}-\frac{1}{2}} \sqrt{\frac{j_3-m_3+\frac{1}{2}+\frac{1}{2}}{(2j_3+1)(2j_3+2)}} = (-1)^{2j_3+1}(-1)^{j_3-m_3} \sqrt{\frac{j_3-m_3+1}{(2j_3+1)(2j_3+2)}}$$

$$\begin{pmatrix} j_3+\frac{1}{2} & \frac{1}{2} & j_3 \\ m_3+\frac{1}{2} & -\frac{1}{2} & -m_3 \end{pmatrix} = \begin{pmatrix} j_3+\frac{1}{2} & j_3 & \frac{1}{2} \\ -m_3-\frac{1}{2} & m_3 & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} j+\frac{1}{2} & j & \frac{1}{2} \\ m & -m-\frac{1}{2} & \frac{1}{2} \end{pmatrix}\Bigg|_{j=j_3;m=-m_3-\frac{1}{2}} =$$

$$= (-1)^{j_3+m_3+\frac{1}{2}-\frac{1}{2}} \sqrt{\frac{j_3+m_3+\frac{1}{2}+\frac{1}{2}}{(2j_3+1)(2j_3+2)}} = (-1)^{j_3+m_3} \sqrt{\frac{j_3+m_3+1}{(2j_3+1)(2j_3+2)}}$$

3.33 Multipole Expansion

Using (3.29.1.1) we get:

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|}=\sum_{l=0}^{\infty}\frac{r_<^l}{r_>^{l+1}}P_l(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}')=\sum_{l,m}\frac{r_<^l}{r_>^{l+1}}\frac{4\pi}{2l+1}Y_{lm}(\hat{\mathbf{r}})Y_{lm}^*(\hat{\mathbf{r}}')$$

where we used the formula:

$$\sum_m\langle\hat{\mathbf{r}}|lm\rangle\langle lm|\hat{\mathbf{r}}'\rangle=\frac{2l+1}{4\pi}\langle\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}'|P_l\rangle$$

3.34 Hypergeometric Functions

The series:

$$\sum_{k=0}^{\infty}t_k$$

with $t_0=1$ is geometric if the ratio of two consecutive terms t_{k+1}/t_k is a constant (with respect to k):

$$\frac{t_{k+1}}{t_k}=x$$

then we get:

$$\sum_{k=0}^{\infty}t_k=\sum_{k=0}^{\infty}x^k$$

It is hypergeometric if the ratio t_{k+1}/t_k is a rational function (with respect to k):

$$\frac{t_{k+1}}{t_k}=\frac{P(k)}{Q(k)}$$

where $P(k)$ and $Q(k)$ are polynomials in k, which we can completely factor into the form

$$\frac{t_{k+1}}{t_k}=\frac{P(k)}{Q(k)}=\frac{(k+a_1)(k+a_2)\cdots(k+a_p)}{(k+b_1)(k+b_2)\cdots(k+b_q)(k+1)}x \tag{3.34.1}$$

where x is a constant and the $(k+1)$ factor is just a convention (if the polynomial $Q(k)$ does not contain the factor $(k+1)$ we can just add it to both numerator and denominator and absorb the "1" into a_p). The hypergeometric series is then given by:

$${}_pF_q(a_1,a_2,\ldots,a_p;b_1,b_2,\ldots,b_q;x)=\sum_{k=0}^{\infty}\frac{(a_1)_k(a_2)_k\cdots(a_p)_k}{(b_1)_k(b_2)_k\cdots(b_q)_k}\frac{x^k}{k!}$$

where

$$(a)_k=\frac{\Gamma(a+k)}{\Gamma(a)}=\begin{cases}a(a+1)(a+2)\cdots(a+k-1), & \text{if } k\geq 1;\\ 1, & \text{if } k=0\end{cases}$$

is the rising factorial function (also called the Pochhammer symbol).

To write a function as a hypergeometric series, we simply expand it in series and then write the ratio t_{k+1}/t_k in the form (3.34.1) and immediately identify the proper ${}_pF_q$ function. If the ratio cannot be put into the form (3.34.1) then the function is not hypergeometric.

3.34.1 Convergence Conditions

If any $a_i = 0, -1, -2, \ldots$, then the series is a polynomial of degree $-a_i$.

If any $b_i = 0, -1, -2, \ldots$ then the denominators eventually become 0 (unless the series is terminated as a polynomial before that, due to the previous point) and the series is undefined.

Except the previous two cases, the radius of convergence R of the hypergeometric series is:

$$R = \begin{cases} \infty & \text{if } p \leq q; \\ 1 & \text{if } p = q + 1; \\ 0 & \text{if } p > q + 1. \end{cases}$$

3.34.2 Elementary and Special Functions

The hypergeometric functions for low p and q have special names:

${}_0F_1$	confluent hypergeometric limit function
${}_1F_1$	Kummer's confluent hypergeometric function of the first kind
${}_2F_1$	Gauss' hypergeometric function

Most common functions can be expressed using ${}_pF_q$ as follows:

The Series 0F0

Elementary functions:

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = {}_0F_0(x)$$

The Series 1F0

Elementary functions:

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k = {}_1F_0(1; x)$$

$$\frac{1}{(1-x)^a} = \sum_{k=0}^{\infty} \frac{(a+k-1)!}{(a-1)!k!} x^k = {}_1F_0(a; x)$$

$$x^a = {}_1F_0(-a; 1-x)$$

$$\sqrt{x} = {}_1F_0(-\tfrac{1}{2}; 1-x)$$

The Series 0F1

Elementary functions:

$$\sin z = z\ {}_0F_1(\tfrac{3}{2}; -\frac{z^2}{4})$$
$$\cos z = {}_0F_1(\tfrac{1}{2}; -\frac{z^2}{4})$$
$$\sinh z = z\ {}_0F_1(\tfrac{3}{2}; \frac{z^2}{4})$$
$$\cosh z = {}_0F_1(\tfrac{1}{2}; \frac{z^2}{4})$$

Bessel function:

$$J_\alpha(x) = \sum_{k=0}^{\infty} \frac{(-1)^k \left(\frac{x}{2}\right)^{2k+\alpha}}{k!(k+\alpha)!} = \frac{\left(\frac{x}{2}\right)^\alpha}{\Gamma(\alpha+1)}\ {}_0F_1\left(\alpha+1; -\frac{x^2}{4}\right)$$

Spherical Bessel function of the first kind:

$$j_\nu(x) = \sqrt{\frac{\pi}{2x}} J_{\nu+\frac{1}{2}}(x) = \frac{\sqrt{\pi}\left(\frac{x}{2}\right)^\nu}{2\Gamma(\nu+\frac{3}{2})}\ {}_0F_1\left(\nu+\frac{3}{2}; -\frac{x^2}{4}\right)$$

Modified Bessel functions:

$$I_\nu(z) = i^{-\nu} J_\nu(iz) = \sum_{k=0}^{\infty} \frac{\left(\frac{x}{2}\right)^{2k+\nu}}{k!(k+\nu)!} = \frac{1}{\Gamma(\nu+1)}\left(\frac{z}{2}\right)^\nu\ {}_0F_1\left(\nu+1; \frac{z^2}{4}\right)$$
$$K_\nu(z) = \frac{\Gamma(\nu)}{2}\left(\frac{2}{z}\right)^\nu\ {}_0F_1\left(1-\nu; \frac{z^2}{4}\right) + \frac{\Gamma(-\nu)}{2}\left(\frac{z}{2}\right)^\nu\ {}_0F_1\left(\nu+1; \frac{z^2}{4}\right)$$

The Series 1F1

Elementary functions:

$$z^a e^z = {}_1F_1(a; a-\tfrac{1}{2}; -2z)$$

Lower incomplete gamma function:

$$\gamma(z,x) = x^z\Gamma(z)e^{-x}\sum_{k=0}^{\infty}\frac{x^k}{\Gamma(z+k+1)} = x^z z^{-1} e^{-x}\ {}_1F_1(1; z+1; x) = x^z z^{-1}\ {}_1F_1(z; z+1; -x)$$

Error function:

$$\text{erf}(x) = \frac{1}{\sqrt{\pi}}\gamma(\tfrac{1}{2}, x^2) = \frac{2x}{\sqrt{\pi}}\ {}_1F_1(\tfrac{1}{2}; \tfrac{3}{2}, -x^2)$$

Hermite polynomials:

$$H_{2n}(x) = (-1)^n \frac{(2n)!}{n!}\ {}_1F_1(-n; \tfrac{1}{2}; x^2)$$
$$H_{2n+1}(x) = (-1)^n \frac{(2n+1)!}{n!} 2x\ {}_1F_1(-n; \tfrac{3}{2}; x^2)$$

Laguerre polynomials:

$$L_n^\alpha(x) = \binom{n+\alpha}{n}\ {}_1F_1(-n; \alpha+1; x) \tag{3.34.2.1}$$

Solution $P_{nl}(r) = rR_{nl}(r)$ of the radial Schrödinger equation in the Coulomb potential $V(r) = -Z/r$ (we use (3.34.2.1) in the second equation below):

$$P_{nl}(r) = N_{nl}\left(\frac{2Zr}{n}\right)^{l+1} e^{-\frac{Zr}{n}}\, {}_1F_1\left(-n+l+1; 2l+2; \frac{2Zr}{n}\right) =$$
$$= N_{nl}\left(\frac{2Zr}{n}\right)^{l+1} e^{-\frac{Zr}{n}}\, L_{n-l-1}^{2l+1}\left(\frac{2Zr}{n}\right) \frac{(2l+1)!(n-l-1)!}{(n+l)!} =$$
$$= \frac{1}{n}\sqrt{\frac{Z(n-l-1)!}{(n+l)!}}\left(\frac{2Zr}{n}\right)^{l+1} e^{-\frac{Zr}{n}}\, L_{n-l-1}^{2l+1}\left(\frac{2Zr}{n}\right)$$
$$N_{nl} = \frac{1}{n(2l+1)!}\sqrt{\frac{Z(n+l)!}{(n-l-1)!}}$$

The Series 2F1

Elementary functions:

$$\log(1+z) = z\, {}_2F_1(1,1;2;-z)$$
$$\log(z) = (z-1)\, {}_2F_1(1,1;2;1-z)$$
$$\arcsin z = z\, {}_2F_1(\tfrac{1}{2},\tfrac{1}{2};\tfrac{3}{2};z^2)$$
$$\arccos z = \frac{\pi}{2} - z\, {}_2F_1(\tfrac{1}{2},\tfrac{1}{2};\tfrac{3}{2};z^2)$$
$$\arctan z = z\, {}_2F_1(1,\tfrac{1}{2};\tfrac{3}{2};-z^2)$$

Legendre polynomials (and associated Legendre polynomials):

$$P_n(z) = {}_2F_1\left(-n, n+1; 1; \frac{1-z}{2}\right)$$
$$P_n^{\mu}(z) = \frac{1}{\Gamma(1-\mu)}\left(\frac{1+z}{1-z}\right)^{\frac{\mu}{2}} {}_2F_1\left(-n, n+1; 1-\mu; \frac{1-z}{2}\right)$$

Chebyshev polynomials:

$$T_n(z) = {}_2F_1\left(-n, n; \tfrac{1}{2}; \frac{1-z}{2}\right)$$
$$U_n(z) = (n+1)\, {}_2F_1\left(-n, n+2; \tfrac{3}{2}; \frac{1-z}{2}\right)$$

Gegenbauer polynomials:

$$C_n^{\alpha}(z) = \frac{(2\alpha)_n}{n!}\, {}_2F_1\left(-n, 2\alpha+n; \alpha+\tfrac{1}{2}; \frac{1-z}{2}\right)$$

Jacobi polynomials:

$$P_n^{(\alpha,\beta)}(z) = \frac{(\alpha+1)_n}{n!}\, {}_2F_1\left(-n, 1+\alpha+\beta+n; \alpha+1; \frac{1-z}{2}\right)$$

Complete elliptic integrals:

$$K(k) = \frac{\pi}{2}\, {}_2F_1(\tfrac{1}{2},\tfrac{1}{2};1;k^2)$$
$$E(k) = \frac{\pi}{2}\, {}_2F_1(-\tfrac{1}{2},\tfrac{1}{2};1;k^2)$$

The Series 3F2

Elementary functions:

$$\tan(z) = \frac{8z}{\pi^2 - 4z^2}\, {}_3F_2(1, \tfrac{1}{2} - \frac{z}{\pi}, \tfrac{1}{2} + \frac{z}{\pi}; \tfrac{3}{2} - \frac{z}{\pi}, \tfrac{3}{2} + \frac{z}{\pi}; 1)$$

Dilogarithm:

$$\mathrm{Li}_2(z) = z\, {}_3F_2(1, 1, 1; 2, 2; z)$$

Digamma:

$$\psi(z) = (z-1)\, {}_3F_2(1, 1, 2 - z; 2, 2; 1) - \gamma$$

The Wigner 3j symbol:

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = (-1)^{-j_1+j_2+m_3} \delta_{-m_3, m_1+m_2} \frac{1}{(-j_2+j_3+m_1)!(-j_1+j_3-m_2)!}$$
$$\frac{\sqrt{(j_1-j_2+j_3)!(-j_1+j_2+j_3)!(j_1+m_1)!(j_2-m_2)!(j_3+m_3)!(j_3-m_3)!}}{\sqrt{(j_1+j_2-j_3)!(j_1+j_2+j_3+1)!(j_1-m_1)!(j_2+m_2)!}}$$
$${}_3F_2(-j_1-j_2+j_3, m_1-j_1, -j_2-m_2; -j_1+j_3-m_2+1, -j_2+j_3+m_1+1; 1)$$

The Series pFq

Polylogarithm:

$$\mathrm{Li}_s(z) = z\, {}_{s+1}F_s(1, 1, \ldots, 1; 2, \ldots, 2; z)$$

Fermi-Dirac integral:

$$I_\nu(x) = \int_0^\infty \frac{t^\nu}{1 + e^{t-x}} \mathrm{d}t = -\Gamma(\nu+1)\mathrm{Li}_{\nu+1}(-e^x)$$

3.34.3 Example I

By writing out the series expansion for the t_{k+1}/t_k ratio we can prove that:

$$p\, {}_1F_1(a; b; x) + q\, {}_1F_1(a+1; b; x) = (p+q)\, {}_2F_2\left(a, a\left(\frac{p}{q}+1\right)+1; b, a\left(\frac{p}{q}+1\right); x\right)$$

The left hand side is equal to:

$$p\, {}_1F_1(a; b; x) + q\, {}_1F_1(a+1; b; x) = \sum_{k=0}^{\infty} \frac{p(a)_k + q(a+1)_k}{(b)_k k!} x^k$$

We simplify the t_k term:

$$t_k = \frac{p(a)_k + q(a+1)_k}{(b)_k k!} x^k = \frac{(a)_k \left(p + q + \frac{qk}{a}\right)}{(b)_k k!} x^k$$

We calculate the ratio t_{k+1}/t_k as well as t_0 to get the normalization:

$$t_0 = p + q$$
$$\frac{t_{k+1}}{t_k} = \frac{(k+a)\left(p + q + \frac{q(k+1)}{a}\right)}{(k+b)(k+1)\left(p + q + \frac{qk}{a}\right)} x = \frac{(k+a)\left(k + a\left(\frac{p}{q}+1\right)+1\right)}{(k+b)\left(k + a\left(\frac{p}{q}+1\right)\right)(k+1)} x$$

From which we read the arguments of the hypergeometric function ${}_2F_2$ on the right hand side and we need to multiply it by the normalization factor $t_0 = p + q$.

3.34.4 Example II

By writing out the series expansion for the t_{k+1}/t_k ratio we can prove that:

$$e^{-x}\ {}_1F_1(1;2;2x) = {}_0F_1\left(\tfrac{3}{2};\frac{x^2}{4}\right)$$

We can also use the substitution $z = \frac{x^2}{4}$:

$$e^{-2\sqrt{z}}\ {}_1F_1(1;2;4\sqrt{z}) = {}_0F_1\left(\tfrac{3}{2};z\right)$$

Which is a special case of

$${}_0F_1\left(a;z\right) = e^{-2\sqrt{z}}\ {}_1F_1(a-\tfrac{1}{2};2a-1;4\sqrt{z})$$

for $a = \frac{3}{2}$.

3.34.5 Example III

One way to express $\sinh(z)$ is:

$$\sinh z = ze^{-z}\ {}_1F_1(1;2;2z)$$

using the previous example, this is equal to:

$$\sinh z = ze^{-z}\ {}_1F_1(1;2;2z) = z\ {}_0F_1\left(\tfrac{3}{2};\frac{z^2}{4}\right)$$

So the lowest hypergeometric function that can express $\sinh(z)$ is ${}_0F_1$.

3.35 Feynman Parameters

When integrating a denominator like $\frac{1}{AB}$, the idea is to introduce auxiliary parameters in order to make the denominator simpler. We start with the identity:

$$\frac{1}{AB} = \int_0^1 \mathrm{d}x \frac{1}{\left(xA+(1-x)B\right)^2} = \int_0^1 \mathrm{d}x \int_0^1 \mathrm{d}y \frac{\delta(x+y-1)}{\left(xA+yB\right)^2} \tag{3.35.1}$$

which can be proven easily:

```
>>> var("A B")
(A, B)
>>> integrate(1/(x*A + (1-x)*B)**2, (x, 0, 1))
1/(A*B - A**2) - 1/(-A*B + B**2)
>>> simplify(_)
1/(A*B)
```

By repeatedly differentiating with respect to B:

$$\begin{aligned}
\frac{1}{AB^2} &= \int_0^1 \mathrm{d}x \int_0^1 \mathrm{d}y \frac{2y\delta(x+y-1)}{\left(xA+yB\right)^3} \\
\frac{1}{AB^3} &= \int_0^1 \mathrm{d}x \int_0^1 \mathrm{d}y \frac{3y^2\delta(x+y-1)}{\left(xA+yB\right)^4} \\
\frac{1}{AB^n} &= \int_0^1 \mathrm{d}x \int_0^1 \mathrm{d}y \frac{ny^{n-1}\delta(x+y-1)}{\left(xA+yB\right)^{n+1}}
\end{aligned} \tag{3.35.2}$$

Then we prove:

$$\frac{1}{A_1 A_2 \cdots A_n} = \int_0^1 \mathrm{d}x_1 \cdots \int_0^1 \mathrm{d}x_n \frac{(n-1)!\,\delta(x_1 + \cdots + x_n - 1)}{(x_1 A_1 + \cdots + x_n A_n)^n} \tag{3.35.3}$$

For $n = 2$ we get (3.35.1) and if it holds for n it also holds for $n+1$, because we multiply (3.35.3) by $\frac{1}{A_{n+1}}$ and get:

$$\begin{aligned}
&\frac{1}{A_1 A_2 \cdots A_n}\frac{1}{A_{n+1}} = \\
&= \int_0^1 \mathrm{d}x_1 \cdots \int_0^1 \mathrm{d}x_n (n-1)!\,\delta(x_1 + \cdots + x_n - 1)\frac{1}{(x_1 A_1 + \cdots + x_n A_n)^n A_{n+1}} = \\
&= \int_0^1 \mathrm{d}x_1 \cdots \int_0^1 \mathrm{d}x_n (n-1)!\,\delta(x_1 + \cdots + x_n - 1) \int_0^1 \mathrm{d}x \int_0^1 \mathrm{d}y \frac{n y^{n-1}\delta(x+y-1)}{(x A_{n+1} + y\,(x_1 A_1 + \cdots + x_n A_n))^{n+1}} \\
&= \int_0^1 \mathrm{d}x_1 \cdots \int_0^1 \mathrm{d}x_n \int_0^1 \mathrm{d}y \frac{n!\,\delta(x_1 + \cdots + x_n - 1) y^{n-1}}{((1-y) A_{n+1} + y\,(x_1 A_1 + \cdots + x_n A_n))^{n+1}} = \\
&= \int_0^{\frac{1}{y}} \mathrm{d}x_1 \cdots \int_0^{\frac{1}{y}} \mathrm{d}x_n \int_0^1 \mathrm{d}y \frac{n!\,\delta(y x_1 + \cdots + y x_n - y) y^n}{((1-y) A_{n+1} + y\,(x_1 A_1 + \cdots + x_n A_n))^{n+1}} = \\
&= \int_0^{\frac{1}{y}} y\mathrm{d}x_1 \cdots \int_0^{\frac{1}{y}} y\mathrm{d}x_n \int_0^1 \mathrm{d}y \frac{n!\,\delta(y x_1 + \cdots + y x_n - y)}{((1-y) A_{n+1} + (y x_1 A_1 + \cdots + y x_n A_n))^{n+1}} = \\
&= \int_0^1 \mathrm{d}z_1 \cdots \int_0^1 \mathrm{d}z_n \int_0^1 \mathrm{d}y \frac{n!\,\delta(z_1 + \cdots + z_n - y)}{((1-y) A_{n+1} + (z_1 A_1 + \cdots + z_n A_n))^{n+1}} = \\
&= -\int_0^1 \mathrm{d}z_1 \cdots \int_0^1 \mathrm{d}z_n \int_1^0 \mathrm{d}y' \frac{n!\,\delta(z_1 + \cdots + z_n + y' - 1)}{(y' A_{n+1} + (z_1 A_1 + \cdots + z_n A_n))^{n+1}} = \\
&= \int_0^1 \mathrm{d}x_1 \cdots \int_0^1 \mathrm{d}x_{n+1} \frac{n!\,\delta(x_1 + \cdots + x_{n+1} - 1)}{(x_1 A_1 + \cdots + x_{n+1} A_{n+1})^{n+1}}
\end{aligned}$$

Where we used (3.35.2) and the fact, that $\delta(x_1 + \cdots + x_n - 1) = y\,\delta(y x_1 + \cdots + y x_n - y)$, after the substituation we also restricted the limits of integration from 1 to $\frac{1}{y}$, since $x_1, x_2, \ldots$ are all positive.

3.35.1 Example 1

$$\int \frac{\mathrm{d}^4 k}{(k-p)^2 (k^2 - m^2)} = \int d^4 k \int_0^1 \mathrm{d}x \mathrm{d}y \frac{\delta(x+y-1)}{D^2}$$

where

$$D = x(k-p)^2 + y(k^2 - m^2) = (x+y)k^2 - 2xk \cdot p + xp^2 - ym^2 = k^2 - 2xk \cdot p + xp^2 - ym^2$$

In the last part we used $x + y = 1$. We now shift k by introducing:

$$\begin{aligned} l &= k - xp \\ \mathrm{d}^4 k &= \mathrm{d}^4 l \end{aligned}$$

and we get:

$$D = k^2 - 2xk \cdot p + xp^2 - ym^2 = l^2 - x^2 p^2 + xp^2 - ym^2$$

thus:

$$\begin{aligned}
&\int d^4 k \int_0^1 \mathrm{d}x \mathrm{d}y \frac{\delta(x+y-1)}{D^2} = \\
&= \int l^4 k \int_0^1 \mathrm{d}x \mathrm{d}y \frac{\delta(x+y-1)}{(l^2 - x^2 p^2 + xp^2 - ym^2)^2}
\end{aligned}$$

3.35.2 Example 2

$$\int \frac{\mathrm{d}^4k}{(k^2-m^2+i\epsilon)((k+p)^2-m^2+i\epsilon)((k-p)^2+i\epsilon)} = \int d^4k \int_0^1 \mathrm{d}x\mathrm{d}y\mathrm{d}z \frac{2\delta(x+y+z-1)}{D^3}$$

where

$$\begin{aligned} D &= x(k^2-m^2+i\epsilon)+y((k+p)^2-m^2+i\epsilon)+z((k-p)^2+i\epsilon) = \\ &= (x+y+z)k^2+2k\cdot(yq-zp)+yq^2+zp^2-(x+y)m^2+(x+y+z)i\epsilon = \\ &= k^2+2k\cdot(yq-zp)+yq^2+zp^2-(x+y)m^2+i\epsilon \end{aligned}$$

In the last part we used $x+y+z=1$. We now shift k by introducing:

$$\begin{aligned} l &= k+yq-zp \\ \mathrm{d}^4k &= \mathrm{d}^4l \end{aligned}$$

and we get:

$$\begin{aligned} D &= k^2+2k\cdot(yq-zp)+yq^2+zp^2-(x+y)m^2+i\epsilon = \\ &= l^2-\Delta+i\epsilon \end{aligned}$$

where

$$\Delta = -xyq^2+(1-z)^2m^2$$

thus:

$$\begin{aligned} &\int d^4k \int_0^1 \mathrm{d}x\mathrm{d}y\mathrm{d}z \frac{2\delta(x+y+z-1)}{D^3} = \\ &= \int \mathrm{d}^4l \int_0^1 \mathrm{d}x\mathrm{d}y\mathrm{d}z \frac{2\delta(x+y+z-1)}{(l^2-\Delta+i\epsilon)^3} = \\ &= (-i)\int \mathrm{d}^4l_E \int_0^1 \mathrm{d}x\mathrm{d}y\mathrm{d}z \frac{2\delta(x+y+z-1)}{(l_E^2+\Delta)^3} = \\ &= (-i)\int \mathrm{d}\Omega_4 \int_0^\infty \mathrm{d}l_E \int_0^1 \mathrm{d}x\mathrm{d}y\mathrm{d}z \frac{2\delta(x+y+z-1)l_E^3}{(l_E^2+\Delta)^3} = \\ &= (-i4\pi^2)\int_0^1 \mathrm{d}x\mathrm{d}y\mathrm{d}z\delta(x+y+z-1)\int_0^\infty \mathrm{d}l_E \frac{l_E^3}{(l_E^2+\Delta)^3} = \\ &= (-i4\pi^2)\int_0^1 \mathrm{d}x\mathrm{d}y\mathrm{d}z\delta(x+y+z-1)\int_\Delta^\infty \mathrm{d}h \frac{h-\Delta}{2h^3} = \\ &= (-i4\pi^2)\int_0^1 \mathrm{d}x\mathrm{d}y\mathrm{d}z\delta(x+y+z-1)\frac{1}{4\Delta} = \\ &= (-i\pi^2)\int_0^1 \mathrm{d}x\mathrm{d}y\mathrm{d}z \frac{\delta(x+y+z-1)}{\Delta} = \\ &= (-i\pi^2)\int_0^1 \mathrm{d}x\mathrm{d}y\mathrm{d}z \frac{\delta(x+y+z-1)}{(1-z)^2m^2-xyq^2} \end{aligned}$$

This integral has an infrared divergence. We can cure this by pretending that the photon has a small nonzero mass μ, then in the denominator of the photon propagator we need to change:

$$(k-p)^2 \to (k-p)^2-\mu^2$$

This denominator is multiplied by z later on, so at the end we need to do the change:

$$\Delta \to \Delta+z\mu^2$$

and we get:

$$(-i\pi^2)\int_0^1 \mathrm{d}x\mathrm{d}y\mathrm{d}z\frac{\delta(x+y+z-1)}{(1-z)^2m^2-xyq^2} \to$$
$$\to (-i\pi^2)\int_0^1 \mathrm{d}x\mathrm{d}y\mathrm{d}z\frac{\delta(x+y+z-1)}{(1-z)^2m^2-xyq^2+z\mu^2}$$

for $q^2=0$ we get:

$$(-i\pi^2)\int_0^1 \mathrm{d}x\mathrm{d}y\mathrm{d}z\frac{\delta(x+y+z-1)}{(1-z)^2m^2-xyq^2+z\mu^2} \to$$
$$\to (-i\pi^2)\int_0^1 \mathrm{d}x\mathrm{d}y\mathrm{d}z\frac{\delta(x+y+z-1)}{(1-z)^2m^2+z\mu^2} =$$
$$= (-i\pi^2)\int_0^1 \mathrm{d}z\int_0^{1-z}\mathrm{d}y\frac{1}{(1-z)^2m^2+z\mu^2} =$$
$$= (-i\pi^2)\int_0^1 \mathrm{d}z\frac{1-z}{(1-z)^2m^2+z\mu^2}$$

We can use the following integral:

$$\int_0^1 \frac{1-z}{1-2z+z^2+z\mu^2}\,dz = \frac{1}{2}\log\left(\mu^{-2}\right)+\frac{\operatorname{atan}\left(\frac{1}{\sqrt{-1+\frac{4}{\mu^2}}}\right)}{\sqrt{-1+\frac{4}{\mu^2}}}-\frac{\operatorname{atan}\left(\frac{1-\frac{2}{\mu^2}}{\sqrt{-1+\frac{4}{\mu^2}}}\right)}{\sqrt{-1+\frac{4}{\mu^2}}}$$

that is equal to $\frac{1}{2}\log(\frac{1}{\mu^2})$ in the limit $\mu \to 0$.

here are a few special cases for $\mu=1$, $\mu=1/2$ and $\mu=1/3$:

$$\int_0^1 \frac{1-z}{1-z+z^2}\,dz = \frac{1}{9}\pi\sqrt{3}$$
$$\int_0^1 \frac{1-z}{1-\frac{7}{4}z+z^2}\,dz = \frac{1}{2}\log(4)+\frac{1}{15}\sqrt{15}\operatorname{atan}\left(\frac{1}{15}\sqrt{15}\right)+\frac{1}{15}\sqrt{15}\operatorname{atan}\left(\frac{7}{15}\sqrt{15}\right)$$
$$\int_0^1 \frac{1-z}{1-\frac{17}{9}z+z^2}\,dz = \frac{1}{2}\log(9)+\frac{1}{35}\sqrt{35}\operatorname{atan}\left(\frac{1}{35}\sqrt{35}\right)+\frac{1}{35}\sqrt{35}\operatorname{atan}\left(\frac{17}{35}\sqrt{35}\right)$$

Code:

```
>>> from sympy import log, atan, var, sqrt, Eq, Integral, S
>>> var("z m mu")
>>> F = -log(z*(1 - 2/m) + 1/m + z**2/m)/2 + \
        atan((1 - 2/m + 2*z/m)/sqrt(-1 + 4/m))/sqrt(-1 + 4/m)
>>> f = F.diff(z).simplify()
>>> print f
(1 - z)/(1 - 2*z + m*z + z**2)
>>> integ_f_0_1 = F.subs(z, 1) - F.subs(z, 0)
>>> e = Eq(Integral(f.subs(m, mu**2), (z, 0, 1)), integ_f_0_1.subs(m, mu**2))
>>> print e
Integral((1 - z)/(1 - 2*z + z**2 + z*mu**2), (z, 0, 1)) == log(mu**(-2))/2 + atan((-1␣
↪+ 4/mu**2)**(-1/2))/(-1 + 4/mu**2)**(1/2) - atan((1 - 2/mu**2)/(-1 + 4/mu**2)**(1/
↪2))/(-1 + 4/mu**2)**(1/2)
>>> print e.subs(mu, 1)
Integral((1 - z)/(1 - z + z**2), (z, 0, 1)) == pi*3**(1/2)/9
>>> print e.subs(mu, S(1)/2)
```

(continues on next page)

(continued from previous page)

```
Integral((1 - z)/(1 - 7*z/4 + z**2), (z, 0, 1)) == log(4)/2 + 15**(1/2)*atan(15**(1/
↪2)/15)/15 + 15**(1/2)*atan(7*15**(1/2)/15)/15
>>> print e.subs(mu, S(1)/3)
Integral((1 - z)/(1 - 17*z/9 + z**2), (z, 0, 1)) == log(9)/2 + 35**(1/2)*atan(35**(1/
↪2)/35)/35 + 35**(1/2)*atan(17*35**(1/2)/35)/35
```

Then for $m = 1$ and small μ we get:

$$(-i\pi^2)\int_0^1 \mathrm{d}z \frac{1-z}{(1-z)^2 m^2 + z\mu^2} = \\ = (-i\pi^2)\tfrac{1}{2}\log\frac{1}{\mu^2}$$

3.36 Groups

These are notes of Karel Výborný and Ondřej Čertík on the group theory as a result of the first VDNK (Výprava do neznámých krajů) held between October 30 and November 2, 2006 in Prague. So that the next time we look at it we should be able to quickly recover our forgotten ideas.

3.36.1 Theory

Definition of a group:

- I1: $x, y \in G \Rightarrow xy \in G$
- I2: there exist e such that $ex = xe = x$ for each $x \in G$
- I3: there exist x^{-1} such that $xx^{-1} = x^{-1}x = e$ for each $x \in G$
- I4: $(xy)z = x(yz)$ for each $x, y, z \in G$

Every finite group is isomorphic to a subgroup of S_n (permutations).

Representation

Set of linear operators $T(x)$ (for each $x \in G$ there is one $T(x)$)

$$T(x)T(y) = T(xy), \quad T(e) = \mathbb{1}\,.$$

$T(x)$ fulfills all the group axioms I1, I2, I3, I4. The requirement $T(e) = 1$ is non-trivial, consider for example the following 4 matrices

$$T(\boldsymbol{\sigma}) = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & 0 \end{pmatrix}, \quad T(e) = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & 0 \end{pmatrix},$$

that fulfill $T(x)T(y) = T(xy)$ but not $T(e) = \mathbb{1}$.

The representation $T(x)$ is said to be *faithful* if there is a one-to-one relationship between $T(x)$ and x (an isomorphism).

Equivalent representations T_1 and T_2: there exist S such that $T_2 = ST_1(x)S^{-1}$ for each $x \in G$.

Reducible representation $T(x)$: there exist an equivalent representation that is diagonal:

$$T'(x) = ST(x)S^{-1} = \begin{pmatrix} T_1' & 0 \\ 0 & T_2' \end{pmatrix}, \qquad \forall x \in G\,. \tag{3.36.1.1}$$

We say that T' is a direct sum of T'_1 and T'_2: $T' = T'_1 \oplus T'_2$.

Irreducible representation: is not reducible.

Conjugate element: x is conjugate to y ($x \sim y$) if there exist c such that:

$$x = cyc^{-1}$$

if $x \sim y$ and $y \sim z$ then $x \sim z$.

Conjugate class: elements which are all conjugate to each other

No element may belong to more than one class $\Rightarrow$ every group may be broken up into separate classes.

Character χ of the representation $T(x)$: set of numbers $\chi(x) = \operatorname{Tr} T(x)$ as the group element x runs through the group.

Equivalent representations have the same character:

$$\chi'(x) = \operatorname{Tr} T'(x) = \operatorname{Tr} ST(x)S^{-1} = \operatorname{Tr} T(x) = \chi(x)$$

Representations having the same character are equivalent.

Proof: Characters can be thought of as elements of a q-dimensional vector space where q is the number of conjugacy classes. Using the orthogonality relations derived above, we find that the q characters for the q inequivalent irreducible representations forms a basis set. Also, according to Maschke's theorem, both representations can be expressed as the direct sum of irreducible representations. Since the character of the direct sum of representations is the sum of their characters, from linear algebra, we see they are equivalent.

All the elements in the same class have the same character.

Maschke's theorem: for finite groups, every class of equivalent representations contains unitary representations. The theorem is also true for most infinite groups of interest in physics.

Let T be a reducible representation, then:

$$T = m_1 T^{(1)} \oplus m_2 T^{(2)} \oplus m_3 T^{(3)} \oplus \cdots$$

where $T^{(1)}$, $T^{(2)}$, $T^{(3)}$ dots are all the inequivalent irreducible representations and m_α ($\alpha = 1, 2, 3, \ldots$) gives the number of times that the irreducible representation $T^{(\alpha)}$ occurs in the reduction.

Similar relation holds for group characters:

$$\chi = m_1 \chi^{(1)} + m_2 \chi^{(2)} + m_3 \chi^{(3)} + \cdots$$

and it can be shown *[Elliott]* (eq. 4.28, page 63):

$$\begin{aligned} m_\alpha &= \frac{1}{g} \sum_{x \in G} \chi^{(\alpha)*}(x) \chi(x) = \\ &= \frac{1}{g} \sum_p c_p \chi_p^{(\alpha)*} \chi_p \end{aligned}$$

where c_p is the number of elements in the class p, g is the number of elements in G (the order of the group).

Example

Consider the S_3 permutation group. The character table is:

S_3	e	$3(12)$	$2(123)$
A_1	1	1	1
A_2	1	-1	1
E	2	0	-1

From the table we read $c_1 = 1$, $c_2 = 3$, $c_3 = 2$ and $g = c_1 + c_2 + c_3 = 1 + 3 + 2 = 6$. There are 3 classes and 3 irreducible representations.

Case I We are given a representation given by the following matrices:

$$e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad a = \tfrac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}, \qquad b = \tfrac{1}{2}\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix},$$
$$c = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad d = \tfrac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}, \qquad f = \tfrac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}.$$

These 6 matrices belong to the following three classes $\{e\}$, $\{a, b, c\}$, $\{d, f\}$ and the corresponding characters for each class are:

$$\chi_1 = 2$$
$$\chi_2 = 0$$
$$\chi_3 = -1$$

and we get:

$$m_1 = \frac{1}{6}(1 \cdot 1 \cdot 2 + 3 \cdot 1 \cdot 0 + 2 \cdot 1 \cdot (-1)) = 0$$
$$m_2 = \frac{1}{6}(1 \cdot 1 \cdot 2 + 3 \cdot (-1) \cdot 0 + 2 \cdot 1 \cdot (-1)) = 0$$
$$m_3 = \frac{1}{6}(1 \cdot 2 \cdot 2 + 3 \cdot 0 \cdot 0 + 2 \cdot (-1) \cdot (-1)) = 1$$

So this representation is irreducible and it is equivalent to $m_1 A_1 \oplus m_2 A_2 \oplus m_3 E = E$.

Case II We are given a representation given by the following matrices:

$$e = d = f = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad a = b = c = \tfrac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}.$$

These 6 matrices belong to the following three classes $\{e\}$, $\{a, b, c\}$, $\{d, f\}$ and the corresponding characters for each class are:

$$\chi_1 = 2$$
$$\chi_2 = 0$$
$$\chi_3 = 2$$

and we get:

$$m_1 = \frac{1}{6}(1 \cdot 1 \cdot 2 + 3 \cdot 1 \cdot 0 + 2 \cdot 1 \cdot 2) = 1$$
$$m_2 = \frac{1}{6}(1 \cdot 1 \cdot 2 + 3 \cdot (-1) \cdot 0 + 2 \cdot 1 \cdot 2) = 1$$
$$m_3 = \frac{1}{6}(1 \cdot 2 \cdot 2 + 3 \cdot 0 \cdot 0 + 2 \cdot (-1) \cdot 2) = 0$$

So this representation is reducible and it is equivalent to $m_1 A_1 \oplus m_2 A_2 \oplus m_3 E = A_1 \oplus A_2$. The matrices are equivalent to:

$$e = d = f = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad a = b = c = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Other facts

Number of irreducible representations = the number of classes.

Regular representation of G: Take R^n with $n = \#G$ and assign a canonical basis to the elements g_i of G. A matrix A_a assigned to $a \in G$ now describes the mapping $(g_1, g_2, \ldots) \mapsto (ag_1, ag_2, \ldots)$, i.e. in if $ag_1 = g_5$, then the fifth element of the first row is one and others of that row are zero in A_a. Each IR of the reg. rep. occurs in its decomposition with the multiplicity equal to its dimension. Thus (p. 65, *[Sternberg]*)

$$\#G = \sum p_i^2 \,.$$

Element Order The order n of an element g is the least integer for which $g^n = e$. The order n can be determined from the group multiplication table for example. Theorem: n must divide the size (order) of the group (for finite groups). Example: For a group of six elements, the only possible orders are 1, 2, 3 and 6. Note: the element order is the same for the whole conjugacy class because: $x^n = \left(cyc^{-1}\right)^n = cy^n c^{-1} = cec^{-1} = e$.

Schur's lemma (a) Be r an IR of G. If $[r(a), T] = 0, \forall a \in G$, then $T = cI$.

(b) Be r_1, r_2 two inequivalent IRs of G. Then $r_1(a)T = Tr_2(a)$ valid $\forall a \in G$ implies $T = 0$. See p. 55 in *[Sternberg]*. This can be used to derive the orthogonality relations for characters.

Complete reducibility Every rep can be decomposed into IRs: true for finite (p. 52) and compact (p. 179 in *[Sternberg]*) groups. Counterexample for larger groups, p. 53.

Sum of reps. Opposite process to reduction, $\rho \oplus \sigma$, it lives on the direct sum of the two vector spaces of ρ and σ.

Take an IR ρ of G. Then ρ will also be a rep. of any subgroup $H \subset G$, but it need not be an IR, because the condition for reducibility, Eq. (3.36.1.1), is less strict: it suffices if the matrices $T(g)$ are simultaneously block diagonal only for $g \in H$, not for all $g \in G$. This is called *restriction* and it is denoted by $\downarrow$.

Induced representation, denoted by $\uparrow$, is an opposite of the restriction. It works as follows: if $F = G \otimes H$, then $\rho(f) = \rho(g)$, when $f = g \otimes h$.

Product of representations, $\rho \otimes \sigma$ lives on the direct product of the two vector spaces. Product of IRs need not be an IR. Most prominent example: adding of angular momenta.

Interesting examples

O and T_d (see *Crystallographic Point Groups*) are isomorphic to S_4 (p. 35 in *[Sternberg]*). Written as matrices in 3D, they are 3D representations. Since O has only $\det A = 1$ matrices unlike T_d, they are inequivalent.

Homeomorphism of $SL(2, C)$ into the Lorentz group [or $SU(2)$ into $SO(3)$], p. 7 *[Sternberg]*}. Start with the following $1-1$ correspondence between $\vec{x}$ and x:

$$\vec{x} = (x_0, x_1, x_2, x_3)^T \,, \qquad x = \begin{pmatrix} x_0 + x_3 & x_1 - ix_2 \\ x_1 + ix_2 & x_0 - x_3 \end{pmatrix} \,.$$

For any matrix of $A \in SL(2, C)$ take $AxA^* = x'$. Decode x' into $\vec{x}'$ and the relation between $\vec{x}$ and $\vec{x}'$ defines uniquely a Lorentz transformation; thus A was mapped into some Lorentz group element. If $x_0 = 0$ this gives a mapping from $SU(2)$ into $SO(3)$. The mapping is $2-1$ because A and $-A$ give the same x'.

$SO(3)$ is not simply connected. Consider matrices $U_\theta = diag(e^{-i\theta}, e^{i\theta}) \in SU(2), \theta \in [0, \pi]$. These map into $SO(3)$ rotations by 2θ around the z–axis. These matrices $A_\theta = R_{z,2\theta}$ in $SO(3)$ form a closed loop, $R_{z,0} = R_{z,2\pi}$. If $SO(3)$ were simply connected it would be possible to contract this loop into a point while keeping A_0 and A_π unchanged. But then the same would have to happen with the original curve of matrices U_θ while keeping U_0 and U_π at their place. Since $U_\pi = -I \neq U_0 = I$, this curve is not closed and such a contraction is not possible.

All IRs of S_3 are in *[Sternberg]*, p. 57.

3.36.2 Crystallographic Point Groups

Point group is a subgroup of O(3).

Crystallographic point groups are all subgroups of $O(3)$, which leave a monoatomic crystal lattice invariant. Those can be symmetries of an infinite crystal (e.g. C_5 is excluded since pentagons cannot cover the plane).

There are only 7 crystallographic point groups: S_2 (triclinic), C_{2h} (monoclinic), D_{2h} (orthorhombic), D_{3d} (rhombohedral), D_{4h} (tetragonal), D_{6h} (hexagonal) and O_h (cubic).

For simple monoatomic crystals with one atom per unit cell these seven are the only possible crystallographic point groups. For more complicated crystals with a molecule or an arrangement of atoms in the unit cell, the symmetry will be reduced to the subgroup which leaves not only the lattice but also the unit cell invariant.

The complete list of all possible crystallographic point groups will therefore be given by the above seven together with all their subgroups (Tab. 3 in *[Birss]* or Tab. 4 in *[Sternberg]*):

S_2	C_{1h}, S_2
C_{2h}	C_2, C_{1h}, C_{2h}
D_{2h}	D_2, C_{2v}, D_{2h}
D_{3d}	$C_3, S_6, D_3, C_{3v}, D_{3d}$
D_{4h}	$C_4, S_4, C_{4h}, D_4, C_{4v}, D_{2d}, D_{4h}$
D_{6h}	$C_3, S_6, D_3, C_{3v}, D_{3d}, C_6, C_{3h}, C_{6h}, D_6, C_{6v}, D_{3h}, D_{6h}$
O_h	T, T_h, O, T_d, O_h

There are 37 subgroups together. D_{3d} is a subgroup of D_{6h} (so all 5 subgroups of D_{3d} are also subgroups of D_{6h}). Together we get 37-5 = 32 distinct subgroups. Groups, which might at first sight appear to be missing from the list are C_{1v}, D_1, D_{1h}, S_1, and S_3, but these are the same as C_{1h}, C_2, C_{2v}, C_{1h} and C_{3h} respectively.

The following groups are isomorphic:

C_{1h}, S_2, C_2

S_4, C_4

S_6, C_{3h}, C_6

C_{2h}, C_{2v}, D_2

C_{3v}, D_3

D_{2d}, C_{4v}, D_4

D_{3d}, D_{3h}, C_{6v}, D_6

T_d, O

The way to derive the above lists is the following.

Procedure:

1. Find all finite crystallographic subgroups of $SO(3)$ called *rotation subgroups*
2. Take each subgroup from 1) and add $-I$ and close the subgroup ('non-rot containing $-I$')
3. for each subgroup $G^\wedge$ in 1), find whether it has some normal subgroups G^+ of index 2 (half a size of $G^\wedge$) and construct $G^+ \cup (-I)aG^+$, where $a \notin G^+$ and $a \in G^\wedge$; this will be a 'non-rot not containing $-I$' (for each $G^\wedge$ there can be zero, one or more such G^+).

The sum of 1.,2.,3. are all finite crystallographic groups of $O(3)$. The procedure is described in *[Sternberg]*, p. 28-40.

An example: O (all rot. symm. of a cube, i.e. no mirroring) is 1), O^h (all symm. of a cube) is 2) made of O and T_d (all symm. of a tetrahedron) is 3) made of 1).

Zoology

Schönflies notation: C_n is an n–fold rotation (2_z, 3_z ...) group (planar polygon), D_n is a diedric group, i.e. C_n plus turn-the-page two-fold rotations (e.g. 2_x, $2_\perp$), T, O and I (= Y) are the rotational symmetries of a tetrahedron, octahedron (identical to those of a cube) and icosahedron (identical to those of a dodecahedron), respectively. Additional indexes mean reflection planes, horizontal, vertical, diagonal (h,v,d) or $-I$ (i). Some atypical notation: $S_2 = C_i$, $S_6 = C_{3i}$, $S_4 = C_{2i}$, $C_s = C_{1h}$.

Hermann–Mauguin (HM, international) notation: 2,3,4 means C_n, $\bar{4}$ means rotation-inversion axis (rotation followed by $-I$), m is a vertical mirror plane, $/m$ is a horizontal mirror plane.

Symmetry operations (in Table 3 of *[Birss]*): like HM, 2_x means a two-fold rotation around x–axis, $2_\perp$ means some other axis in the xy plane than x, y or xy (diagonal), $\bar{3}_z$ is a rotation followed by $-I$. $3(2_\perp)$ means three different two-fold axes $2_\perp$.

Construction and usage of the character table

For simpler groups the character tables can be fully constructed by the following rules:

1. The sum of the squares of the dimensions n_i of the irreducible representations is equal to the order g of the point group:

$$\sum_{\mu=1}^{k} n_\mu^2 = g$$

   The dimension n_μ is given by the character of the identity matrix (first column) $n_\mu = \chi^\mu(E)$, so the sum of squares of the first column is g. It is customary to put the characters of the one dimensional representation ($\chi^1(C_i) = 1$) into the first row, so the first row is filled with 1s. Also, n_i must divide g.

2. The number of irreducible representations r (rows) is equal to the number of classes k (columns)

3. The rows must satisfy

$$\sum_{i=1}^{k} g_i \chi^\mu(C_i)\chi^\nu(C_i)^* = g\delta_{\mu\nu}$$

4. The columns must satisfy

$$\sum_{\nu=1}^{k} \chi^\nu(C_i)\chi^\nu(C_j)^* = \frac{g}{g_i}\delta_{ij}$$

5. Characters of all one-dimensional representation must be roots of unity, equal to $\chi = e^{\frac{2\pi ik}{n}}$, where n is the element order, which must divide the group size g (and it is the same for the whole conjugacy class). In general, k is any integer (for faithful representation it would be $k = 1$). This follows from the fact, that the character is also the one-dimensional representation matrix and they all commute, thus the group is Abelian. Also, the characters (=representation matrices) must respect the group operations, so for example if for two group elements $g_1^2 = g_2$, then their characters must also obey $\chi_1^2 = \chi_2$.

6. Character of an element is the complex conjugate of its inverse. If they both belong to the same conjugacy class, the character must be real. If the character is complex, it means that its inverse is not from the same conjugacy class and then there must be a complex conjugate for another conjugacy class from the same irreducible representation.

7. Characters come in complex conjugate pairs, since complex conjugate of a representation is also a representation. If there is only one representation of the dimension d, then it must be real (since it is its own conjugate). If there are two representations of dimension d and one is complex, then the other one must be its complex conjugate. Another way to look at this is that if we conjugate each entry of the character table, then we must get the same character table (up to a possible reordering of rows within the same dimension).
8. If there is one dimensional representation A_1 (with characters χ_1) and any other representation T of dimension d (with characters χ), then there must be a representation of dimension d with characters $\chi_1\chi$ (corresponding to the tensor product $A_1 \otimes T$).

There exists a systematic approach that works for any group, but it is complicated (see for example *[Dixon67]*, *[Blokker72]*, *[Cannon69]* and *[Chillag86]*).

The notation for irreducible representation: One-dimensional irreducible representations are labeled either A or B according to whether the character of a $\frac{2\pi}{n}$ (proper or improper) rotation about the symmetry axis of highest order n is $+1$ or -1. If there is no symmetry axis, all one-dimensional representations are labeled A.

For general information, see *[Elliott]* (sec. 4.15, page 67) and *[Bishop]*, page 128.

Example I

Let's take the group C_{3v}, which has three classes E (1 element), C_3 (2 elements) and σ_v (3 elements).

So $g_1 = 1$, $g_2 = 2$ and $g_3 = 3$ and the order is $g = g_1 + g_2 + g_3 = 6$. Therefore it has three irreducible representations, whose dimensions must satisfy:

$$n_1^2 + n_2^2 + n_3^2 = 6$$

The only integer solution (up to a permutation) is $n_1 = n_2 = 1$ and $n_3 = 2$. So we immediately have:

C_{3v}	E	$2C_3$	$3\sigma_v$
	1	1	1
	1	a	b
	2	c	d

The rule 3. generates the following equations for all μ and ν:

μ	ν	$\sum_{i=1}^{k} g_i \chi^{\mu}(C_i)\chi^{\nu}(C_i)^* = g\delta_{\mu\nu}$
1	1	$6 = 6$
1	2	$1 + 2a + 3b = 0$
1	3	$2 + 2c + 3d = 0$
2	2	$1 + 2a^2 + 3b^2 = 6$
2	3	$2 + 2ac + 3bd = 0$
3	3	$4 + 2c^2 + 3d^2 = 6$

Solving all these equations simultaneously, we get two independent solutions. One is:

$$\begin{aligned} a &= 1 \\ b &= -1 \\ c &= -1 \\ d &= 0 \end{aligned}$$

and the other is:

$$a = -\frac{7}{5}$$
$$b = \frac{3}{5}$$
$$c = \frac{1}{5}$$
$$d = -\frac{4}{5}$$

The rule 4. generates the following equations for all i and j:

i	j	$\sum_{\nu=1}^{k} \chi^{\nu}(C_i)\chi^{\nu}(C_j)^* = \frac{g}{g_i}\delta_{ij}$
1	1	$6 = 6$
1	2	$1 + a + 2c = 0$
1	3	$1 + b + 2d = 0$
2	2	$1 + a^2 + c^2 = 3$
2	3	$1 + ab + cd = 0$
3	3	$1 + b^2 + d^2 = 2$

Both of the above solutions for (a, b, c, d) satisfy all of these equations, so the column equations are redundant.

Now we use the rule 5. and see that the second solution is not a root of unity, so we discard it. The final character table is:

C_{3v}	E	$2C_3$	$3\sigma_v$
A_1	1	1	1
A_2	1	1	-1
E	2	-1	0

Code:

```
from sympy import var, solve, Matrix
var("a, b, c, d")
g = [1, 2, 3]
M = Matrix([
    [1, 1, 1],
    [1, a, b],
    [2, c, d]])

def rows(mu, nu, M, g):
    eq = 0
    for i in range(len(g)):
        eq += g[i] * M[mu, i] * M[nu, i]
    if mu == nu:
        eq -= sum(g)
    return eq

def cols(i, j, M, g):
    eq = 0
    for nu in range(len(g)):
        eq += M[nu, i] * M[nu, j]
    if i == j:
        eq -= sum(g) / g[i]
    return eq

print "Character table:"
```

(continues on next page)

(continued from previous page)

```
print M
print "Rows conditions:"
eqs = []
for mu in range(3):
    for nu in range(mu, 3):
        eq = rows(mu, nu, M, g)
        eqs.append(eq)
        print mu+1, nu+1, ":    ", eq
print "-"*40
print "Columns conditions:"
eqs2 = []
for i in range(3):
    for j in range(i, 3):
        eq = cols(i, j, M, g)
        eqs2.append(eq)
        print i+1, j+1, ":    ", eq
print "-"*40
print "Solving the 1, 2, 4, 5 equations out of 0..5 from the rows conditions"
s = solve(eqs[1:3]+eqs[4:], [a, b, c, d])
print s
print "Test that all the solutions satisfy the rest of the equations:"
for a, b, c, d in s:
    print
    print "Solution:", a, b, c, d
    r = eqs[3].subs({
        "a": a,
        "b": b,
        "c": c,
        "d": d,
        })
    print "Equation 3 from rows conditions, result: ", r
    assert r == 0
    print "Columns conditions:"
    for i, eq in enumerate(eqs2):
        r = eq.subs({
                "a": a,
                "b": b,
                "c": c,
                "d": d,
                })
        print "Equation %i from columns conditions, result: %r" % (i, r)
        assert r == 0
```

This prints:

```
Character table:
[1, 1, 1]
[1, a, b]
[2, c, d]
Rows conditions:
1 1 :     0
1 2 :     2*a + 3*b + 1
1 3 :     2*c + 3*d + 2
2 2 :     2*a**2 + 3*b**2 - 5
2 3 :     2*a*c + 3*b*d + 2
3 3 :     2*c**2 + 3*d**2 - 2
----------------------------------------
```

(continues on next page)

(continued from previous page)

```
Columns conditions:
1 1 :     0
1 2 :     a + 2*c + 1
1 3 :     b + 2*d + 1
2 2 :     a**2 + c**2 - 2
2 3 :     a*b + c*d + 1
3 3 :     b**2 + d**2 - 1
--------------------------------------
Solving the 1, 2, 4, 5 equations out of 0..5 from the rows conditions
[(-7/5, 3/5, 1/5, -4/5), (1, -1, -1, 0)]
Test that all the solutions satisfy the rest of the equations:

Solution: -7/5 3/5 1/5 -4/5
Equation 3 from rows conditions, result:  0
Columns conditions:
Equation 0 from columns conditions, result: 0
Equation 1 from columns conditions, result: 0
Equation 2 from columns conditions, result: 0
Equation 3 from columns conditions, result: 0
Equation 4 from columns conditions, result: 0
Equation 5 from columns conditions, result: 0

Solution: 1 -1 -1 0
Equation 3 from rows conditions, result:  0
Columns conditions:
Equation 0 from columns conditions, result: 0
Equation 1 from columns conditions, result: 0
Equation 2 from columns conditions, result: 0
Equation 3 from columns conditions, result: 0
Equation 4 from columns conditions, result: 0
Equation 5 from columns conditions, result: 0
```

Example II

We derive the character table for C_{3v} again, using another approach. First we determine the element orders, that must divide the size of the group (possible values are 1, 2, 3, 6). Element order of the class E is 1, because $E^2 = 1$. The element order of C_3 is 3, because $C_3^3 = 1$. Finally, the element order of σ_v is 2, because $\sigma_v^2 = 1$.

C_{3v}	E	$2C_3$	$3\sigma_v$
class sizes	1	2	3
element orders	1	3	2
A_1	1	1	1
A_2	1	a	b
E	2	c	d

Rule 7: The characters of the representation A_2 must be real, because otherwise A_1 would have to be a complex conjugate. E is the only representation of dimension 2, so it must be real as well.

Rule 5: A_2 is Abelian, with element orders 1, 3 and 2. As such, we must have:

$$a = e^{\frac{2\pi i k}{3}}$$

$$b = e^{\frac{2\pi i l}{2}}$$

Where k and l are unknown integers. However, since both a and b is real, the only solution is $k = 0$ (corresponding to $a = 1$) and $l = 0, 1$ (corresponding to $b = \pm 1$).

Rule 3 gives:

$$1 + 2a + 3b = 0$$

And plugging in $a = 1$ this implies $b = -1$, consistent with the previous paragraph.

Rule 8: multiplying A_2 by E must give characters of dimension 2, which is E, so we get:

$$+1 \cdot c = c$$
$$-1 \cdot d = d$$

From which $d = 0$. Rule 3 gives:

$$2 + 2c + 3d = 0$$

Where we use $d = 0$ and we get $c = -1$. The final character table is:

C_{3v}	E	$2C_3$	$3\sigma_v$
A_1	1	1	1
A_2	1	1	-1
E	2	-1	0

Example III

We derive the character table for C_2.

C_2	E	C_2
class sizes	1	1
element orders	1	2
A_1	1	1
A_2	1	a

We have two classes, group order is 2, so we must have two representations of dimension 1. Using the rule 3. we get:

$$1 + a = 0$$

so $a = -1$ and the final character table is:

C_2	E	C_2
A_1	1	1
A_2	1	-1

Example IV

We derive the character table for C_3.

C_3	E	C_3	C_3^2
class sizes	1	1	1
element orders	1	3	3

We have 3 classes and representations, group order is 3, so they must be one dimensional:

C_3	E	C_3	C_3^2
class sizes	1	1	1
element orders	1	3	3
A_1	1	1	1
A_2	1	a	b
A_3	1	c	d

Rule 3 says:

$$1 + a + b = 0$$

Rule 5 says:

$$a = \omega^k$$
$$b = \omega^l$$

where $\omega = e^{\frac{2\pi i}{3}}$, so:

$$1 + \omega^k + \omega^l = 0$$

Which has only two solutions: $k = 1$, $l = 2$ and $k = 2$, $l = 1$. If we choose the first solution, we get $a = \omega$ and $b = \omega^2 = \bar{\omega}$. Using the rule 7. it follows that $c = \bar{a} = \bar{\omega} = \omega^2$ and $d = \bar{b} = \omega$. If we choose the second solution, we get the pairs a, b and c, d interchanged, however, we can reorder the rows, so these two options are equivalent. The final character table is:

C_3	E	C_3	C_3^2
A_1	1	1	1
A_2	1	ω	ω^2
A_3	1	ω^2	ω

$$\omega = e^{\frac{2\pi i}{3}} = \frac{-1 + i\sqrt{3}}{2}$$

Example V

Group C_4:

C_4	E	C_4	C_4^2	C_4^3
class sizes	1	1	1	1
element orders	1	4	2	4
A_1	1	1	1	1
A_2	1	a	b	c
A_3	1			
A_4	1			

Rule 5 gives:

$$a = i^k$$
$$b = (-1)^l$$
$$c = i^m$$

Rule 3 gives:

$$1 + a + b + c = 0 \tag{3.36.2.1}$$

Using the rule 7. we know that at least one of A_2, A_3 and A_4 must be real, so let it be A_2. The only real solutions of the equation (3.36.2.1) are $a = 1$, $b = -1$, $c = -1$ and permutations. The representation however must be isomorphic to the C_4 group, so in particular $a^2 = b$, from which $b = 1$ and then $a = -1$ and $c = -1$.

The group operations give:

$$a^2 = b$$
$$ab = c$$
$$ac = 1$$

which gives:

$$\begin{aligned} 2k &= l \\ k + l &= m \\ k + m &= 0, 4, 8, 12, ... \end{aligned}$$

The possible solutions are:

k	l	m
2	0	2
1	2	3
3	2	1

The first solution is real and it is equal to A_2. The other two solutions are complex conjugate and they must be solutions of A_3 and A_4, because A_3 and A_4 cannot be real (otherwise they would have to be equal to A_2 and the orthogonality relation for columns would not hold). The final character table is:

C_4	E	C_4	C_4^2	C_4^3
A_1	1	1	1	1
A_2	1	-1	1	-1
A_3	1	i	-1	$-i$
A_4	1	$-i$	-1	i

Example VI

Group T:

T	E	$4C_3$	$4C_3^2$	$3C_2$
class sizes	1	4	4	3
element orders	1	3	3	2
A_1	1	1	1	1
A_2	1	a	b	c
A_3	1			
T	3	d	e	f

The group size is 1 + 4 + 4 + 3 = 12, so the only possible option for dimensions of the 4 representations is 1, 1, 1 and 3.

Rule 5 gives:

$$\begin{aligned} a &= \omega^k \\ b &= \omega^l \\ c &= (-1)^m \end{aligned}$$

where $\omega = e^{\frac{2\pi i}{3}}$. Rule 3 gives:

$$1 + 4\omega^k + 4\omega^l + 3(-1)^m = 0$$

The only solution is $m = 0$, $k = 1$ and $l = 2$ (and k with l interchanged). This fully determines A_2 and A_3. The last row is determined from column orthogonality conditions (we compare the given column with the first column):

$$\begin{aligned} 1 + \omega + \omega^2 + 3d &= 0 \\ 1 + \omega^2 + \omega + 3e &= 0 \\ 1 + 1 + 1 + 3f &= 0 \end{aligned}$$

Using the relation $1 + \omega + \omega^2 = 0$ we get $d = 0$, $e = 0$ and $f = -1$.

The final character table is:

T	E	$4C_3$	$4C_3^2$	$3C_2$
A_1	1	1	1	1
A_2	1	ω	ω^2	1
A_3	1	ω^2	ω	1
T	3	0	0	-1

3.36.3 Applications of finite groups

Distinct energy levels ('vibrations')

Assume that we know number of atoms in a molecule and measure the number of its distinct vibrational modes (frequencies) in a multiplet. We want to know its symmetry.

We go through the list of all possible (point) symmetries S of such a molecule and look at what all reps. S has. If an n–tuplet was observed among the vibrational modes and there is no n-dimensional IR of S, then can be excluded.

This procedure assumes that (a) the original symmetry S is slightly disturbed because of something and (b) two multiplets (m and n dimensional) do not accidentally happen to have the same frequencies ('accidental degeneracy').

Selection rules ('transitions')

According to *[Pilar]*, p. 572.

Probability of an optical transition is proportional to

$$\langle i|H_1|f\rangle \ , \tag{3.36.3.1}$$

where $|i\rangle$, $|f\rangle$ are the initial and final states and H_1 is the operator of the interaction causing the transition. This is the Fermi golden rule (first order time dep. perturbation theory).

The integral ((3.36.3.1)) may vanish because of the symmetry. A simple 1D example is that $|f\rangle$ is an even function $f(x)$, $|i\rangle$ is an odd function $i(x)$ and H_1 is an even function $h_1(x)$. Then $i^*(x)h_1(x)f(x)$ is odd and thus the integral over $(-\infty, \infty)$ vanishes. The group theory only generalizes this observation.

The procedure is: find the IRs ρ_i, ρ_f to which $|i\rangle$, $|f\rangle$ belong and also ρ, the regular rep of H_1 in order to catch all IRs of H_1 (is this procedure correct?). Then construct $\rho_i \otimes \rho \otimes \rho_f$, decompose it into IRs and see if the trivial rep is present. If not, the integral ((3.36.3.1)) vanishes. This procedure is claimed to be equivalent to checking whether $\rho_i \otimes \rho_f$ and ρ contain at least one common IR.

The infrared absorption (IRa) is described by $H_1 \propto x$ (or y, z, depending probably on the polarization of light), the Raman scattering has $H_1 \propto x^2$ (it comes from the second order perturbation theory?).

Zoology

Todo:

- Describe the representations A_1, A_2, B_1, E etc.
- Reps are specified by the generating functions $f(x, y, z)$ and the symmetry operations T acting on these functions $f(x, y, z) \mapsto f(x', y', z')$ then transform the arguments, $(x, y, z) \mapsto (x', y', z') = T(x, y, z)$. Explain what functions are commonly used (x, R_x, ...) and give maybe some examples.
- Further reading: Davydov, p. 318, 195. Joe Penrose: Symmetry in Science.

3.36.4 Continuous Groups

Lie groups+algebras

A continuous group with metrics is a Lie group (more exactly a differentiable manifold and $a \mapsto ag$ and $a \mapsto a^{-1}$ are differentiable $\forall g$, p. 172 in *[Sternberg]*) usually a subgroup of $GL(n)$ is meant, a linear Lie group (i.e. matrices). Peter–Weyl theorem (p. 179 in *[Sternberg]*) looks like that compact Lie groups are practically as nice as finite groups.

Consider $G = O(n)$, p. 234 in *[Sternberg]*. If $A \in G$ then $\exp(-tA) \subset G$ where $t \in R$. At least in $O(3)$ and probably in any $O(n)$, any element of G can be written as $\exp(-tA)$ where A is a $\pi/2$ rotation around some axis. These A's are the generators of G.

Typical example: for $G = SO(3)$ there are three generators, iA_x, iA_y, iA_z, where A_x is the rotation by $\pi/2$ around x–axis in R^3. The generators form a vector space (here the linear span of iA_x, iA_y, iA_z) with an additional operation of commutation. This structure is closed and it is called the Lie algebra of the group G. The commutation relations between the generators fully specify the Lie algebra. E.g. $[iA_x, iA_y] = iA_z$ and the two other ones.

This is a great simplification because a continuous (infinite) group was thus mapped on a vector space, the algebra, where it suffices to look at the basis elements, the generator. The net effect is that we have to watch only three objects instead of infinitely many in the example above.

Todo: weights, roots and Dynkin diagrams. Octets and decuplets. Classification of IRs of $SU(n)$. From *[Georgi]*.

IRs of SU(2)

16. 181 in *[Sternberg]*; alternative somewhere in *[Georgi]*.

The Peter-Weyl theorem concerns also the orthogonality of characters and that in turn strongly restricts any possible characters of $SU(2)$. The conjugacy classes of $SU(2)$ are exemplified by matrices $U_\theta = \mathrm{diag}(e^{i\theta}, e^{-i\theta})$ and their possible characters can only be

$$\chi(\theta) = \sum_{k=-s}^{s} \exp(-i2k\theta)$$

with $2s$ integer.

All the corresponding reps exist, they are defined on the space $z_1^{2s}, z_1^{2s-1}z_2, \ldots, z_2^{2s}$ by $U_{-\theta} z_1^{2s-k} z_2^k \mapsto [\exp i(2s - 2k)\theta] z_1^{2s-k} z_2^k$.

For an IR of $SU(2)$ the complex conjugate is just the original. For other $SU(n)$ it is not necessarily the case, p. 182 in *[Sternberg]*.

IRs of $SO(3)$ are just those of $SU(2)$ but s must be an integer.

Young diagrams

YD is a systematic method to find all IRs of any symmetric group S_n (permutations of an n-element set). The idea:

- find all conjugacy classes of S_n
- assign an IR to each of them

Char'n of the conjugacy classes: each permutation can be decomposed into cycles. This cycle structure (i.e. how many cycles of length 1, how many of length 2, etc. $= [\nu_1, \nu_2 \ldots, \nu_n]$) is a unique mark of each conjugation class. The Young diagram is written by rows, each row has λ_i empty boxes and $\lambda_i - \lambda_{i+i} = \nu_i \geq 0$. Each conjugacy class has one YD. An YD of S_n has n boxes.

A Young tabloid (YTd) is obtained by filling an YD with numbers $1, \ldots, n$ where ordering in each row does not matter. A Young tableau is an YTd where all orderings (thus also in rows) matter.

The IRs of S_n. Take an YD λ. On the space of all corresponding YTd's (M_λ) a rep. of the S_n is created. It is decomposed into IRs and shown to have some 'new' IR compared to $\mu > \lambda$.

Details are explained in *[Sternberg]*, p. 76 or in the lecture notes of J. Niederle.

Comments from p. 82 of *[Sternberg]*: Basis of M_λ is defined (e_t; $\delta_{\{t\}}$ means probably a function on M_λ which is zero for all $\{y\}$ unless $\{y\} = \{t\}$). The action of $a \in S_n$ on this basis functions is described.

3.36.5 Literature

Books:

Articles:

3.37 Wigner D Function

The Wigner D function gives the matrix elements of the rotation operator R in the jm-representation. For the Euler angles α, β, γ, the D function is defined as:

$$\langle j, m|R(\alpha, \beta, \gamma)|j', m'\rangle = \delta_{jj'} D(j, m, m', \alpha, \beta, \gamma)$$

Where the rotation operator $R(\alpha, \beta, \gamma)$ is defined using the z-y-z convention:

$$R(\alpha, \beta, \gamma) = e^{-i\alpha J_z} e^{-i\beta J_y} e^{-i\gamma J_z}$$

Here J_i is the projection of the total angular momentum on an i-axis. The $|jm\rangle$ is the eigenstate of the operators J^2 and J_z. Using the fact that $e^{-i\gamma J_z}|jm\rangle = e^{-im\gamma}|jm\rangle$, we can see that the Wigner D function can always be written using the Wigner small-d function as:

$$\begin{aligned} D(j, m, m', \alpha, \beta, \gamma) = \langle j, m|R(\alpha, \beta, \gamma)|j, m'\rangle = \langle j, m|e^{-i\alpha J_z} e^{-i\beta J_y} e^{-i\gamma J_z}|j, m'\rangle = \\ = e^{-im\alpha} \langle j, m|e^{-i\beta J_y}|j, m'\rangle e^{-im'\gamma} = e^{-im\alpha} d(j, m, m', \beta) e^{-im'\gamma} \end{aligned}$$

where

$$d(j, m, m', \beta) = \langle j, m|e^{-i\beta J_y}|j, m'\rangle$$

We can use the following relations to evaluate $d(j, m, m', \beta)$:

$$d(j, m, m', \beta) = i^{2j-m-m'}(-1)^{2m} \sum_{m''=-j}^{j} d(j, m, m'', \frac{\pi}{2}) e^{-im''\beta} d(j, m'', -m', \frac{\pi}{2})$$

$$d(j, m, m', \frac{\pi}{2}) = (-1)^{m-m'} \frac{1}{2^j} \sqrt{\frac{(j+m)!(j-m)!}{(j+m')!(j-m')!}} \sum_k (-1)^k \binom{j+m'}{k} \binom{j-m'}{k+m-m'}$$

3.37.1 Derivation

The small-d function formula above can be derived from the following formula:

$$d(j, m, m', \beta) = \sum_k (-1)^k \frac{\sqrt{(j+m)!(j-m)!(j+m')!(j-m')!}}{(j-m'-k)!(j+m-k)!k!(k+m'-m)!} \cos^{2j+m-m'-2k}\frac{\beta}{2} \sin^{2k+m'-m}\frac{\beta}{2}$$

by substituting

$$a = +e^{-\frac{1}{2}i\alpha}\cos\frac{\beta}{2}e^{-\frac{1}{2}i\gamma}$$
$$b = -e^{-\frac{1}{2}i\alpha}\sin\frac{\beta}{2}e^{+\frac{1}{2}i\gamma}$$

into

$$\sum_k (-1)^k \frac{\sqrt{(j+m)!(j-m)!(j+m')!(j-m')!}}{(j-m'-k)!(j+m-k)!k!(k+m'-m)!} a^{j-m'-k} a^{*j+m-k} b^k b^{*k+m'-m}$$

This follows from:

$$\epsilon' = a\epsilon + b\zeta$$
$$\zeta' = -b^*\epsilon + a^*\zeta$$

let the polynomial be:

$$f_m(\epsilon,\zeta) = \frac{\epsilon^{j+m}\zeta^{j-m}}{\sqrt{(j+m)!(j-m)!}}$$

and (using binomial theorem in the process):

$$\mathbf{P_u} f_m(\epsilon,\zeta) = f_m(a^*\epsilon - b\zeta, b^*\epsilon + a\zeta) = \frac{(a^*\epsilon - b\zeta)^{j+m}(b^*\epsilon + a\zeta)^{j-m}}{\sqrt{(j+m)!(j-m)!}} =$$
$$= \sum_{k=0}^{j+m}\sum_{k'=0}^{j-m} (-1)^k \frac{\sqrt{(j+m)!(j-m)!}}{k!k'!(j+m-k)!(j-m-k')!} a^{k'} a^{*j+m-k} b^k b^{*j-m-k'} \epsilon^{2j-k-k'} \zeta^{k+k'} =$$
$$= \sum_{m'}\sum_k (-1)^k \frac{\sqrt{(j+m)!(j-m)!(j+m')!(j-m')!}}{(j-m'-k)!(j+m-k)!k!(k+m'-m)!} a^{j-m'-k} a^{*j+m-k} b^k b^{*k+m'-m} f_{m'}(\epsilon,\zeta)$$

And it is the coefficient of $f_{m'}$.

3.38 Ordinary Differential Equations

3.38.1 Finite Difference Formulas

We define the backward difference operator ∇_h by:

$$\nabla_h f(a) = f(a) - f(a-h)$$

Repeated application gives:

$$\nabla_h^2 f(a) = \nabla_h(f(a) - f(a-h)) = f(a) - f(a-h) - f(a-h) + f(a-2h) = f(a) - 2f(a-h) + f(a-2h)$$
$$\nabla_h^3 f(a) = f(a) - 3f(a-h) + 3f(a-2h) - f(a-3h)$$
$$\nabla_h^n f(a) = \sum_{k=0}^{n} \binom{n}{k} (-1)^k f(a-kh)$$

We can also derive a formula for $f(a+t)$ where t is any real number, independent of h:

$$f(a-h) = (1-\nabla_h) f(a)$$
$$(1-\nabla_h)^{-1} f(a-h) = f(a)$$
$$(1-\nabla_h)^{-1} f(a) = f(a+h)$$
$$(1-\nabla_h)^{-n} f(a) = f(a+nh)$$
$$(1-\nabla_h)^{-\frac{t}{h}} f(a) = f(a+t)$$

Now we can express the following general integral using the function value from either left ($f(a)$) or right ($f(a+h)$) hand side of the interval h:

$$\int_a^{a+h} f(t)\,\mathrm{d}t = \int_0^h f(a+t)\,\mathrm{d}t = \int_0^h (1-\nabla_h)^{-\frac{t}{h}} f(a)\,\mathrm{d}t =$$
$$= -\frac{h\nabla_h}{(1-\nabla_h)\log(1-\nabla_h)} f(a) =$$
$$= h\left(1 + \tfrac{1}{2}\nabla_h + \frac{5}{12}\nabla_h^2 + \frac{3}{8}\nabla_h^3 + \cdots\right) f(a) =$$
$$= -\frac{h\nabla_h}{\log(1-\nabla_h)}(1-\nabla_h)^{-1} f(a) = -\frac{h\nabla_h}{\log(1-\nabla_h)} f(a+h) =$$
$$= h\left(1 - \tfrac{1}{2}\nabla_h - \frac{1}{12}\nabla_h^2 - \frac{1}{24}\nabla_h^3 + \cdots\right) f(a+h)$$

Code:

```
>>> from sympy import var, simplify, integrate
>>> var("nabla t h")
(nabla, t, h)
>>> s = integrate((1-nabla)**(-t/h), (t, 0, h))
>>> simplify(s)
h*nabla/(-log(1 - nabla) + nabla*log(1 - nabla))
>>> s.series(nabla, 0, 5)
h + h*nabla/2 + 5*h*nabla**2/12 + 3*h*nabla**3/8 + 251*h*nabla**4/720 + O(nabla**5)
>>> s2 = s*(1-nabla)
>>> simplify(s2)
-h*nabla/log(1 - nabla)
>>> s2.series(nabla, 0, 5)
h - h*nabla/2 - h*nabla**2/12 - h*nabla**3/24 - 19*h*nabla**4/720 + O(nabla**5)
```

Keeping terms only to third-order, we obtain:

$$\int_a^{a+h} f(t)\,\mathrm{d}t = -\frac{h\nabla_h}{(1-\nabla_h)\log(1-\nabla_h)} f(a) \approx h\left(1 + \tfrac{1}{2}\nabla_h + \frac{5}{12}\nabla_h^2 + \frac{3}{8}\nabla_h^3\right) f(a) =$$
$$= hf(a) + h\tfrac{1}{2}\left(f(a) - f(a-h)\right) + h\frac{5}{12}\left(f(a) - 2f(a-h) + f(a-2h)\right) +$$
$$+h\frac{3}{8}\left(f(a) - 3f(a-h) + 3f(a-2h) - f(a-3h)\right) =$$
$$= h\left(1 + \tfrac{1}{2} + \frac{5}{12} + \frac{3}{8}\right) f(a) - h\left(\tfrac{1}{2} + \frac{2\cdot 5}{12} + \frac{3\cdot 3}{8}\right) f(a-h) +$$
$$+h\left(\frac{5}{12} + \frac{3\cdot 3}{8}\right) f(a-2h) - h\left(\frac{3}{8}\right) f(a-3h) =$$
$$= h\frac{55}{24} f(a) - h\frac{59}{24} f(a-h) + h\frac{37}{24} f(a-2h) - h\frac{3}{8} f(a-3h) =$$
$$= \frac{h}{24}\left(55f(a) - 59f(a-h) + 37f(a-2h) - 9f(a-3h)\right)$$

Similarly:

$$\int_a^{a+h} f(t)\,\mathrm{d}t = -\frac{h\nabla_h}{\log(1-\nabla_h)} f(a+h) \approx h\left(1 - \tfrac{1}{2}\nabla_h - \frac{1}{12}\nabla_h^2 - \frac{1}{24}\nabla_h^3\right) f(a+h) =$$
$$= \frac{h}{24}\left(9f(a+h) + 19f(a) - 5f(a-h) + f(a-2h)\right)$$

Code:

```
>>> from sympy import var
>>> var("f0 f1 f2 f3")
(f0, f1, f2, f3)
>>> nabla1 = f0 - f1
>>> nabla2 = f0 - 2*f1 + f2
>>> nabla3 = f0 - 3*f1 + 3*f2 - f3
>>> 24*(f0 + nabla1/2 + 5*nabla2/12 + 3*nabla3/8)
-59*f1 - 9*f3 + 37*f2 + 55*f0
>>> 24*(f0 - nabla1/2 - nabla2/12 - nabla3/24)
f3 - 5*f2 + 9*f0 + 19*f1
```

3.38.2 Integrating ODE

Set of linear ODEs can be written in the form:

$$\frac{\mathrm{d}y}{\mathrm{d}r} = Gy \tag{3.38.2.1}$$

For example for the Schrödinger we have

$$y = \begin{pmatrix} P \\ Q \end{pmatrix}$$
$$G = \begin{pmatrix} 0 & 1 \\ -2(E-V) - \frac{l(l+1)}{r^2} & 0 \end{pmatrix}$$

Now we need to choose a grid $r = r(t)$, where t is some uniform grid. For example $r = r_0(e^t - 1)$:

$$r_i = r_0(e^{t_i} - 1)$$
$$t_i = (i-1)h$$

where $i = 1, 2, 3, \ldots, N$. We also need the derivative, for the exampe above we get:

$$\frac{\mathrm{d}r}{\mathrm{d}t} = r_0 e^t$$

Now we substitute this into (3.38.2.1):

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\mathrm{d}r}{\mathrm{d}t} Gy$$

We can integrate this system from a to $a + h$ on a uniform grid t_i:

$$y(a+h) = y(a) + \int_a^{a+h} \frac{\mathrm{d}r}{\mathrm{d}t} Gy \,\mathrm{d}t = y(a) + \int_a^{a+h} f(t) \,\mathrm{d}t$$

where $f(t) = \frac{\mathrm{d}r}{\mathrm{d}t} Gy$ and we use some method to approximate the integral, see the previous section.

3.38.3 Radial Poisson Equation

Radial Poisson equation is:

$$V''(r) + \frac{2}{r} V'(r) = -4\pi n(r) \tag{3.38.3.1}$$

The left hand side can be written as:

$$V'' + \frac{2}{r} V' = \frac{1}{r} \left(rV'' + 2V'\right) = \frac{1}{r} \left(rV\right)''$$

So the Poisson equation can also be written as:

$$(rV)'' = -4\pi r\, n \tag{3.38.3.2}$$

Now we determine the values of $V(0)$, $V'(0)$ and the behavior of $V(\infty)$ and $V'(\infty)$. The equation determines V up to an arbitrary constant, so we set $V(\infty) = 0$ and now the potential V is determined uniquely.

The 3D integral of the (number) density is equal to the total (numeric) charge, which is equal to Z (number of electrons). We can then use the Poisson equation to rewrite the integral in terms of V:

$$\begin{aligned} Z = \int n(\mathbf{x})\mathrm{d}^3x = \int n(r)r^2\mathrm{d}\Omega\mathrm{d}r = \int_0^\infty 4\pi n(r)r^2\mathrm{d}r = \\ = -\int_0^\infty (rV)''r\mathrm{d}r = \\ = \int_0^\infty (rV)'\mathrm{d}r - [(rV)'r]_0^\infty = \\ = [rV]_0^\infty - [(rV)'r]_0^\infty = \\ = [rV - (rV)'r]_0^\infty = \\ = -[V'r^2]_0^\infty = \\ = \lim_{r\to 0} V'(r)r^2 - \lim_{r\to\infty} V'(r)r^2 \end{aligned}$$

Let

$$\lim_{r\to 0} V'(r)r^2 = C$$

Then around $r \to 0$ we get $V'(r) = \frac{C}{r^2}$ and $V(r) = -\frac{C}{r} + D$ (for some constant D). As such, C is a point charge (delta function) at the origin. From now on, we will assume no point charge, i.e. $C = 0$.

In the limit $r \to \infty$, we get the equation:

$$V'(r) = -\frac{Z - C}{r^2} = -\frac{Z}{r^2}$$

by integrating (and requiring that V vanished in infinity to get rid of the integration constant), we get for $r \to \infty$:

$$V(r) = \frac{Z}{r}$$

Integrating (3.38.3.2) directly, we get:

$$\begin{aligned} {}_0^\infty = -4\pi \int_0^\infty rn(r)\mathrm{d}r \\ [V + rV']_0^\infty = -4\pi \int_0^\infty rn(r)\mathrm{d}r \end{aligned}$$

We already know that V' behaves like $-\frac{Z}{r^2}$ in infinity, so rV' vanishes. Requiring V itself to vanish in infinity, the left hand side simplifies to $-V(0)$ and we get:

$$V(0) = 4\pi \int_0^\infty rn(r)\mathrm{d}r$$

Last thing to determine is $V'(0)$. To do that, we expand the charge density and potential (and it's derivatives) into a

series around the origin:

$$n(r) = n_0 + n_1 r + n_2 r^2 + \cdots = \sum_{k=0}^{\infty} n_k r^k$$
$$V(r) = V_0 + V_1 r + V_2 r^2 + \cdots = \sum_{k=0}^{\infty} V_k r^k$$
$$V'(r) = \sum_{k=1}^{\infty} V_k k r^{k-1}$$
$$V''(r) = \sum_{k=2}^{\infty} V_k k(k-1) r^{k-2}$$

And substitute into the equation (3.38.3.1):

$$\sum_{k=2}^{\infty} V_k k(k-1) r^{k-2} + \frac{2}{r} \sum_{k=1}^{\infty} V_k k r^{k-1} = -4\pi \sum_{k=0}^{\infty} n_k r^k$$
$$\sum_{k=2}^{\infty} V_k k(k-1) r^{k-2} + \frac{2}{r} V_1 + \frac{2}{r} \sum_{k=2}^{\infty} V_k k r^{k-1} = -4\pi \sum_{k=0}^{\infty} n_k r^k$$
$$\sum_{k=2}^{\infty} V_k k(k-1) r^{k-2} + \frac{2}{r} V_1 + \sum_{k=2}^{\infty} 2 V_k k r^{k-2} = -4\pi \sum_{k=0}^{\infty} n_k r^k$$
$$\frac{2}{r} V_1 + \sum_{k=2}^{\infty} V_k k \left((k-1) + 2\right) r^{k-2} = -4\pi \sum_{k=0}^{\infty} n_k r^k$$
$$\frac{2}{r} V_1 + \sum_{k=2}^{\infty} V_k k(k+1) r^{k-2} = -4\pi \sum_{k=0}^{\infty} n_k r^k$$
$$\frac{2}{r} V_1 + \sum_{l=0}^{\infty} V_{l+2} (l+2)(l+3) r^l = -4\pi \sum_{k=0}^{\infty} n_k r^k$$
$$\frac{2}{r} V_1 = -\sum_{k=0}^{\infty} \left(4\pi n_k + V_{k+2}(k+2)(k+3)\right) r^k$$

We now multiply the whole equation by r and then set $r = 0$. We get $V_1 = 0$, so $V'(0) = V_1 = 0$. We put it back into the equation to get:

$$\sum_{k=0}^{\infty} \left(4\pi n_k + V_{k+2}(k+2)(k+3)\right) r^k = 0$$

This must hold for all r, so we get the following set of equations for $k = 0, 1, \cdots$:

$$4\pi n_k + V_{k+2}(k+2)(k+3) = 0$$

from which we express V_k for all $k \geq 2$. We already know the values for $k = 1$ and $k = 0$ from earlier, so overall we get:

$$V_0 = 4\pi \int_0^{\infty} r n(r) \mathrm{d}r$$
$$V_1 = 0$$
$$V_{k+2} = -\frac{4\pi n_k}{(k+2)(k+3)}$$

in particular:

$$V_2 = -\frac{4\pi n_0}{6} = -\frac{2\pi}{3} n_0$$
$$V_3 = -\frac{4\pi n_1}{12} = -\frac{\pi}{3} n_1$$
$$V_4 = -\frac{4\pi n_2}{20} = -\frac{\pi}{5} n_2$$
$$V_5 = -\frac{4\pi n_3}{30} = -\frac{2\pi}{15} n_3$$
$$\dots$$

So we get the following series expansion for V and V':

$$V = V_0 - \frac{2\pi}{3} n_0 r^2 - \frac{\pi}{3} n_1 r^3 - \frac{\pi}{5} n_2 r^4 - \frac{2\pi}{15} n_3 r^5 - \dots$$
$$V' = -\frac{4\pi}{3} n_0 r - \pi n_1 r^2 - \frac{4\pi}{5} n_2 r^3 - \frac{2\pi}{3} n_3 r^4 - \dots$$

Examples

It is useful to have analytic solutions to test the numerical solvers. Here we present a few.

Gaussian Charge

The Gaussian charge is simply a Gaussian, normalized in such a way that the total charge is Z:

$$n(r) = \frac{Z\alpha^3}{\pi^{\frac{3}{2}}} e^{-\alpha^2 r^2} \tag{3.38.3.3}$$

Let us verify the normalization by calculating the total charge Q:

$$Q = \int n(\mathbf{x}) \mathrm{d}^3 x = 4\pi \int_0^\infty n(r) r^2 \mathrm{d}r =$$
$$= 4\pi \int_0^\infty \frac{Z\alpha^3}{\pi^{\frac{3}{2}}} e^{-\alpha^2 r^2} r^2 \mathrm{d}r =$$
$$= \frac{4Z\alpha^3}{\sqrt{\pi}} \int_0^\infty e^{-\alpha^2 r^2} r^2 \mathrm{d}r =$$
$$= \frac{4Z\alpha^3}{\sqrt{\pi}} \frac{\sqrt{\pi}}{4\alpha^3} = Z$$

So the total charge is $Q = Z$, as expected. Code:

```
>>> from sympy import var, integrate, exp, Symbol, oo
>>> var("r")
r
>>> alpha=Symbol("alpha", positive=True)
>>> integrate(exp(-alpha**2*r**2)*r**2, (r, 0, oo))
sqrt(pi)/(4*alpha**3)
```

Now we calculate the potential $V(r)$ from the Poisson equation (3.38.3.2):

$$\begin{aligned}(rV(r))'' = -4\pi r n(r) = -\frac{4\alpha^3 r Z}{\sqrt{\pi}} e^{-\alpha^2 r^2} \\ (rV(r))' = \frac{2Z\alpha}{\sqrt{\pi}} e^{-\alpha^2 r^2} + A \\ rV(r) = Z\,\mathrm{erf}(\alpha r) + Ar + B \\ V(r) = Z\,\frac{\mathrm{erf}(\alpha r)}{r} + A + \frac{B}{r}\end{aligned}$$

We have two integration constants A and B. We fix the potential using the condition $V(\infty) = 0$, which implies $A = 0$. The other constant B is a point charge at the origin, which in our case (3.38.3.3) is zero, so $B = 0$.

We finally obtain the potential:

$$V(r) = Z\,\frac{\mathrm{erf}(\alpha r)}{r}$$

We can calculate the electrostatic self-energy, i.e. the energy of interaction of the charge $n(r)$ with the potential generated by this charge $V(r)$:

$$\begin{aligned}E_{\text{self}} = \tfrac{1}{2}\int n(\mathbf{x})V(\mathbf{x})\mathrm{d}^3x = \frac{4\pi}{2}\int_0^\infty n(r)V(r)r^2\mathrm{d}r = \\ = 2\pi\int_0^\infty \frac{Z\alpha^3}{\pi^{\frac{3}{2}}} e^{-\alpha^2 r^2} Z\,\frac{\mathrm{erf}(\alpha r)}{r} r^2\mathrm{d}r = \\ = \frac{2Z^2\alpha^3}{\sqrt{\pi}}\int_0^\infty e^{-\alpha^2 r^2}\mathrm{erf}(\alpha r) r\mathrm{d}r = \\ = \frac{2Z^2\alpha^3}{\sqrt{\pi}}\,\frac{\sqrt{2}}{4\alpha^2} = \\ = \frac{Z^2\alpha}{\sqrt{2\pi}}\end{aligned}$$

Code:

```
>>> from sympy import var, integrate, exp, Symbol, oo, erf
>>> var("r")
r
>>> alpha=Symbol("alpha", positive=True)
>>> integrate(exp(-alpha**2*r**2)*erf(alpha*r)*r, (r, 0, oo))
sqrt(2)/(4*alpha**2)
```

Exponential Charge

The exponential charge is simply an exponential, normalized in such a way that the total charge is Z:

$$n(r) = \frac{Z\alpha^3}{8\pi} e^{-\alpha r} \tag{3.38.3.4}$$

Let us verify the normalization by calculating the total charge Q:

$$\begin{aligned} Q = \int n(\mathbf{x}) \mathrm{d}^3 x = 4\pi \int_0^\infty n(r) r^2 \mathrm{d}r = \\ = 4\pi \int_0^\infty \frac{Z\alpha^3}{8\pi} e^{-\alpha r} r^2 \mathrm{d}r = \\ = \frac{Z\alpha^3}{2} \int_0^\infty e^{-\alpha r} r^2 \mathrm{d}r = \\ = \frac{Z\alpha^3}{2} \frac{2}{\alpha^3} = Z \end{aligned}$$

So the total charge is $Q = Z$, as expected.

Now we calculate the potential $V(r)$ from the Poisson equation (3.38.3.2):

$$\begin{aligned} (rV(r))'' = -4\pi r n(r) = -\frac{Z\alpha^3}{2} r e^{-\alpha r} \\ V(r) = -Z \left(\frac{\alpha}{2} + \frac{1}{r} \right) e^{-\alpha r} + A + \frac{B}{r} \end{aligned}$$

Similarly as for the Gaussian charge, we require the potential $V(r)$ to vanish at infinity, which implies $A = 0$. Then we calculate the point charge at the origin:

$$\begin{aligned} C = \lim_{r \to 0} V'(r) r^2 = \\ = \lim_{r \to 0} \tfrac{1}{2} \left(-2Be^{\alpha r} + Z\alpha r \left(\alpha r + 1 \right) + Z \left(\alpha r + 2 \right) \right) e^{-\alpha r} = \\ = Z - B \end{aligned}$$

We do not have any point charge at the origin, so $C = Z - B = 0$, from which it follows $B = Z$. We finally obtain:

$$V(r) = -Z \left(\frac{\alpha}{2} + \frac{1}{r} \right) e^{-\alpha r} + \frac{Z}{r} = Z \left(\frac{1 - e^{-\alpha r}}{r} - \frac{\alpha e^{-\alpha r}}{2} \right)$$

Let us calculate the self-energy:

$$\begin{aligned} E_{\text{self}} = \tfrac{1}{2} \int n(\mathbf{x}) V(\mathbf{x}) \mathrm{d}^3 x = \frac{4\pi}{2} \int_0^\infty n(r) V(r) r^2 \mathrm{d}r = \\ = 2\pi \int_0^\infty \frac{Z\alpha^3}{8\pi} e^{-\alpha r} Z \left(\frac{1 - e^{-\alpha r}}{r} - \frac{\alpha e^{-\alpha r}}{2} \right) r^2 \mathrm{d}r = \\ = \frac{Z^2\alpha^3}{4} \int_0^\infty e^{-\alpha r} \left(\frac{1 - e^{-\alpha r}}{r} - \frac{\alpha e^{-\alpha r}}{2} \right) r^2 \mathrm{d}r = \\ = \frac{Z^2\alpha^3}{4} \left(\frac{5}{8\alpha^2} \right) = \\ = \frac{5Z^2\alpha}{32} \end{aligned}$$

Code:

```
>>> from sympy import var, integrate, exp, Symbol, oo
>>> var("r Z B")
(r, Z, B)
>>> alpha=Symbol("alpha", positive=True)
>>> integrate(exp(-alpha*r)*r**2, (r, 0, oo))
2/alpha**3
>>> V = integrate(-Z*alpha**3/2 * r * exp(-alpha*r), r, r)/r
```

(continues on next page)

(continued from previous page)

```
>>> V.simplify()
-Z*(alpha*r + 2)*exp(-alpha*r)/(2*r)
>>> ((V+B/r).diff(r)*r**2).simplify()
(-2*B*exp(alpha*r) + Z*alpha*r*(alpha*r + 1) + Z*(alpha*r +
2))*exp(-alpha*r)/2
>>> ((V+B/r).diff(r)*r**2).limit(r, 0)
-B + Z
>>> integrate(exp(-alpha*r)*((1-exp(-alpha*r))/r-alpha*exp(-alpha*r)/2)*r**2, (r, 0,␣
↪oo))
5/(8*alpha**2)
```

Piecewise Polynomial Charge

We will use a second-derivative continuous piecewise polynomial for $n(r)$, normalized in such a way that the total charge is Z:

$$n(r) = \begin{cases} -21Z(r-r_c)^3(6r^2+3rr_c+r_c^2)/(5\pi r_c^8) & \text{for } 0 \le r \le r_c \\ 0 & \text{for } r > r_c \end{cases} \tag{3.38.3.5}$$

Let us verify the normalization by calculating the total charge Q:

$$\begin{aligned} Q = \int n(\mathbf{x})\mathrm{d}^3x = 4\pi \int_0^\infty n(r)r^2\mathrm{d}r = \\ = 4\pi \int_0^{r_c} -21Z(r-r_c)^3(6r^2+3rr_c+r_c^2)/(5\pi r_c^8)r^2\mathrm{d}r = \\ = Z \end{aligned}$$

So the total charge is $Q = Z$, as expected.

Now we calculate the potential $V(r)$ from the Poisson equation (3.38.3.2):

$$(rV(r))'' = -4\pi rn(r) = 4\pi r \cdot 21Z(r-r_c)^3(6r^2+3rr_c+r_c^2)/(5\pi r_c^8)$$

$$V(r) = \begin{cases} \frac{Zr^2}{5r_c^8}\left(9r^5 - 30r^4r_c + 28r^3r_c^2 - 14r_c^5\right) + A_1 + \frac{B_1}{r} & \text{for } 0 \le r \le r_c \\ A_2 + \frac{B_2}{r} & \text{for } r > r_c \end{cases}$$

Similarly as for the Gaussian charge, we require the potential $V(r)$ to vanish at infinity, which implies $A_2 = 0$. Then we calculate the point charge at the origin:

$$\begin{aligned} C = \lim_{r\to 0} V'(r)r^2 = \\ = \lim_{r\to 0}\left(-B_1 + \frac{63Zr^8}{5r_c^8} - \frac{36Z}{r_c^7}r^7 + \frac{28Z}{r_c^6}r^6 - \frac{28Zr^3}{5r_c^3}\right) \\ = -B_1 \end{aligned}$$

We do not have any point charge at the origin, so $C = -B_1 = 0$, from which it follows $B_1 = 0$. Then B_2 is calculated from the condition of a continuous first derivative at $r = r_c$:

$$V'(r_c) = \begin{cases} -\frac{Z}{r_c^2} & \text{for } 0 \le r \le r_c \\ -\frac{B_2}{r_c^2} & \text{for } r > r_c \end{cases}$$

So $B_2 = Z$. Finally, A_1 is calculated from the continuous values of $V(r_c)$:

$$V(r_c) = \begin{cases} A_1 - \frac{7Z}{5r_c} & \text{for } 0 \le r \le r_c \\ \frac{Z}{r_c} & \text{for } r > r_c \end{cases}$$

which implies $A_1 = \frac{12Z}{5r_c}$. We finally obtain:

$$V(r) = \begin{cases} \frac{Z}{5r_c^8}\left(9r^7 - 30r^6 r_c + 28r^5 r_c^2 - 14r^2 r_c^5 + 12r_c^7\right) & \text{for } 0 \le r \le r_c \\ \frac{Z}{r} & \text{for } r > r_c \end{cases}$$

Let us calculate the self-energy:

$$E_{\text{self}} = \frac{1}{2}\int n(\mathbf{x})V(\mathbf{x})\mathrm{d}^3x = \frac{4\pi}{2}\int_0^\infty n(r)V(r)r^2\mathrm{d}r =$$

$$= 2\pi\int_0^{r_c} -21Z(r-r_c)^3(6r^2+3rr_c+r_c^2)/(5\pi r_c^8)\frac{Z}{5r_c^8}\left(9r^7 - 30r^6 r_c + 28r^5 r_c^2 - 14r^2 r_c^5 + 12r_c^7\right)r^2\mathrm{d}r =$$

$$= \frac{15962Z^2}{17875r_c}$$

Let us also calculate the following integral:

$$I_g = \int n(\mathbf{x})\left(\frac{Z}{r} - V(\mathbf{x})\right)\mathrm{d}^3x = 4\pi\int_0^\infty n(r)\left(\frac{Z}{r} - V(r)\right)r^2\mathrm{d}r =$$

$$= \frac{10976Z^2}{17875r_c}$$

Which agrees with *[Pask2012]*, equation (10c). The following integral over the sphere of radius r_c:

$$I_{sph} = \int_{\Omega:r<r_c}\left(\frac{Z}{r} - V(\mathbf{x})\right)\mathrm{d}^3x = 4\pi\int_0^{r_c}\left(\frac{Z}{r} - V(r)\right)r^2\mathrm{d}r =$$

$$= \frac{14\pi Z r_c^2}{75}$$

Again in agreement with *[Pask2012]*, the paragraph after equation (17).

Code:

```
>>> from sympy import var, pi, integrate, solve
>>> var("r r_c Z A B")
(r, r_c, Z, A, B)
>>> n = -21*Z*(r-r_c)**3*(6*r**2+3*r*r_c+r_c**2)/(5*pi* r_c**8)
>>> 4*pi*integrate(n*r**2, (r, 0, r_c))
Z
>>> V = integrate(-4*pi*r*n, r, r)/r
>>> V.simplify()
Z*r**2*(9*r**5 - 30*r**4*r_c + 28*r**3*r_c**2 - 14*r_c**5)/(5*r_c**8)
>>> ((V+A+B/r).diff(r)*r**2).simplify()
-B + 63*Z*r**8/(5*r_c**8) - 36*Z*r**7/r_c**7 + 28*Z*r**6/r_c**6 - 28*Z*r**3/(5*r_c**3)
>>> (V+A).diff(r).subs(r, r_c)
-Z/r_c**2
>>> (V+A).subs(r, r_c)
A - 7*Z/(5*r_c)
>>> A = solve((V+A).subs(r, r_c)-Z/r_c, A)[0]
>>> A
12*Z/(5*r_c)
>>> V = V + A
>>> V.simplify()
Z*(r**2*(9*r**5 - 30*r**4*r_c + 28*r**3*r_c**2 - 14*r_c**5) + 12*r_c**7)/(5*r_c**8)
>>> 2*pi*integrate(n*V*r**2, (r, 0, r_c))
15962*Z**2/(17875*r_c)
>>> 4*pi*integrate(n*(Z/r-V)*r**2, (r, 0, r_c))
10976*Z**2/(17875*r_c)
>>> 4*pi*integrate((Z/r-V)*r**2, (r, 0, r_c))
14*pi*Z*r_c**2/75
```

Alternatively, one can also calculate this using a `Piecewise` function:

```
>>> from sympy import var, pi, integrate, solve, Piecewise, oo, Symbol
>>> var("r Z A B")
(r, Z, A, B)
>>> r_c = Symbol("r_c", positive=True)
>>> n = Piecewise((-21*Z*(r - r_c)**3*(6*r**2 + 3*r*r_c + r_c**2)/(5*pi*r_c**8), r <=
↪r_c), (0, True))
>>> 4*pi*integrate(n*r**2, (r, 0, oo))
Z
>>> V = integrate(-4*pi*r*n, r, r)/r
>>> V.simplify()
Piecewise((Z*r**2*(9*r**5 - 30*r**4*r_c + 28*r**3*r_c**2 - 14*r_c**5)/(5*r_c**8), r
↪<= r_c), (0, True))
>>> ((V+A+B/r).diff(r)*r**2).simplify()
Piecewise((-B + 63*Z*r**8/(5*r_c**8) - 36*Z*r**7/r_c**7 + 28*Z*r**6/r_c**6 -
↪28*Z*r**3/(5*r_c**3), r <= r_c), (-B, True))
>>> (V+A).diff(r).subs(r, r_c)
-Z/r_c**2
>>> (V+A).subs(r, r_c)
A - 7*Z/(5*r_c)
>>> A = solve((V+A).subs(r, r_c)-Z/r_c, A)[0]
>>> A
12*Z/(5*r_c)
>>> V = V + Piecewise((A, r <= r_c), (0, True))
>>> V.simplify()
Piecewise((Z*(r**2*(9*r**5 - 30*r**4*r_c + 28*r**3*r_c**2 - 14*r_c**5)/r_c**7 + 12)/
↪(5*r_c), r <= r_c), (0, True))
>>> 2*pi*integrate(n*V*r**2, (r, 0, oo))
15962*Z**2/(17875*r_c)
>>> 4*pi*integrate(n*(Z/r-V)*r**2, (r, 0, oo))
10976*Z**2/(17875*r_c)
```

3.39 Linear Algebra

3.39.1 Scalar Product

Virtually all spaces used in physics are Hilbert spaces (treated in the weak sense, i.e. equipped with distributions), which means that they have a scalar product and a norm.

The braket $\langle f|g\rangle$ in Dirac notation is a scalar product and we are free to define it anyway we like, as long as it satisfies the following properties:

$$\begin{aligned}\langle f|g\rangle &= \langle g|f\rangle^* \\ \langle f|ag\rangle &= a\langle f|g\rangle \\ \langle f_1+f_2|g\rangle &= \langle f_1|g\rangle + \langle f_2|g\rangle \\ \langle f|f\rangle &\geq 0\end{aligned}$$

where $\langle f|f\rangle = 0$ if, and only if, $|f\rangle = 0$. Scalar product induces the norm:

$$||\,|f\rangle\,|| = \sqrt{\langle f|f\rangle}$$

Any norm has to satisfy the following properties:

$$\begin{aligned} ||a\,|f\rangle\,|| &= |a|\,||\,|f\rangle\,|| \\ ||\,|f\rangle + |g\rangle\,|| &\leq ||\,|f\rangle\,|| + ||\,|g\rangle\,|| \\ ||\,|f\rangle\,|| &\geq 0 \end{aligned}$$

where $||\,|f\rangle\,|| = 0$ if, and only if, $|f\rangle = 0$. Those properties are automatically satisfied by the induced norm.

In general, any of these properties can be weakened, one can study spaces that have a norm, but not a scalar product, or spaces, that have objects resembling a norm (or a scalar product), that only satisfy some of the properties of the norm (or a scalar product). Those are not very important in physics. On the other hand, it is very important to understand how to work with Hilbert spaces (in the weak sense). Dirac notation makes it very easy to understand and remember how to work with such spaces.

Examples

Some examples of frequently used spaces and scalar products follows.

Finite dimensional spaces:

$$\begin{aligned} |f\rangle &= \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix} \\ \langle f| &= \begin{pmatrix} f_1 & f_2 & \cdots & f_n \end{pmatrix} \end{aligned}$$

R^n Euclidean scalar product:

$$\langle f|g\rangle = \begin{pmatrix} f_1 & f_2 & \cdots & f_n \end{pmatrix} \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_n \end{pmatrix} = f_1 g_2 + f_2 g_2 + \cdots + f_n g_n$$

Infinite dimensional spaces:

$$\begin{aligned} |f\rangle &= \begin{pmatrix} f_1 \\ f_2 \\ \vdots \end{pmatrix} \\ \langle f| &= \begin{pmatrix} f_1 & f_2 & \cdots \end{pmatrix} \end{aligned}$$

l^2 scalar product:

$$\langle f|g\rangle = \begin{pmatrix} f_1 & f_2 & \cdots \end{pmatrix} \begin{pmatrix} g_1 \\ g_2 \\ \vdots \end{pmatrix} = f_1 g_2 + f_2 g_2 + \cdots$$

Function spaces:

$$\begin{aligned} |f\rangle &= f(x) \\ \langle f| &= f^*(x) \end{aligned}$$

L^2 scalar product:

$$\langle f|g\rangle = \int_\Omega f^*(x) g(x) \mathrm{d}x$$

H^1 scalar product:

$$\langle f|g\rangle = \int_\Omega f^*(x)g(x) + f'^*(x)g'(x)\mathrm{d}x$$

H^2 scalar product:

$$\langle f|g\rangle = \int_\Omega f^*(x)g(x) + f'^*(x)g'(x) + f''^*(x)g''(x)\mathrm{d}x$$

Energy scalar product:

$$\langle f|g\rangle = \int_\Omega f^*(x)q(x)g(x) + f'^*(x)p(x)g'(x)\mathrm{d}x$$

All of these scalar products automatically satisfy all of the properties of the scalar product, only the energy scalar product doesn't automatically satisfy $\langle f|f\rangle \geq 0$, which imposes some conditions on the parameters $p(x)$ and $q(x)$.

3.39.2 Projections

Projection is a linear idempotent operator P:

$$P^2 = P$$

It takes a vector $|u\rangle$ from V and projects it onto a vector $|w\rangle = P|u\rangle$ from W. Further application of the operator P gains nothing: $P|w\rangle = P^2|u\rangle = P|u\rangle = |w\rangle$. It decomposes the space V into a direct sum $V = W \oplus W^\perp$ of the projection subspace W and its complement $W^\perp$. If $|w\rangle$ is from W then its complement $|u\rangle - P|u\rangle$ is from $W^\perp$.

Orthogonal projection is a projection that is Hermitean:

$$P^\dagger = P$$

The complement of an orthogonal projection is orthogonal to any vector from W:

$$\begin{aligned}\langle u - Pu|w\rangle = \langle u|w\rangle - \langle Pu|w\rangle = \langle u|w\rangle - \langle u|P^\dagger|w\rangle = \\ = \langle u|w\rangle - \langle u|P|w\rangle = \langle u|w\rangle - \langle u|w\rangle = 0\end{aligned}$$

In other words, orthogonal projection projects a vector $|u\rangle$ from the space V into an orthogonal subspace (projection subspace) W. The two spaces W and $W^\perp$ are orthogonal, because any vector from W is orthogonal to all vectors from $W^\perp$. Given the space W, the operator P is unique.

The complement of non-orthogonal projection is not orthogonal to any vector from W:

$$\langle u - Pu|w\rangle = \langle u|w\rangle - \langle u|P|w\rangle \neq \langle u|w\rangle - \langle u|P^\dagger|w\rangle = \langle u|w\rangle - \langle u|w\rangle = 0$$

And the two spaces W and $W^\perp$ are not orthogonal, because any vector from W is not orthogonal to any vector from $W^\perp$. Given both spaces W and $W^\perp$, the operator P is unique.

If we choose any orthonormal basis $|w_0\rangle$, $|w_1\rangle$, $|w_2\rangle$, ..., of the subspace W, then the orthogonal projection P is:

$$P = \sum_{k=0}^{\infty} |w_k\rangle \langle w_k| \tag{3.39.2.1}$$

because:

$$\begin{aligned}P^2 = \sum_{k=0}^{\infty} |w_k\rangle \langle w_k| \sum_{l=0}^{\infty} |w_l\rangle \langle w_l| = \sum_{k,l=0}^{\infty} |w_k\rangle \langle w_k|w_l\rangle \langle w_l| = \\ = \sum_{k,l=0}^{\infty} |w_k\rangle \delta_{kl} \langle w_l| = \sum_{k=0}^{\infty} |w_k\rangle \langle w_k| = P\end{aligned}$$

and

$$P^\dagger = \left(\sum_{k=0}^{\infty} |w_k\rangle \langle w_k|\right)^\dagger = \sum_{k=0}^{\infty} (|w_k\rangle \langle w_k|)^\dagger = \sum_{k=0}^{\infty} |w_k\rangle \langle w_k| = P$$

P is independent of the basis, i.e $\sum_{k=0}^{\infty} |w_k\rangle \langle w_k| = \sum_{l=0}^{\infty} |u_l\rangle \langle u_l|$, as long as $|u_l\rangle$ span the same subspace as $|w_k\rangle$, because the operator P is unique.

To find the closest vector $|w\rangle$ from W to the vector $|u\rangle$ from V, we need to minimize the norm $|| |u\rangle - |w\rangle ||$. So we write $|w\rangle = P|u\rangle + |z\rangle$ for some vector $|z\rangle$ from W and simplify the norm:

$$\begin{aligned} || |u\rangle - |w\rangle ||^2 = \langle u - w | u - w \rangle = \langle u - Pu - z | u - Pu - z \rangle = \\ = \langle u - Pu | u - Pu \rangle + \langle z | z \rangle - \langle u - Pu | z \rangle - \langle z | u - Pu \rangle = \\ = \langle u - Pu | u - Pu \rangle + \langle z | z \rangle \end{aligned}$$

which is minimal for $|z\rangle = 0$, so we found out that the closest vector is $|w\rangle = P|u\rangle$. We used the fact that $\langle u - Pu | z \rangle = 0$, because $|u - Pu\rangle$ is from the orthogonal complement to the subspace W. In other words, orthogonal projection finds the closest vector from a subspace onto which it projects.

Projection Coefficients

Given the basis $|v_k\rangle$ (orthogonal or non-orthogonal), we would like to find a formula for the projection coefficients ϕ_k defined by:

$$P|u\rangle = \sum_{k=0}^{\infty} |v_k\rangle \phi_k \,. \tag{3.39.2.2}$$

This holds, because $P|u\rangle$ belongs to the space W and every vector from it can be expressed as a linear combination of $|v_k\rangle$.

Projecting to Orthogonal Basis

For orthogonal projections we simply substitute the equation (3.39.2.1) into (3.39.2.2) and get:

$$P|u\rangle = \sum_{k=0}^{\infty} |w_k\rangle \langle w_k | u \rangle = \sum_{k=0}^{\infty} |w_k\rangle \phi_k \,,$$

from which the projection coefficients ϕ_k are given by

$$\phi_k = \langle w_k | u \rangle \,. \tag{3.39.2.3}$$

Projecting to Nonorthogonal Basis

In order to project onto a nonorthogonal basis $|v_k\rangle$ (for example a finite element basis), we multiply (3.39.2.2) by $\langle v_l|$ from the left and simplify:

$$\begin{aligned} \langle v_l | P | u \rangle &= \sum_{k=0}^{\infty} \langle v_l | v_k \rangle \phi_k \\ \langle v_l | P^\dagger | u \rangle &= \sum_{k=0}^{\infty} \langle v_l | v_k \rangle \phi_k \\ \langle v_l | u \rangle &= \sum_{k=0}^{\infty} \langle v_l | v_k \rangle \phi_k \end{aligned}$$

so we need to solve a linear system for the coefficients ϕ_k:

$$A_{lk}\phi_k = f_l \,, \tag{3.39.2.4}$$

where

$$\begin{aligned} A_{lk} &= \langle v_l | v_k \rangle \\ f_l &= \langle v_l | u \rangle \ . \end{aligned}$$

This works for any basis, it doesn't have to be normalized nor orthogonal. In the special case of a (normalized) orthogonal basis, we get $A_{lk} = \langle v_l | v_k \rangle = \delta_{lk}$ and from (3.39.2.4) we get

$$A_{lk}\phi_k = \delta_{lk}\phi_k = \phi_l = f_l = \langle v_l | u \rangle \ ,$$

so we recovered the equation (3.39.2.3) as expected.

Examples

R^3 orthogonal projection. Orthogonal basis:

$$\begin{aligned} |w_0\rangle &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\ |w_1\rangle &= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\ P = |w_0\rangle\langle w_0| + |w_1\rangle\langle w_1| &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Different orthogonal basis:

$$\begin{aligned} |w_0\rangle &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \\ |w_1\rangle &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \\ P = |w_0\rangle\langle w_0| + |w_1\rangle\langle w_1| &= \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

R^2 non-orthogonal (oblique) projection ($\alpha \neq 0$):

$$\begin{aligned} P &= \begin{pmatrix} 1 & \alpha \\ 0 & 0 \end{pmatrix} \\ P^2 = \begin{pmatrix} 1 & \alpha \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & \alpha \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} 1 & \alpha \\ 0 & 0 \end{pmatrix} = P \\ P^\dagger = P^T &= \begin{pmatrix} 1 & 0 \\ \alpha & 0 \end{pmatrix} \neq P \\ P\,|u\rangle = \begin{pmatrix} 1 & \alpha \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} x + \alpha y \\ 0 \end{pmatrix} \end{aligned}$$

Because the projection is not orthogonal (in the R^2 Euclidean scalar product), the projected point $(x+\alpha y, 0)$ is not the closest point (in the induced Euclidean R^2 norm) to (x, y). For $\alpha = 0$ the projection becomes orthogonal and indeed the projected point $(x, 0)$ then becomes the closest point to (x, y).

Lagrange interpolation projection onto the space $\{1, x\}$:

$$\begin{aligned} |u\rangle &= f(x) \\ P|u\rangle &= \frac{f(1)+f(-1)}{2} + x\frac{f(1)-f(-1)}{2} \end{aligned}$$

L^2 projection onto the space $\{1, x\}$. Orthogonal basis:

$$\begin{aligned} |u\rangle &= f(x) \\ |w_0\rangle &= \frac{1}{\sqrt{2}} \\ |w_1\rangle &= \sqrt{\frac{3}{2}}x \\ P|u\rangle &= |w_0\rangle\langle w_0|u\rangle + |w_1\rangle\langle w_1|u\rangle = \\ &= \frac{1}{\sqrt{2}}\int_{-1}^{1}\frac{1}{\sqrt{2}}f(x)\mathrm{d}x + \sqrt{\frac{3}{2}}x\int_{-1}^{1}\sqrt{\frac{3}{2}}xf(x)\mathrm{d}x = \frac{1}{2}\int_{-1}^{1}f(x)\mathrm{d}x + \frac{3}{2}x\int_{-1}^{1}xf(x)\mathrm{d}x \end{aligned}$$

Different orthogonal basis:

$$\begin{aligned} |w_0\rangle &= \frac{\sqrt{6}}{4}(1+x) \\ |w_1\rangle &= \frac{\sqrt{2}}{4}(1-3x) \\ P|u\rangle &= |w_0\rangle\langle w_0|u\rangle + |w_1\rangle\langle w_1|u\rangle = \\ &= \frac{\sqrt{6}}{4}(1+x)\int_{-1}^{1}\frac{\sqrt{6}}{4}(1+x)f(x)\mathrm{d}x + \frac{\sqrt{2}}{4}(1-3x)\int_{-1}^{1}\frac{\sqrt{2}}{4}(1-3x)f(x)\mathrm{d}x = \frac{1}{2}\int_{-1}^{1}f(x)\mathrm{d}x + \frac{3}{2}x\int_{-1}^{1}xf(x)\mathrm{d}x \end{aligned}$$

Nonorthogonal basis:

$$|w_0\rangle = 1$$
$$|w_1\rangle = 1 + x$$
$$A_{lk}\phi_k = f_l$$
$$A_{lk} = \begin{pmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{pmatrix} = \begin{pmatrix} \langle w_0|w_0\rangle & \langle w_0|w_1\rangle \\ \langle w_1|w_0\rangle & \langle w_1|w_1\rangle \end{pmatrix} =$$
$$= \begin{pmatrix} \int_{-1}^{1} \mathrm{d}x & \int_{-1}^{1} 1 + x\,\mathrm{d}x \\ \int_{-1}^{1} 1 + x\,\mathrm{d}x & \int_{-1}^{1} (1+x)^2\,\mathrm{d}x \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 2 & \frac{8}{3} \end{pmatrix}$$
$$A^{-1}_{kl} = \begin{pmatrix} 2 & -\frac{3}{2} \\ -\frac{3}{2} & \frac{3}{2} \end{pmatrix}$$
$$f_l = \begin{pmatrix} f_0 \\ f_1 \end{pmatrix} = \begin{pmatrix} \langle w_0|u\rangle \\ \langle w_1|u\rangle \end{pmatrix} = \begin{pmatrix} \int_{-1}^{1} f(x)\,\mathrm{d}x \\ \int_{-1}^{1} (1+x) f(x)\,\mathrm{d}x \end{pmatrix}$$
$$\phi_k = \begin{pmatrix} \phi_0 \\ \phi_1 \end{pmatrix} = A^{-1}_{kl} f_l = \begin{pmatrix} 2 & -\frac{3}{2} \\ -\frac{3}{2} & \frac{3}{2} \end{pmatrix} \begin{pmatrix} \int_{-1}^{1} f(x)\,\mathrm{d}x \\ \int_{-1}^{1} (1+x) f(x)\,\mathrm{d}x \end{pmatrix} =$$
$$= \begin{pmatrix} 2\int_{-1}^{1} f(x) - \frac{3}{2}(1+x)f(x)\mathrm{d}x \\ -\frac{3}{2}\int_{-1}^{1} f(x) + \frac{3}{2}(1+x)f(x)\mathrm{d}x \end{pmatrix} =$$
$$P\,|u\rangle = |w_0\rangle\,\phi_0 + |w_1\rangle\,\phi_1 =$$
$$= 1\left(2\int_{-1}^{1} f(x) - \frac{3}{2}(1+x)f(x)\mathrm{d}x\right) + (1+x)\left(-\frac{3}{2}\int_{-1}^{1} f(x) + \frac{3}{2}(1+x)f(x)\mathrm{d}x\right) =$$
$$= \frac{1}{2}\int_{-1}^{1} f(x)\mathrm{d}x + \frac{3}{2}x\int_{-1}^{1} x f(x)\mathrm{d}x$$

H^1 projection. Nonorthogonal basis:

$$|w_0\rangle = 1$$
$$|w_1\rangle = 1 + x$$
$$A_{lk}\phi_k = f_l$$
$$A_{lk} = \begin{pmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{pmatrix} = \begin{pmatrix} \langle w_0|w_0\rangle & \langle w_0|w_1\rangle \\ \langle w_1|w_0\rangle & \langle w_1|w_1\rangle \end{pmatrix} =$$
$$= \begin{pmatrix} \int_{-1}^{1} \mathrm{d}x & \int_{-1}^{1} 1+x\,\mathrm{d}x \\ \int_{-1}^{1} 1+x\,\mathrm{d}x & \int_{-1}^{1} (1+x)^2+1\,\mathrm{d}x \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 2 & \frac{14}{3} \end{pmatrix}$$
$$A^{-1}_{kl} = \begin{pmatrix} \frac{7}{8} & -\frac{3}{8} \\ -\frac{3}{8} & \frac{3}{8} \end{pmatrix}$$
$$f_l = \begin{pmatrix} f_0 \\ f_1 \end{pmatrix} = \begin{pmatrix} \langle w_0|u\rangle \\ \langle w_1|u\rangle \end{pmatrix} = \begin{pmatrix} \int_{-1}^{1} f(x)\,\mathrm{d}x \\ \int_{-1}^{1} (1+x)f(x) + f'(x)\,\mathrm{d}x \end{pmatrix}$$
$$\phi_k = \begin{pmatrix} \phi_0 \\ \phi_1 \end{pmatrix} = A^{-1}_{kl} f_l = \begin{pmatrix} \frac{7}{8} & -\frac{3}{8} \\ -\frac{3}{8} & \frac{3}{8} \end{pmatrix} \begin{pmatrix} \int_{-1}^{1} f(x)\,\mathrm{d}x \\ \int_{-1}^{1} (1+x)f(x) + f'(x)\,\mathrm{d}x \end{pmatrix} =$$
$$= \begin{pmatrix} \int_{-1}^{1} \frac{7}{8}f(x) - \frac{3}{8}(1+x)f(x) - \frac{3}{8}f'(x)\mathrm{d}x \\ \int_{-1}^{1} -\frac{3}{8}f(x) + \frac{3}{8}(1+x)f(x) + \frac{3}{8}f'(x)\mathrm{d}x \end{pmatrix} =$$
$$P\,|u\rangle = |w_0\rangle\,\phi_0 + |w_1\rangle\,\phi_1 =$$
$$= 1\left(\int_{-1}^{1} \frac{7}{8}f(x) - \frac{3}{8}(1+x)f(x) - \frac{3}{8}f'(x)\mathrm{d}x\right) + (1+x)\left(\int_{-1}^{1} -\frac{3}{8}f(x) + \frac{3}{8}(1+x)f(x) + \frac{3}{8}f'(x)\mathrm{d}x\right) =$$
$$= \int_{-1}^{1} \frac{1}{2}f(x) - \frac{3}{8}f'(x)\mathrm{d}x + x\int_{-1}^{1} \frac{3}{8}xf(x) + \frac{3}{8}f'(x)\mathrm{d}x =$$
$$= \int_{-1}^{1} \frac{1}{2}f(x)\mathrm{d}x + x\int_{-1}^{1} \frac{3}{8}xf(x)\mathrm{d}x - \frac{3}{8}(f(1) - f(-1)) + \frac{3}{8}x(f(1) - f(-1)) =$$
$$= \frac{1}{2}\int_{-1}^{1} f(x)\mathrm{d}x + \frac{3}{8}x\int_{-1}^{1} xf(x)\mathrm{d}x + \frac{3}{8}(-1+x)(f(1) - f(-1))$$

3.40 Differential Geometry

3.40.1 Manifolds

Scalars, Vectors, Tensors

Differentiable manifold U is a space covered by an atlas of maps, each map covers part of the manifold and is a one to one mapping to an euclidean space $\mathbf{R}^n$:

$$\phi : U \to \mathbf{R}^n$$

Let's have a one-to-one transformation between x^μ and x'^μ coordinates (we simply write $x \equiv x^\mu$, etc.):

$$x' = x'(x)$$

$$x = x(x')$$

Scalar $\phi(x)$ is such a field that transforms as ($\phi'(x')$ is it's value in x' coordinates):

$$\phi'(x') = \phi(x)$$

One form $p_\alpha(x)$ is such a field that transforms the same as the gradient $\frac{\partial\phi(x)}{\partial x^\mu}$ of a scalar, that transforms as ($\frac{\partial\phi'(x')}{\partial x'^\mu}$ is it's value in x' coordinates):

$$\frac{\partial\phi'(x')}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial\phi'(x')}{\partial x^\nu} = \frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial\phi(x)}{\partial x^\nu}$$

so

$$p'_\mu(x') = \frac{\partial x^\nu}{\partial x'^\mu}p_\nu(x)$$

Vector V^α is such a field that produces a scalar $\phi = V^\alpha p_\alpha$ when contracted with a one form and this fact is used to deduce how it transforms:

$$\phi' = V'^\alpha p'_\alpha = V'^\alpha\frac{\partial x^\beta}{\partial x'^\alpha}p_\beta = \phi = V^\beta p_\beta$$

so we have

$$V'^\alpha\frac{\partial x^\beta}{\partial x'^\alpha} = V^\beta$$

multiplying by $\frac{\partial x'^\mu}{\partial x^\beta}$ and using the fact that $\frac{\partial x^\beta}{\partial x'^\alpha}\frac{\partial x'^\mu}{\partial x^\beta} = \frac{\partial x'^\mu}{\partial x'^\alpha} = \delta^\mu_\alpha$ we get

$$V'^\mu = \frac{\partial x'^\mu}{\partial x^\beta}V^\beta$$

Higher tensors are build up and their transformation properties derived from the fact, that by contracting with either a vector or a form we get a lower rank tensor that we already know how it transforms.

Having now defined scalar, vector and tensor fields, one may then choose a basis at each point for each field, the only requirement being that the basis is not singular. For example for vectors, each point in U has a basis $\vec{e}_\alpha$, so a vector (field) $\vec{V}$ has components V^α with respect to this basis:

$$\vec{V} = V^\alpha\vec{e}_\alpha$$

Covariant differentiation

The derivative of the basis vector $\frac{\partial\vec{e}_\alpha}{\partial x^\beta}$ is a vector, thus it can be written as a linear combination of the basis vectors:

$$\frac{\partial\vec{e}_\alpha}{\partial x^\beta} = \Gamma^\mu_{\alpha\beta}\vec{e}_\mu$$

Differentiating a vector is then easy:

$$\frac{\partial\vec{V}}{\partial x^\beta} \equiv \nabla_\beta\vec{V} = \frac{\partial V^\alpha}{\partial x^\beta}\vec{e}_\alpha + V^\alpha\frac{\partial\vec{e}_\alpha}{\partial x^\beta} = \frac{\partial V^\alpha}{\partial x^\beta}\vec{e}_\alpha + V^\alpha\Gamma^\mu_{\alpha\beta}\vec{e}_\mu = \left(\frac{\partial V^\alpha}{\partial x^\beta} + \Gamma^\alpha_{\mu\beta}V^\mu\right)\vec{e}_\alpha$$

So we define a covariant derivative:

$$\nabla_\beta V^\alpha = \frac{\partial V^\alpha}{\partial x^\beta} + \Gamma^\alpha_{\mu\beta}V^\mu$$

and write

$$\frac{\partial\vec{V}}{\partial x^\beta} = \nabla_\beta\vec{V} = \left(\nabla_\beta\vec{V}\right)^\alpha\vec{e}_\alpha = \left(\nabla_\beta V^\alpha\right)\vec{e}_\alpha$$

I.e. we have:

$$\nabla_\beta \vec{V} = \nabla_\beta (V^\alpha \vec{e}_\alpha) = (\nabla_\beta V^\alpha)\vec{e}_\alpha$$

We also define:

$$\nabla_{\vec{X}} \vec{V} = \nabla_{X^\beta \vec{e}_\beta} \vec{V} \equiv X^\beta \nabla_\beta \vec{V} = X^\beta (\nabla_\beta V^\alpha)\vec{e}_\alpha$$

A scalar doesn't depend on basis vectors, so its covariant derivative is just its partial derivative

$$\nabla_\alpha \phi = \frac{\partial \phi}{\partial x^\alpha}$$

Differentiating a one form p_α is done using the fact, that $\phi = p_\alpha V^\alpha$ is a scalar, thus

$$\nabla_\beta \phi = \frac{\partial p_\alpha V^\alpha}{\partial x^\beta} = \frac{\partial p_\alpha}{\partial x^\beta} V^\alpha + p_\alpha \frac{\partial V^\alpha}{\partial x^\beta} = \frac{\partial p_\alpha}{\partial x^\beta} V^\alpha + p_\alpha \left(\nabla_\beta V^\alpha - \Gamma^\alpha_{\mu\beta} V^\mu \right) =$$

$$= V^\alpha \left(\frac{\partial p_\alpha}{\partial x^\beta} - \Gamma^\mu_{\alpha\beta} p_\mu \right) + p_\alpha \nabla_\beta V^\alpha = V^\alpha \nabla_\beta p_\alpha + p_\alpha \nabla_\beta V^\alpha$$

where we have defined

$$\nabla_\beta p_\alpha = \frac{\partial p_\alpha}{\partial x^\beta} - \Gamma^\mu_{\alpha\beta} p_\mu$$

This is obviously a tensor, because the above equation has a tensor on the left hand side ($\nabla_\beta \phi$) and tensors on the right hand side ($p_\alpha \nabla_\beta V^\alpha$ and V^α). Similarly for the derivative of the tensor $A^{\mu\nu}$ we use the fact that $V^\mu = A^{\mu\nu} p_\nu$ is a vector:

$$\nabla_\beta V^\mu = \nabla_\beta (A^{\mu\nu} p_\nu) = \partial_\beta (A^{\mu\nu} p_\nu) + \Gamma^\mu_{\alpha\beta} A^{\alpha\nu} p_\nu = p_\nu \partial_\beta A^{\mu\nu} + A^{\mu\nu} \partial_\beta p_\nu + \Gamma^\mu_{\alpha\beta} A^{\alpha\nu} p_\nu =$$

$$= p_\nu \partial_\beta A^{\mu\nu} + A^{\mu\nu} \left(\nabla_\beta p_\nu + \Gamma^\mu_{\nu\beta} p_\mu \right) + \Gamma^\mu_{\alpha\beta} A^{\alpha\nu} p_\nu = p_\nu \nabla_\beta A^{\mu\nu} + A^{\mu\nu} \nabla_\beta p_\nu$$

where we define

$$\nabla_\beta A^{\mu\nu} = \partial_\beta A^{\mu\nu} + \Gamma^\mu_{\alpha\beta} A^{\alpha\nu} + \Gamma^\nu_{\alpha\beta} A^{\mu\alpha}$$

and so on for other tensors, for example:

$$\nabla_\beta A^\mu{}_\nu = \partial_\beta A^\mu{}_\nu + \Gamma^\mu_{\alpha\beta} A^\alpha{}_\nu - \Gamma^\alpha_{\nu\beta} A^\mu{}_\alpha$$

$$\nabla_\beta A_{\mu\nu} = \partial_\beta A_{\mu\nu} - \Gamma^\alpha_{\mu\beta} A_{\alpha\nu} - \Gamma^\alpha_{\nu\beta} A_{\mu\alpha}$$

One can now easily proof some common relations simply by rewriting it to components and back:

$$\nabla_{\vec{X}} (f\vec{Y}) = (\nabla_{\vec{X}} f)\vec{Y} + f \nabla_{\vec{X}} \vec{Y}$$

$$\nabla_{\vec{X}} (\vec{Y} + \vec{Z}) = \nabla_{\vec{X}} \vec{Y} + \nabla_{\vec{X}} \vec{Z}$$

$$\nabla_{f\vec{X}} \vec{Y} = f \nabla_{\vec{X}} \vec{Y}$$

Change of variable:

$$\Gamma'^\alpha{}_{\beta\gamma} = \frac{\partial x^\mu}{\partial x'^\beta} \frac{\partial x^\nu}{\partial x'^\gamma} \Gamma^\sigma{}_{\mu\nu} \frac{\partial x'^\alpha}{\partial x^\sigma} + \frac{\partial x'^\alpha}{\partial x^\sigma} \frac{\partial^2 x^\sigma}{\partial x'^\beta \partial x'^\gamma}$$

Parallel transport

If the vectors $\vec{V}$ at infinitesimally close points of the curve $x^\mu(\lambda)$ are parallel and of equal length, then $\vec{V}$ is said to be parallel transported along the curve, i.e.:

$$\frac{\mathrm{d}\vec{V}}{\mathrm{d}\lambda} = 0$$

So

$$\frac{\mathrm{d}\vec{V}}{\mathrm{d}\lambda} = \frac{\mathrm{d}(V^\alpha \vec{e}_\alpha)}{\mathrm{d}\lambda} = \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda}\partial_\beta(V^\alpha \vec{e}_\alpha) = \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda}(\nabla_\beta V^\alpha)\vec{e}_\alpha = 0$$

In components (using the tangent vector $U^\beta = \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda}$):

$$\frac{\mathrm{d}V^\alpha}{\mathrm{d}\lambda} = U^\beta \nabla_\beta V^\alpha = 0$$

Fermi-Walker transport

In local inertial frame:

$$U_0^\lambda = (1,0,0,0)$$

$$\frac{\mathrm{d}S^i}{\mathrm{d}t} = 0$$

We require orthogonality $S_\mu U^\mu = 0$, in a general frame:

$$\frac{\mathrm{d}S^\alpha}{\mathrm{d}\tau} = \lambda U^\alpha = S_\mu \frac{\mathrm{d}U^\mu}{\mathrm{d}\tau} U^\alpha$$

where λ was calculated by differentiating the orthogonality condition. This is called a Thomas precession.

For any vector, we define: the vector X^μ is Fermi-Walker tranported along the curve if:

$$\frac{\mathrm{d}X^\mu}{\mathrm{d}\lambda} = X_\alpha \frac{\mathrm{d}U^\alpha}{\mathrm{d}\lambda} U^\mu - X_\alpha U^\alpha \frac{\mathrm{d}U^\mu}{\mathrm{d}\lambda}$$

If X^μ is perpendicular to U^μ, the second term is zero and the result is called a Fermi transport.

Why: the U^μ is transported by Fermi-Walker and also this is the equation for gyroscopes, so the natural, nonrotating tetrade is the one with $\vec{e}_0^\mu \equiv U^\mu$, which is then correctly transported along any curve (not just geodesics).

Geodesics

Geodesics is a curve $x^\alpha(\lambda)$ that locally looks like a line, i.e. it parallel transports its own tangent vector:

$$U^\beta \nabla_\beta U^\alpha = 0$$

so

$$U^\beta \partial_\beta U^\alpha + \Gamma^\alpha_{\beta\gamma} U^\beta U^\gamma = 0$$

or equivalently (using the fact $U^\beta \partial_\beta U^\alpha = \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda}\frac{\partial}{\partial x^\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda} = \frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}\lambda^2}$):

$$\frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}\lambda^2} + \Gamma^\alpha_{\beta\gamma}\frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\gamma}{\mathrm{d}\lambda} = 0 \tag{3.40.1.1}$$

Let's determine all possible reparametrizations that leave the geodesic equation invariant:

$$\begin{aligned}
\lambda' &= f(\lambda) \\
\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda} &= \frac{\mathrm{d}\lambda'}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda'} = f'(\lambda)\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda'} \\
\frac{\mathrm{d}^2x^\alpha}{\mathrm{d}\lambda^2} = \frac{\mathrm{d}}{\mathrm{d}\lambda}\left(f'(\lambda)\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda'}\right) &= f''(\lambda)\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda'} + f'(\lambda)\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda'} = \\
&= f''(\lambda)\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda'} + f'^2(\lambda)\frac{\mathrm{d}}{\mathrm{d}\lambda'}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda'} = \\
&= f''(\lambda)\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda'} + f'^2(\lambda)\frac{\mathrm{d}^2x^\alpha}{\mathrm{d}\lambda'^2}
\end{aligned}$$

Substituting into the geodesic equation, we get:

$$f''(\lambda)\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda'} + f'^2(\lambda)\left(\frac{\mathrm{d}^2x^\alpha}{\mathrm{d}\lambda'^2} + \Gamma^\alpha_{\beta\gamma}\frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda'}\frac{\mathrm{d}x^\gamma}{\mathrm{d}\lambda'}\right) = 0$$

So we can see that the equation is invariant as long as $f''(\lambda) = 0$, which gives:

$$f(\lambda) = a\lambda + b$$

This is called an affine reparametrization.

Another way to derive the geodesic equation is by finding a curve that extremizes the proper time:

$$\tau = \int \mathrm{d}\tau = \int \sqrt{-\frac{1}{c^2}\mathrm{d}s^2} = \int \sqrt{-\frac{1}{c^2}g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu} = \int \sqrt{-\frac{1}{c^2}g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}}\mathrm{d}\lambda = \int K\mathrm{d}\lambda$$

Here λ can be *any* parametrization. We have introduced K to make the formulas shorter:

$$K = \sqrt{-\frac{1}{c^2}g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}}$$

We vary this action with respect to x^μ:

$$\begin{aligned}
\delta\tau = \delta\int K\mathrm{d}\lambda = \\
= \delta\int\sqrt{-\frac{1}{c^2}g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}}\mathrm{d}\lambda = \\
= \int\frac{-\frac{1}{c^2}\left((\delta g_{\mu\nu})\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}+g_{\mu\nu}\left(\delta\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\right)\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}+g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\left(\delta\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}\right)\right)}{2K}\mathrm{d}\lambda = \\
= \frac{1}{c^2}\int\frac{\left(-\frac{1}{2}(\delta g_{\mu\nu})\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}-g_{\mu\nu}\left(\delta\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\right)\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}\right)}{K}\mathrm{d}\lambda = \\
= \frac{1}{c^2}\int\frac{\left(-\frac{1}{2}(\delta x^\alpha)\partial_\alpha g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}-g_{\alpha\nu}\frac{\mathrm{d}(\delta x^\alpha)}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}\right)}{K}\mathrm{d}\lambda = \\
= \frac{1}{c^2}\int\left(\frac{-\frac{1}{2}(\delta x^\alpha)\partial_\alpha g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}}{K}+\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{g_{\alpha\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}}{K}\right)(\delta x^\alpha)\right)\mathrm{d}\lambda = \\
= \frac{1}{c^2}\int\frac{1}{K}\left(K\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{g_{\alpha\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}}{K}\right)-\frac{1}{2}\partial_\alpha g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}\right)(\delta x^\alpha)\mathrm{d}\lambda = \\
= \frac{1}{c^2}\int\frac{1}{K}\left(K\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{1}{K}\right)+\frac{\mathrm{d}g_{\alpha\nu}}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}+g_{\alpha\nu}\frac{\mathrm{d}^2x^\nu}{\mathrm{d}\lambda^2}-\frac{1}{2}\partial_\alpha g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}\right)(\delta x^\alpha)\mathrm{d}\lambda = \\
= \frac{1}{c^2}\int\frac{1}{K}\left(K\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{1}{K}\right)+g_{\alpha\nu}\frac{\mathrm{d}^2x^\nu}{\mathrm{d}\lambda^2}+\partial_\mu g_{\alpha\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}-\frac{1}{2}\partial_\alpha g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}\right)(\delta x^\alpha)\mathrm{d}\lambda = \\
= \frac{1}{c^2}\int\frac{1}{K}\left(K\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{1}{K}\right)+g_{\alpha\nu}\frac{\mathrm{d}^2x^\nu}{\mathrm{d}\lambda^2}+\frac{1}{2}(\partial_\mu g_{\alpha\nu}+\partial_\nu g_{\alpha\mu}-\partial_\alpha g_{\mu\nu})\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}\right)(\delta x^\alpha)\mathrm{d}\lambda = \\
= \frac{1}{c^2}\int\frac{1}{K}\left(K\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{1}{K}\right)+g_{\alpha\nu}\frac{\mathrm{d}^2x^\nu}{\mathrm{d}\lambda^2}+\frac{1}{2}(\partial_\mu g_{\alpha\nu}+\partial_\nu g_{\alpha\mu}-\partial_\alpha g_{\mu\nu})\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}\right)g^{\alpha\rho}(\delta x_\rho)\mathrm{d}\lambda = \\
= \frac{1}{c^2}\int\frac{1}{K}\left(K\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{1}{K}\right)g^{\alpha\rho}+\delta_\nu{}^\rho\frac{\mathrm{d}^2x^\nu}{\mathrm{d}\lambda^2}+\frac{1}{2}g^{\alpha\rho}(\partial_\mu g_{\alpha\nu}+\partial_\nu g_{\alpha\mu}-\partial_\alpha g_{\mu\nu})\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}\right)(\delta x_\rho)\mathrm{d}\lambda = \\
= \frac{1}{c^2}\int\frac{1}{K}\left(K\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{1}{K}\right)g^{\alpha\rho}+\frac{\mathrm{d}^2x^\rho}{\mathrm{d}\lambda^2}+\Gamma^\rho_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}\right)(\delta x_\rho)\mathrm{d}\lambda
\end{aligned}$$

By setting the variation $\delta\tau = 0$ we obtain the geodesic equation:

$$K\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{1}{K}\right)g^{\alpha\rho}+\frac{\mathrm{d}^2x^\rho}{\mathrm{d}\lambda^2}+\Gamma^\rho_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}=0 \tag{3.40.1.2}$$

We have a freedom of choosing λ, so we choose such parametrization so that $\mathrm{d}\lambda = \mathrm{d}\tau$, which makes $K = 1$ and we recover (3.40.1.1):

$$\frac{\mathrm{d}^2x^\rho}{\mathrm{d}\lambda^2}+\Gamma^\rho_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}=0$$

Note that the equation (3.40.1.2) is parametrization invariant, but (3.40.1.1) is not (only affine reparametrization leaves (3.40.1.1) invariant).

Riemann Curvature Tensor

Curvature means that we take a vector V^μ, parallel transport it around a closed loop (which is just applying a commutator of the covariant derivatives $[\nabla_\alpha, \nabla_\beta]V^\mu$) and see how it changes. We express the result in terms of the vector V^μ:

$$[\nabla_\alpha, \nabla_\beta]V^\mu \equiv R^\mu{}_{\nu\alpha\beta}V^\nu$$

The coefficients $R^{\mu}{}_{\nu\alpha\beta}$ form a tensor called Riemann curvature tensor. Expanding the left hand side:

$$\begin{aligned}
V^\mu = \nabla_\alpha\nabla_\beta V^\mu - \nabla_\alpha\nabla_\beta V^\mu = \\
= \nabla_\alpha\nabla_\beta V^\mu - (\alpha \leftrightarrow \beta) = \\
= \partial_\alpha\nabla_\beta V^\mu - \Gamma^\sigma_{\alpha\beta}\nabla_\sigma V^\mu + \Gamma^\mu_{\alpha\sigma}\nabla_\beta V^\sigma - (\alpha \leftrightarrow \beta) = \\
= \partial_\alpha\left(\partial_\beta V^\mu + \Gamma^\mu_{\beta\sigma}V^\sigma\right) - \Gamma^\sigma_{\alpha\beta}\left(\partial_\sigma V^\mu + \Gamma^\mu_{\sigma\nu}V^\nu\right) + \Gamma^\mu_{\alpha\sigma}\left(\partial_\beta V^\sigma + \Gamma^\sigma_{\beta\nu}V^\nu\right) - (\alpha \leftrightarrow \beta) = \\
= \partial_\alpha\left(\partial_\beta V^\mu + \Gamma^\mu_{\beta\sigma}V^\sigma\right) + \Gamma^\mu_{\alpha\sigma}\left(\partial_\beta V^\sigma + \Gamma^\sigma_{\beta\nu}V^\nu\right) - (\alpha \leftrightarrow \beta) = \\
= \partial_\alpha\partial_\beta V^\mu + \partial_\alpha\Gamma^\mu_{\beta\sigma}V^\sigma + \Gamma^\mu_{\beta\sigma}\partial_\alpha V^\sigma + \Gamma^\mu_{\alpha\sigma}\partial_\beta V^\sigma + \Gamma^\mu_{\alpha\sigma}\Gamma^\sigma_{\beta\nu}V^\nu - (\alpha \leftrightarrow \beta) = \\
= \partial_\alpha\Gamma^\mu_{\beta\sigma}V^\sigma + \Gamma^\mu_{\alpha\sigma}\Gamma^\sigma_{\beta\nu}V^\nu - (\alpha \leftrightarrow \beta) = \\
= \left(\partial_\alpha\Gamma^\mu_{\beta\nu} + \Gamma^\mu_{\alpha\sigma}\Gamma^\sigma_{\beta\nu} - (\alpha \leftrightarrow \beta)\right)V^\nu
\end{aligned}$$

Where we have used the fact that all terms symmetric in $\alpha\beta$ (in particular $\Gamma^\sigma_{\alpha\beta}$ and $\partial_\alpha\partial_\beta V^\mu$ and $\Gamma^\mu_{\beta\sigma}\partial_\alpha V^\sigma + \Gamma^\mu_{\alpha\sigma}\partial_\beta V^\sigma$) get canceled by the same term in the $(\alpha \leftrightarrow \beta)$. We get

$$R^{\mu}{}_{\nu\alpha\beta} = \partial_\alpha\Gamma^\mu_{\beta\nu} + \Gamma^\mu_{\alpha\sigma}\Gamma^\sigma_{\beta\nu} - (\alpha \leftrightarrow \beta) = \partial_\alpha\Gamma^\mu_{\beta\nu} - \partial_\beta\Gamma^\mu_{\alpha\nu} + \Gamma^\mu_{\alpha\sigma}\Gamma^\sigma_{\beta\nu} - \Gamma^\mu_{\beta\sigma}\Gamma^\sigma_{\alpha\nu}$$

In order to see all the symmetries, that the Riemann tensor has, we lower the first index

$$R_{\mu\nu\alpha\beta} = g_{\mu\lambda}R^{\lambda}{}_{\nu\alpha\beta} = g_{\mu\lambda}\left(\partial_\alpha\Gamma^\lambda_{\beta\nu} + \Gamma^\lambda_{\alpha\sigma}\Gamma^\sigma_{\beta\nu} - (\alpha \leftrightarrow \beta)\right)$$

and use local inertial frame coordinates, where all Christoffel symbols vanish (not their derivatives though):

$$\begin{aligned}
R_{\mu\nu\alpha\beta} = g_{\mu\lambda}\left(\partial_\alpha\Gamma^\lambda_{\beta\nu} - (\alpha \leftrightarrow \beta)\right) \\
= g_{\mu\lambda}\left(\partial_\alpha\left(\tfrac{1}{2}g^{\lambda\sigma}(\partial_\beta g_{\sigma\nu} + \partial_\nu g_{\sigma\beta} - \partial_\sigma g_{\beta\nu})\right) - (\alpha \leftrightarrow \beta)\right) = \\
= \tfrac{1}{2}g_{\mu\lambda}g^{\lambda\sigma}\left(\partial_\alpha\left(\partial_\beta g_{\sigma\nu} + \partial_\nu g_{\sigma\beta} - \partial_\sigma g_{\beta\nu}\right) - (\alpha \leftrightarrow \beta)\right) = \\
= \tfrac{1}{2}\delta_\mu{}^\sigma\left(\partial_\alpha\left(\partial_\beta g_{\sigma\nu} + \partial_\nu g_{\sigma\beta} - \partial_\sigma g_{\beta\nu}\right) - (\alpha \leftrightarrow \beta)\right) = \\
= \tfrac{1}{2}\left(\partial_\alpha\partial_\beta g_{\mu\nu} + \partial_\alpha\partial_\nu g_{\mu\beta} - \partial_\alpha\partial_\mu g_{\beta\nu} - (\alpha \leftrightarrow \beta)\right) = \\
= \tfrac{1}{2}\left(\partial_\alpha\partial_\nu g_{\mu\beta} - \partial_\alpha\partial_\mu g_{\beta\nu} - (\alpha \leftrightarrow \beta)\right) = \\
= \tfrac{1}{2}\left(\partial_\alpha\partial_\nu g_{\mu\beta} - \partial_\alpha\partial_\mu g_{\beta\nu} - \partial_\beta\partial_\nu g_{\mu\alpha} + \partial_\beta\partial_\mu g_{\alpha\nu}\right)
\end{aligned}$$

We will also need:

$$\begin{aligned}
\nabla_\lambda R_{\mu\nu\alpha\beta} = \tfrac{1}{2}\partial_\lambda\left(\partial_\alpha\partial_\nu g_{\mu\beta} - \partial_\alpha\partial_\mu g_{\beta\nu} - (\alpha \leftrightarrow \beta)\right) = \\
= \tfrac{1}{2}\left(\partial_\lambda\partial_\alpha\partial_\nu g_{\mu\beta} - \partial_\lambda\partial_\alpha\partial_\mu g_{\beta\nu} - \partial_\lambda\partial_\beta\partial_\nu g_{\mu\alpha} + \partial_\lambda\partial_\beta\partial_\mu g_{\alpha\nu}\right)
\end{aligned}$$

Using these expressions for the curvature tensor in a local inertial frame, we derive the following 5 symmetries of the curvature tensor by simply substituting for the left hand side and verify that it is equal to the right hand side:

$$\begin{aligned}
R_{\mu\nu\alpha\beta} = -R_{\mu\nu\beta\alpha} \\
R_{\mu\nu\alpha\beta} = -R_{\nu\mu\alpha\beta} \\
R_{\mu\nu\alpha\beta} = R_{\alpha\beta\mu\nu} \\
R_{\mu\nu\alpha\beta} + R_{\mu\alpha\beta\nu} + R_{\mu\beta\nu\alpha} = 0 \\
\nabla_\lambda R_{\mu\nu\alpha\beta} + \nabla_\mu R_{\nu\lambda\alpha\beta} + \nabla_\nu R_{\lambda\mu\alpha\beta} = 0
\end{aligned}$$

These are tensor expressions and so even though we derived them in a local inertial frame, they hold in all coordinates. The last identity is called a Bianchi identity.

The Ricci tensor is defined as:

$$R_{\mu\nu} = R^{\lambda}{}_{\mu\lambda\nu} = g^{\lambda\sigma}R_{\sigma\mu\lambda\nu}$$

From the last equality we can see that it is symmetric in $\mu\nu$. A Ricci scalar is defined as:

$$R = R^{\mu}{}_{\mu} = g^{\mu\nu}R_{\mu\nu}$$

The Einstein tensor is defined as:

$$G_{\mu\nu} = R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu}$$

It is symmetric in $\mu\nu$ due to the symmetry of the metric and Ricci tensors. By contracting the Bianchi identity twice, we can show that Einstein tensor has zero divergence:

$$\begin{aligned}
g^{\lambda\alpha}g^{\nu\beta}\left(\nabla_\lambda R_{\mu\nu\alpha\beta} + \nabla_\mu R_{\nu\lambda\alpha\beta} + \nabla_\nu R_{\lambda\mu\alpha\beta}\right) = \\
= \nabla^\alpha g^{\nu\beta}R_{\mu\nu\alpha\beta} + \nabla_\mu g^{\lambda\alpha}R^{\beta}{}_{\lambda\alpha\beta} + \nabla^\beta g^{\lambda\alpha}R_{\lambda\mu\alpha\beta} = \\
= \nabla^\alpha g^{\nu\beta}R_{\nu\mu\beta\alpha} - \nabla_\mu g^{\lambda\alpha}R^{\beta}{}_{\lambda\beta\alpha} + \nabla^\beta R^{\alpha}{}_{\mu\alpha\beta} = \\
= \nabla^\alpha R^{\beta}{}_{\mu\beta\alpha} - \nabla_\mu g^{\lambda\alpha}R_{\lambda\alpha} + \nabla^\beta R_{\mu\beta} = \\
= \nabla^\alpha R_{\mu\alpha} - \nabla_\mu R^{\alpha}{}_{\alpha} + \nabla^\beta R_{\mu\beta} = \\
= 2\nabla^\alpha R_{\mu\alpha} - \nabla_\mu R = \\
= 2\nabla^\alpha\left(R_{\mu\alpha} - \tfrac{1}{2}g_{\mu\alpha}R\right) = \\
= 2\nabla^\alpha G_{\mu\alpha} = 0
\end{aligned}$$

Lie derivative

Definition of the Lie derivative of any tensor T is:

$$\mathcal{L}_{\vec{U}}T = \lim_{t\to 0}\frac{\phi_{t*}T(\phi_t(p)) - T(p)}{t}$$

it can be shown directly from this definition, that the Lie derivative of a vector is the same as a Lie bracket:

$$\mathcal{L}_{\vec{U}}\vec{V} \equiv [\vec{U}, \vec{V}]$$

and in components

$$\mathcal{L}_{\vec{U}}V^\alpha = [\vec{U}, \vec{V}]^\alpha \equiv U^\beta\nabla_\beta V^\alpha - V^\beta\nabla_\beta U^\alpha = U^\beta\partial_\beta V^\alpha - V^\beta\partial_\beta U^\alpha$$

Lie derivative of a scalar is

$$\mathcal{L}_{\vec{V}}f = V^\mu\partial_\mu f$$

and of a one form p_μ is derived using the observation that $f = p_\mu V^\mu$ is a scalar:

$$\mathcal{L}_{\vec{V}}p_\mu = V^\nu\nabla_\nu p_\mu + p_\nu\nabla_\mu V^\nu = V^\nu\partial_\nu p_\mu + p_\nu\partial_\mu V^\nu$$

and so on for other tensors, for example:

$$\mathcal{L}_{\vec{V}}g_{\mu\nu} = V^\alpha\nabla_\alpha g_{\mu\nu} + g_{\alpha\nu}\nabla_\mu V^\alpha + g_{\mu\alpha}\nabla_\nu V^\alpha = V^\alpha\partial_\alpha g_{\mu\nu} + g_{\alpha\nu}\partial_\mu V^\alpha + g_{\mu\alpha}\partial_\nu V^\alpha$$

Metric

In general, the Christoffel symbols are not symmetric and there is no metric that generates them. However, if the manifold is equipped with metrics, then the fundamental theorem of Riemannian geometry states that there is a unique Levi-Civita connection, for which the metric tensor is preserved by parallel transport:

$$\nabla_\mu g_{\alpha\beta} = 0$$

We define the commutation coefficients of the basis ${c^\alpha}_{\mu\nu}$ by

$${c^\alpha}_{\mu\nu}\vec{e}_\alpha = \nabla_{\vec{e}_\mu}\vec{e}_\nu - \nabla_{\vec{e}_\nu}\vec{e}_\mu$$

In general these coefficients are not zero (as an example, take the units vectors in spherical or cylindrical coordinates), but for coordinate bases they are. It can be proven, that

$$\Gamma^\mu_{\alpha\beta} = \tfrac{1}{2}g^{\mu\sigma}\left(\partial_\beta g_{\sigma\alpha} + \partial_\alpha g_{\sigma\beta} - \partial_\sigma g_{\alpha\beta} + c_{\alpha\sigma\beta} + c_{\beta\sigma\alpha} - c_{\sigma\alpha\beta}\right)$$

and for coordinate bases ${c^\alpha}_{\mu\nu} = 0$, so

$$\Gamma^\mu_{\alpha\beta} = \Gamma^\mu_{\beta\alpha}$$
$$\Gamma^\mu_{\alpha\beta} = \tfrac{1}{2}g^{\mu\sigma}\left(\partial_\beta g_{\sigma\alpha} + \partial_\alpha g_{\sigma\beta} - \partial_\sigma g_{\alpha\beta}\right)$$

As a special case:

$$\begin{aligned}\Gamma^\mu_{\mu\beta} = \tfrac{1}{2}g^{\mu\sigma}\left(\partial_\beta g_{\sigma\mu} + \partial_\mu g_{\sigma\beta} - \partial_\sigma g_{\mu\beta}\right) = \tfrac{1}{2}g^{\mu\sigma}\partial_\beta g_{\sigma\mu} = \\ = \tfrac{1}{2}\operatorname{Tr} g^{-1}\partial_\beta g = \tfrac{1}{2}\operatorname{Tr}\partial_\beta \log g = \tfrac{1}{2}\partial_\beta \operatorname{Tr}\log g = \tfrac{1}{2}\partial_\beta \log|\det g| = \partial_\beta \log\sqrt{|\det g|} = \\ = \frac{1}{2\det g}\partial_\beta \det g = \frac{1}{\sqrt{|\det g|}}\partial_\beta\sqrt{|\det g|}\end{aligned}$$

All last 3 expressions are used (but the last one is probably the most common). g is the matrix of coefficients $g_{\mu\nu}$. At the beginning we used the usual trick that $g^{\mu\sigma}$ is symmetric but $\partial_\mu g_{\sigma\beta} - \partial_\sigma g_{\mu\beta}$ is antisymmetric. Later we used the identity $\operatorname{Tr}\log g = \log|\det g|$, which follows from the well-known identity $\det\exp A = \exp\operatorname{Tr} A$ by substituting $A = \log g$ and taking the logarithm of both sides.

Diagonal Metric

Many times the metric is diagonal, e.g. in 3D:

$$g_{ij} = \begin{pmatrix} h_1^2 & 0 & 0 \\ 0 & h_2^2 & 0 \\ 0 & 0 & h_3^2 \end{pmatrix}$$

(in general $g_{ij} = \operatorname{diag}(h_1^2, h_2^2, \dots)$), then the Christoffel symbols Γ^k_{ij} can be calculated very easily (below we do not sum over i, j and k):

$$\Gamma^k_{ij} = \tfrac{1}{2}g^{kl}\left(\partial_j g_{li} + \partial_i g_{lj} - \partial_l g_{ij}\right) = \tfrac{1}{2}g^{kk}\left(\partial_j g_{ki} + \partial_i g_{kj} - \partial_k g_{ij}\right)$$

If $k = i$ or $k = j$ then

$$\Gamma^k_{ij} = \Gamma^i_{ij} = \Gamma^i_{ji} = \tfrac{1}{2}g^{ii}\left(\partial_j g_{ii} + \partial_i g_{ij} - \partial_i g_{ij}\right) = \tfrac{1}{2}g^{ii}\partial_j g_{ii} = \tfrac{1}{2}\frac{1}{h_i^2}\partial_j h_i^2 = \frac{1}{h_i}\partial_j h_i \tag{3.40.1.3}$$

otherwise (i.e. $k \neq i$ and $k \neq j$) then either $i = j$:

$$\Gamma^k_{ij} = \Gamma^k_{ii} = \tfrac{1}{2}g^{kk}\left(\partial_i g_{ki} + \partial_i g_{ki} - \partial_k g_{ii}\right) = -\tfrac{1}{2}g^{kk}\partial_k g_{ii} = -\tfrac{1}{2}\frac{1}{h_k^2}\partial_k h_i^2 = -\frac{h_i}{h_k^2}\partial_k h_i \tag{3.40.1.4}$$

or $i \neq j$ (i.e. $i \neq j \neq k$):

$$\Gamma^k_{ij} = \tfrac{1}{2} g^{kk} \left(\partial_j g_{ki} + \partial_i g_{kj} - \partial_k g_{ij} \right) = 0$$

In other words, the symbols can only be nonzero if at least two of i, j or k are the same and one can use the two formulas (3.40.1.3) and (3.40.1.4) to quickly evaluate them. A systematic way to do it is to write (3.40.1.3) and (3.40.1.4) in the following form:

$$\begin{aligned} \Gamma^i_{ij} = \Gamma^i_{ji} = \frac{1}{h_i} \partial_j h_i \qquad & i, j \text{ arbitrary} \\ \Gamma^j_{ii} = -\frac{h_i}{h_j^2} \partial_j h_i \qquad & i \neq j \end{aligned} \tag{3.40.1.5}$$

Then find all i and j for which $\partial_j h_i$ is nonzero and then immediately write all nonzero Christoffel symbols using the equations (3.40.1.5).

For example for cylindrical coordinates we have $h_\rho = h_z = 1$ and $h_\phi = \rho$, so $\partial_j h_i$ is only nonzero for $i = \phi$ and $j = \rho$ and we get:

$$\begin{aligned} \Gamma^\phi_{\phi\rho} = \Gamma^\phi_{\rho\phi} = \frac{1}{h_\phi} \partial_\rho h_\phi = \frac{1}{\rho} \partial_\rho \rho = \frac{1}{\rho} \\ \Gamma^\rho_{\phi\phi} = -\frac{h_\phi}{h_\rho^2} \partial_\rho h_\phi = -\frac{\rho}{1^2} \partial_\rho \rho = -\rho \end{aligned}$$

all other Christoffel symbols are zero. For spherical coordinates we have $h_\rho = 1$, $h_\theta = \rho$ and $h_\phi = \rho \sin\theta$, so $\partial_j h_i$ is only nonzero for $i = \theta$, $j = \rho$ or $i = \phi$, $j = \rho$ or $i = \phi$, $j = \theta$ and we get:

$$\begin{aligned} \Gamma^\theta_{\theta\rho} = \Gamma^\theta_{\rho\theta} = \frac{1}{h_\theta} \partial_\rho h_\theta = \frac{1}{\rho} \partial_\rho \rho = \frac{1}{\rho} \\ \Gamma^\rho_{\theta\theta} = -\frac{h_\theta}{h_\rho^2} \partial_\rho h_\theta = -\frac{\rho}{1^2} \partial_\rho \rho = -\rho \\ \Gamma^\phi_{\phi\rho} = \Gamma^\phi_{\rho\phi} = \frac{1}{h_\phi} \partial_\rho h_\phi = \frac{1}{\rho \sin\theta} \partial_\rho (\rho \sin\theta) = \frac{1}{\rho} \\ \Gamma^\rho_{\phi\phi} = -\frac{h_\phi}{h_\rho^2} \partial_\rho h_\phi = -\frac{\rho \sin\theta}{1^2} \partial_\rho (\rho \sin\theta) = -\rho \sin^2\theta \\ \Gamma^\phi_{\phi\theta} = \Gamma^\phi_{\theta\phi} = \frac{1}{h_\phi} \partial_\theta h_\phi = \frac{1}{\rho \sin\theta} \partial_\theta (\rho \sin\theta) = \frac{\cos\theta}{\sin\theta} \\ \Gamma^\theta_{\phi\phi} = -\frac{h_\phi}{h_\theta^2} \partial_\theta h_\phi = -\frac{\rho \sin\theta}{\rho^2} \partial_\theta (\rho \sin\theta) = -\sin\theta \cos\theta \end{aligned}$$

All other symbols are zero.

Symmetries, Killing vectors

We say that a diffeomorphism ϕ is a symmetry of some tensor T if the tensor is invariant after being pulled back under ϕ:

$$\phi_* T = T$$

Let the one-parameter family of symmetries ϕ_t be generated by a vector field $V^\mu(x)$, then the above equation is equivalent to:

$$\mathcal{L}_{\vec{V}} T = 0$$

If T is the metric $g_{\mu\nu}$ then the symmetry is called isometry and V^μ is called a Killing vector field and can be calculated from:

$$\mathcal{L}_{\vec{V}} g_{\mu\nu} = V^\alpha \nabla_\alpha g_{\mu\nu} + g_{\alpha\nu}\nabla_\mu V^\alpha + g_{\mu\alpha}\nabla_\nu V^\alpha = \nabla_\mu V_\nu + \nabla_\nu V_\mu = 0$$

The last equality is Killing's equation. If x^μ is a geodesics with a tangent vector U^μ and V^μ is a Killing vector, then the quantity $V_\mu U^\mu$ is conserved along the geodesics, because:

$$\frac{\mathrm{d}(V_\mu U^\mu)}{\mathrm{d}\lambda} = U^\nu \nabla_\nu (V_\mu U^\mu) = U^\nu U^\mu \nabla_\nu V_\mu + V_\mu U^\nu \nabla_\nu U^\mu = 0$$

where the first term is both symmetric and antisymmetric in (μ, ν), thus zero, and the second term is the geodesics equation, thus also zero.

Symmetry and Antisymmetry

Every tensor can be decomposed into symmetric and antisymmetric parts:

$$T_{\alpha\beta} = \tfrac{1}{2}(T_{\alpha\beta} + T_{\beta\alpha}) + \tfrac{1}{2}(T_{\alpha\beta} - T_{\beta\alpha})$$

In particular, for a symmetric tensor $S_{\alpha\beta} = S_{\beta\alpha}$ we get:

$$S_{\alpha\beta} = \tfrac{1}{2}(S_{\alpha\beta} + S_{\beta\alpha})$$

and for antisymmetric tensor $A_{\alpha\beta} = -A_{\beta\alpha}$ we get:

$$A_{\alpha\beta} = \tfrac{1}{2}(A_{\alpha\beta} - A_{\beta\alpha})$$

When contracting a symmetric tensor with an antisymmetric tensor we get zero:

$$\begin{aligned} S_{\alpha\beta}A^{\alpha\beta} &= \\ &= \tfrac{1}{2}S_{\alpha\beta}(A^{\alpha\beta} - A^{\beta\alpha}) = \\ &= \tfrac{1}{2}(S_{\alpha\beta}A^{\alpha\beta} - S_{\alpha\beta}A^{\beta\alpha}) = \\ &= \tfrac{1}{2}(S_{\alpha\beta}A^{\alpha\beta} - S_{\beta\alpha}A^{\beta\alpha}) = \\ &= \tfrac{1}{2}(S_{\alpha\beta}A^{\alpha\beta} - S_{\alpha\beta}A^{\alpha\beta}) = \\ &= 0 \end{aligned}$$

When contracting a general tensor T with a symmetric tensor S, only the symmetric part of T contributes:

$$\begin{aligned} T_{\alpha\beta}S^{\alpha\beta} &= \\ &= \tfrac{1}{2}(T_{\alpha\beta} + T_{\beta\alpha})S^{\alpha\beta} + \tfrac{1}{2}(T_{\alpha\beta} - T_{\beta\alpha})S^{\alpha\beta} = \\ &= \tfrac{1}{2}(T_{\alpha\beta} + T_{\beta\alpha})S^{\alpha\beta} \end{aligned}$$

When contracting a general tensor T with an antisymmetric tensor A, only the antisymmetric part of T contributes:

$$\begin{aligned} T_{\alpha\beta}A^{\alpha\beta} &= \\ &= \tfrac{1}{2}(T_{\alpha\beta} + T_{\beta\alpha})A^{\alpha\beta} + \tfrac{1}{2}(T_{\alpha\beta} - T_{\beta\alpha})A^{\alpha\beta} = \\ &= \tfrac{1}{2}(T_{\alpha\beta} - T_{\beta\alpha})A^{\alpha\beta} \end{aligned}$$

Example I

We want to rewrite:

$$2\partial_\nu g_{\mu\lambda}\frac{\mathrm{d}x^\mu}{\mathrm{d}s}\frac{\mathrm{d}x^\nu}{\mathrm{d}s}$$

So we write the left part as a sum of symmetric and antisymmetric parts:

$$(\partial_\nu g_{\mu\lambda}+\partial_\mu g_{\nu\lambda}+\partial_\nu g_{\mu\lambda}-\partial_\mu g_{\nu\lambda})\frac{\mathrm{d}x^\mu}{\mathrm{d}s}\frac{\mathrm{d}x^\nu}{\mathrm{d}s}$$

Here $\partial_\nu g_{\mu\lambda}-\partial_\mu g_{\nu\lambda}$ is antisymmetric and $\frac{\mathrm{d}x^\mu}{\mathrm{d}s}\frac{\mathrm{d}x^\nu}{\mathrm{d}s}$ is symmetric in μ,ν, so the contraction is zero. The final result is:

$$(\partial_\nu g_{\mu\lambda}+\partial_\mu g_{\nu\lambda})\frac{\mathrm{d}x^\mu}{\mathrm{d}s}\frac{\mathrm{d}x^\nu}{\mathrm{d}s}$$

Example II

Let $F_{\mu\nu}=\partial_\mu A_\nu-\partial_\nu A_\mu$. Then we can simplify:

$$F^{\mu\nu}F_{\mu\nu}=F^{\mu\nu}(\partial_\mu A_\nu-\partial_\nu A_\mu)=$$
$$=2F^{\mu\nu}\tfrac{1}{2}(\partial_\mu A_\nu-\partial_\nu A_\mu)=2F^{\mu\nu}\partial_\mu A_\nu$$

Here $\frac{1}{2}(\partial_\mu A_\nu-\partial_\nu A_\mu)$ is the antisymmetric part (the only one that contributes, because $F^{\mu\nu}$ is antisymmetric) of $\partial_\mu A_\nu$.

Example III

Let $F_{\mu\nu}=\partial_\mu A_\nu-\partial_\nu A_\mu$. Then we get

$$F_{\mu\nu}v^\mu v^\nu=0\,,$$

because $F_{\mu\nu}$ is an antisymmetric tensor, while $v^\mu v^\nu$ is a symmetric tensor.

Divergence Operator

$$\nabla_\mu A^\mu=\partial_\mu A^\mu+\Gamma^\mu_{\mu\sigma}A^\sigma=$$
$$=\partial_\mu A^\mu+\frac{1}{\sqrt{|\det g|}}\left(\partial_\sigma\sqrt{|\det g|}\right)A^\sigma=$$
$$=\frac{1}{\sqrt{|\det g|}}\partial_\mu\left(\sqrt{|\det g|}\,A^\mu\right)$$

If the metric is diagonal (let's show this in 3D):

$$g_{ij}=\begin{pmatrix}h_1^2&0&0\\0&h_2^2&0\\0&0&h_3^2\end{pmatrix}$$

then

$$\sqrt{|\det g_{ij}|}=h_1h_2h_3$$
$$g^{ij}=\begin{pmatrix}\frac{1}{h_1^2}&0&0\\0&\frac{1}{h_2^2}&0\\0&0&\frac{1}{h_3^2}\end{pmatrix}$$

and

$$\nabla \cdot \mathbf{A} = \nabla_i A^i = \frac{1}{h_1 h_2 h_3} \partial_i \left(h_1 h_2 h_3 A^i \right)$$

Laplace Operator

$$\begin{aligned} \nabla^2 \varphi = \nabla_\mu \nabla^\mu \varphi = \partial_\mu \nabla^\mu \varphi + \Gamma^\mu_{\mu\sigma} \nabla^\sigma \varphi = \partial_\mu \partial^\mu \varphi + \Gamma^\mu_{\mu\sigma} \partial^\sigma \varphi = \\ = \partial_\mu \partial^\mu \varphi + \frac{1}{\sqrt{|\det g|}} \left(\partial_\sigma \sqrt{|\det g|} \right) \partial^\sigma \varphi = \\ = \frac{1}{\sqrt{|\det g|}} \partial_\mu \left(\sqrt{|\det g|}\, \partial^\mu \varphi \right) = \frac{1}{\sqrt{|\det g|}} \partial_\mu \left(\sqrt{|\det g|}\, g^{\mu\sigma} \partial_\sigma \varphi \right) \end{aligned}$$

If the metric is diagonal (let's show this in 3D):

$$g_{ij} = \begin{pmatrix} h_1^2 & 0 & 0 \\ 0 & h_2^2 & 0 \\ 0 & 0 & h_3^2 \end{pmatrix}$$

then

$$\sqrt{|\det g_{ij}|} = h_1 h_2 h_3$$

$$g^{ij} = \begin{pmatrix} \frac{1}{h_1^2} & 0 & 0 \\ 0 & \frac{1}{h_2^2} & 0 \\ 0 & 0 & \frac{1}{h_3^2} \end{pmatrix}$$

and

$$\nabla^2 \varphi = \sum_i \frac{1}{h_1 h_2 h_3} \partial_i \left(\frac{h_1 h_2 h_3}{h_i^2} \partial_i \varphi \right)$$

Covariant integration

If $f(x)$ is a scalar, then the integral $\int f(x) \mathrm{d}^4 x$ depends on coordinates. The correct way to integrate $f(x)$ in any coordinates is:

$$\int f(x) \sqrt{|g|} \mathrm{d}^4 x$$

where $g \equiv \det g_{\mu\nu}$. The Gauss theorem in curvilinear coordinates is:

$$\begin{aligned} \int_\Omega \nabla_\mu u^\mu \sqrt{|g|} \mathrm{d}^4 x = \int_\Omega \frac{1}{\sqrt{|g|}} \partial_\mu \left(\sqrt{|g|} u^\mu \right) \sqrt{|g|} \mathrm{d}^4 x = \int_\Omega \partial_\mu \left(\sqrt{|g|} u^\mu \right) \mathrm{d}^4 x = \\ = \int_{\partial\Omega} \sqrt{|g|} u^\mu n_\mu \mathrm{d}^3 x = \int_{\partial\Omega} u^\mu n_\mu \sqrt{|g|} \mathrm{d}^3 x \end{aligned}$$

where $\partial\Omega$ is the boundary (surface) of Ω and n_ν is the normal vector to this surface.

3.40.2 Examples

Weak Formulation of Laplace Equation

As an example, we write the weak formulation of the Laplace equation in arbitrary coordintes:

$$\begin{aligned}
\nabla^2\varphi - f = 0 \\
\int \left(\nabla^2\varphi v - fv\right)\sqrt{|g|}d^3x = 0 \\
\int \left(\frac{1}{\sqrt{|g|}}\partial_i\left(\sqrt{|g|}g^{ij}\partial_j\varphi\right)v - fv\right)\sqrt{|g|}d^3x = 0 \\
\int \left(\partial_i\left(\sqrt{|g|}g^{ij}\partial_j\varphi\right)v - fv\sqrt{|g|}\right)d^3x = 0
\end{aligned}$$

Now we apply per-partes (assuming the boundary integral vanishes):

$$\begin{aligned}
\int \left(-\sqrt{|g|}g^{ij}\partial_j\varphi\partial_i v - fv\sqrt{|g|}\right)d^3x = 0 \\
\int \left(-g^{ij}\partial_j\varphi\partial_i v - fv\right)\sqrt{|g|}d^3x = 0
\end{aligned}$$

For diagonal metric this evaluates to:

$$\int \left(-\sum_i \frac{1}{h_i^2}\partial_i\varphi\partial_i v - fv\right)h_1h_2h_3d^3x = 0$$

Cylindrical Coordinates

$$\begin{aligned}
x &= \rho\cos\phi \\
y &= \rho\sin\phi \\
z &= z
\end{aligned}$$

The transformation matrix is

$$\frac{\partial(x,y,z)}{\partial(\rho,\phi,z)} = \begin{pmatrix}\cos\phi & -\rho\sin\phi & 0 \\ \sin\phi & \rho\cos\phi & 0 \\ 0 & 0 & 1\end{pmatrix}$$

The metric tensor of the cartesian coordinate system $\hat{x}^a = (x,y,z)$ is $\hat{g}_{ab} = \mathrm{diag}(1,1,1)$, so by transformation we get the metric tensor g_{ij} in the cylindrical coordinates $x^i = (\rho,\phi,z)$:

$$g_{ij} = \frac{\partial\hat{x}^a}{\partial x^i}\frac{\partial\hat{x}^b}{\partial x^j}\hat{g}_{ab} = \left(\frac{\partial\hat{x}}{\partial x}\right)^T\hat{g}\frac{\partial\hat{x}}{\partial x} =$$

$$= \left(\frac{\partial(x,y,z)}{\partial(\rho,\phi,z)}\right)^T\begin{pmatrix}1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1\end{pmatrix}\frac{\partial(x,y,z)}{\partial(\rho,\phi,z)} =$$

$$= \begin{pmatrix}\cos\phi & \sin\phi & 0 \\ -\rho\sin\phi & \rho\cos\phi & 0 \\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\cos\phi & -\rho\sin\phi & 0 \\ \sin\phi & \rho\cos\phi & 0 \\ 0 & 0 & 1\end{pmatrix} = \begin{pmatrix}1 & 0 & 0 \\ 0 & \rho^2 & 0 \\ 0 & 0 & 1\end{pmatrix}$$

$$g^{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\rho^2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\det g = \det g_{ij} = \rho^2$$

$$h_\rho = h_z = 1$$

$$h_\phi = \rho$$

$$\Gamma^\phi_{\phi\rho} = \Gamma^\phi_{\rho\phi} = \frac{1}{h_\phi}\partial_\rho h_\phi = \frac{1}{\rho}\partial_\rho \rho = \frac{1}{\rho}$$

$$\Gamma^\rho_{\phi\phi} = -\frac{h_\phi}{h_\rho^2}\partial_\rho h_\phi = -\frac{\rho}{1^2}\partial_\rho \rho = -\rho$$

$$\nabla \cdot \mathbf{A} = \nabla_i A^i = \frac{1}{h_1 h_2 h_3}\partial_i \left(h_1 h_2 h_3 A^i\right) = \frac{1}{\rho}\partial_i \left(\rho A^i\right) =$$

$$= \frac{1}{\rho}\partial_\rho(\rho A^\rho) + \partial_\theta A^\theta + \partial_z A^z = \partial_\rho A^\rho + \frac{1}{\rho}A^\rho + \partial_\theta A^\theta + \partial_z A^z$$

$$\nabla^2 \varphi = \nabla^i \nabla_i \varphi = \frac{1}{\sqrt{|\det g|}}\partial_i \left(\sqrt{|\det g|}\, g^{ij}\partial_j \varphi\right) =$$

$$= \frac{1}{\rho}\partial_i \left(\rho g^{ij}\partial_j \varphi\right) = \frac{1}{\rho}\partial_\rho \left(\rho \partial_\rho \varphi\right) + \frac{1}{\rho}\partial_\phi \left(\rho \frac{1}{\rho^2}\partial_\phi \varphi\right) + \frac{1}{\rho}\partial_z \left(\rho \partial_z \varphi\right) =$$

$$= \frac{1}{\rho}\partial_\rho \left(\rho \partial_\rho \varphi\right) + \frac{1}{\rho^2}\partial_\phi \partial_\phi \varphi + \partial_z \partial_z \varphi =$$

$$= \partial_\rho \partial_\rho \varphi + \frac{1}{\rho}\partial_\rho \varphi + \frac{1}{\rho^2}\partial_\phi \partial_\phi \varphi + \partial_z \partial_z \varphi$$

As a particular example, let's write the Laplace equation with nonconstant conductivity for axially symmetric field. The Laplace equation is:

$$\nabla \cdot \sigma \nabla \varphi = 0$$

so we use the formulas above to get:

$$0 = \nabla \cdot \sigma \nabla \varphi = \nabla^i \sigma \nabla_i \varphi = \frac{\partial}{\partial \rho}\sigma \frac{\partial \varphi}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial}{\partial \phi}\sigma \frac{\partial \varphi}{\partial \phi} + \frac{\partial}{\partial z}\sigma \frac{\partial \varphi}{\partial z} + \frac{\sigma}{\rho}\frac{\partial \varphi}{\partial \rho}$$

but we know that $\varphi = \varphi(\rho, z)$, so $\frac{\partial \varphi}{\partial \phi} = 0$ and the final equation is:

$$\frac{\partial}{\partial \rho}\sigma \frac{\partial \varphi}{\partial \rho} + \frac{\partial}{\partial z}\sigma \frac{\partial \varphi}{\partial z} + \frac{\sigma}{\rho}\frac{\partial \varphi}{\partial \rho} = 0$$

To write the weak formulation for it, we need to integrate covariantly (e.g. $\rho \,\mathrm{d}\rho\mathrm{d}\phi\mathrm{d}z$ in our case) and rewrite it using per partes. We did exactly this in the previous example in a coordinate free maner, so we just use the final formula we got there for a diagonal metric:

$$\int \left(-\partial_\rho \varphi \partial_\rho v - \frac{1}{\rho^2}\partial_\phi \varphi \partial_\phi v - \partial_z \varphi \partial_z v\right) \sigma \rho \,\mathrm{d}\rho\mathrm{d}\phi\mathrm{d}z = 0$$

and for $\partial_\phi \varphi = 0$, we get:

$$-2\pi \int \left(\partial_\rho \varphi \partial_\rho v + \partial_z \varphi \partial_z v\right) \sigma \rho \,\mathrm{d}\rho\mathrm{d}z = 0$$

Spherical Coordinates

The relation between cartesian coordinates $\hat{x}^a = (x, y, z)$ and spherical coordinates $x^i = (\rho, \theta, \phi)$ is:

$$\begin{aligned} x &= \rho \sin\theta \cos\phi \\ y &= \rho \sin\theta \sin\phi \\ z &= \rho \cos\theta \end{aligned} \tag{3.40.2.1}$$

The transformation matrix (Jacobian) is calculated by differentiating (3.40.2.1):

$$\frac{\partial \hat{x}}{\partial x} = \frac{\partial(x, y, z)}{\partial(\rho, \theta, \phi)} = \begin{pmatrix} \sin\theta\cos\phi & \rho\cos\theta\cos\phi & -\rho\sin\theta\sin\phi \\ \sin\theta\sin\phi & \rho\cos\theta\sin\phi & \rho\sin\theta\cos\phi \\ \cos\theta & -\rho\sin\theta & 0 \end{pmatrix} \tag{3.40.2.2}$$

The inverse Jacobian is calculated by inverting the matrix (3.40.2.2):

$$\frac{\partial x}{\partial \hat{x}} = \frac{\partial(\rho, \theta, \phi)}{\partial(x, y, z)} = \begin{pmatrix} \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \\ \frac{\cos\theta\cos\phi}{\rho} & \frac{\cos\theta\sin\phi}{\rho} & -\frac{\sin\theta}{\rho} \\ -\frac{\sin\phi}{\rho\sin\theta} & \frac{\cos\phi}{\rho\sin\theta} & 0 \end{pmatrix}$$

We expressed the above Jacobians using ρ, θ, ϕ and we can use (3.40.2.1) to express them using x, y, z. Code:

```
from sympy import var, sin, cos, zeros, Matrix, simplify, latex
var("rho theta phi")
x_hat = Matrix([
    rho * sin(theta) * cos(phi),
    rho * sin(theta) * sin(phi),
    rho * cos(theta)])
x = Matrix([rho, theta, phi])

M = zeros(3, 3)
for i in range(3):
    for j in range(3):
        M[i, j] = x_hat[i].diff(x[j])

N = M.inv(method="ADJ")
one = sin(phi)**2*sin(theta)**2 + cos(phi)**2*cos(theta)**2 + \
        cos(phi)**2 + cos(theta)**2
one_simple = one.subs(sin(phi)**2, 1-cos(phi)**2).expand().simplify()
N.simplify()
# one_simple is equal to 1, but simplify() can't do this automatically yet:
N = N.subs(one, one_simple)

print "J =", latex(M)
print
print "J^{-1} =", latex(N)
```

Output:

$$J = \begin{bmatrix} \sin(\theta)\cos(\phi) & \rho\cos(\phi)\cos(\theta) & -\rho\sin(\phi)\sin(\theta) \\ \sin(\phi)\sin(\theta) & \rho\sin(\phi)\cos(\theta) & \rho\sin(\theta)\cos(\phi) \\ \cos(\theta) & -\rho\sin(\theta) & 0 \end{bmatrix}$$

$$J^{-1} = \begin{bmatrix} \sin(\theta)\cos(\phi) & \sin(\phi)\sin(\theta) & \cos(\theta) \\ \frac{\cos(\phi)\cos(\theta)}{\rho} & \frac{\sin(\phi)\cos(\theta)}{\rho} & -\frac{\sin(\theta)}{\rho} \\ -\frac{\sin(\phi)}{\rho\sin(\theta)} & \frac{\cos(\phi)}{\rho\sin(\theta)} & 0 \end{bmatrix}$$

The transformation matrices (Jacobians) are then used to convert vectors

$$V_i = \frac{\partial \hat{x}^a}{\partial x^i} \hat{V}_a$$

and tensors

$$T_{ij} = \frac{\partial \hat{x}^a}{\partial x^i} \frac{\partial \hat{x}^b}{\partial x^j} \hat{T}_{ab}$$

between spherical and cartesian coordinates. For example the partial derivatives from cartesian to spherical coordinates transform as:

$$\partial_i = \frac{\partial \hat{x}^a}{\partial x^i} \hat{\partial}_a$$

$$\begin{pmatrix} \partial_\rho & \partial_\theta & \partial_\phi \end{pmatrix} = \begin{pmatrix} \partial_x & \partial_y & \partial_z \end{pmatrix} \begin{pmatrix} \sin\theta\cos\phi & \rho\cos\theta\cos\phi & -\rho\sin\theta\sin\phi \\ \sin\theta\sin\phi & \rho\cos\theta\sin\phi & \rho\sin\theta\cos\phi \\ \cos\theta & -\rho\sin\theta & 0 \end{pmatrix}$$

and from spherical to cartesian as:

$$\hat{\partial}_a = \frac{\partial x^i}{\partial \hat{x}^a} \partial_i$$

$$\begin{pmatrix} \partial_x & \partial_y & \partial_z \end{pmatrix} = \begin{pmatrix} \partial_\rho & \partial_\theta & \partial_\phi \end{pmatrix} \begin{pmatrix} \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \\ \frac{\cos\theta\cos\phi}{\rho} & \frac{\cos\theta\sin\phi}{\rho} & -\frac{\sin\theta}{\rho} \\ -\frac{\sin\phi}{\rho\sin\theta} & \frac{\cos\phi}{\rho\sin\theta} & 0 \end{pmatrix}$$

Care must be taken when rewriting the index expression into matrices – the top index of the Jacobian is the row index, the bottom index is the column index.

The metric tensor of the cartesian coordinate system $\hat{x}^a$ is $\hat{g}_{ab} = \mathrm{diag}(1,1,1)$, so by transformation we get the metric tensor g_{ij} in the spherical coordinates x^i:

$$g_{ij} = \frac{\partial \hat{x}^a}{\partial x^i} \frac{\partial \hat{x}^b}{\partial x^j} \hat{g}_{ab} = \left(\frac{\partial \hat{x}}{\partial x}\right)^T \hat{g} \frac{\partial \hat{x}}{\partial x} =$$

$$= \left(\frac{\partial (x,y,z)}{\partial (\rho,\theta,\phi)}\right)^T \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \frac{\partial (x,y,z)}{\partial (\rho,\theta,\phi)} =$$

$$= \begin{pmatrix} \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \\ \rho\cos\theta\cos\phi & \rho\cos\theta\sin\phi & -\rho\sin\theta \\ -\rho\sin\theta\sin\phi & \rho\sin\theta\cos\phi & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sin\theta\cos\phi & \rho\cos\theta\cos\phi & -\rho\sin\theta\sin\phi \\ \sin\theta\sin\phi & \rho\cos\theta\sin\phi & \rho\sin\theta\cos\phi \\ \cos\theta & -\rho\sin\theta & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \rho^2 & 0 \\ 0 & 0 & \rho^2\sin^2\theta \end{pmatrix}$$

Once we have the metric tensor expressed in spherical coordinates, we don't need the cartesian coordinates anymore. All formulas only contain the spherical coordinates and the metric tensor.

$$g^{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\rho^2} & 0 \\ 0 & 0 & \frac{1}{\rho^2\sin^2\theta} \end{pmatrix}$$

$$\det g = \det g_{ij} = \rho^4 \sin^2\theta$$

$$\nabla^i \nabla_i \varphi = \partial^i \partial_i \varphi + \frac{1}{2 \det g} \partial_j (\det g)\ g^{jk} \partial_k \varphi =$$

$$= g^{ij} \partial_i \partial_j \varphi + \frac{1}{2\rho^4 \sin^2\theta} \left(\partial_\rho (\rho^4 \sin^2\theta)\ g^{\rho\rho} \partial_\rho \varphi + \partial_\theta (\rho^4 \sin^2\theta)\ g^{\theta\theta} \partial_\theta \varphi \right)$$

$$= g^{ij} \partial_i \partial_j \varphi + \frac{2}{\rho} \partial_\rho \varphi + \frac{\cos\theta}{\rho^2 \sin\theta} \partial_\theta \varphi =$$

$$= \partial_\rho \partial_\rho \varphi + \frac{1}{\rho^2} \partial_\theta \partial_\theta \varphi + \frac{1}{\rho^2 \sin^2\theta} \partial_\phi \partial_\phi \varphi + \frac{2}{\rho} \partial_\rho \varphi + \frac{\cos\theta}{\rho^2 \sin\theta} \partial_\theta \varphi$$

Rotating Disk

Let's have a laboratory Euclidean system $x^\mu = (t, x, y, z)$ and a rotating disk system $x'^\mu = (t', x', y', z')$. The relation between the frames is

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\omega t & \sin\omega t & 0 \\ 0 & -\sin\omega t & \cos\omega t & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} t \\ x\cos\omega t + y\sin\omega t \\ -x\sin\omega t + y\cos\omega t \\ z \end{pmatrix}$$

The inverse transformation can be calculated by simply inverting the matrix:

$$\begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\omega t' & -\sin\omega t' & 0 \\ 0 & \sin\omega t' & \cos\omega t' & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix}$$

so the transformation matrices are:

$$\frac{\partial x'^\mu}{\partial x^\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -x\omega\sin\omega t + y\omega\cos\omega t & \cos\omega t & \sin\omega t & 0 \\ -x\omega\cos\omega t - y\omega\sin\omega t & -\sin\omega t & \cos\omega t & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \frac{\partial x'}{\partial x}$$

$$\frac{\partial x^\nu}{\partial x'^\mu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -x'\omega\sin\omega t' - y'\omega\cos\omega t' & \cos\omega t' & -\sin\omega t' & 0 \\ x'\omega\cos\omega t' - y'\omega\sin\omega t' & \sin\omega t' & \cos\omega t' & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \frac{\partial x}{\partial x'}$$

The problem now is that Newtonian mechanics has a degenerated spacetime metrics (see later). Let's pretend we have the following metrics in the x^μ system:

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = g$$

and

$$g'_{\alpha\beta} = \frac{\partial x^\mu}{\partial x'^\alpha}\frac{\partial x^\nu}{\partial x'^\beta} g_{\mu\nu} = \left(\frac{\partial x}{\partial x'}\right)^T g \left(\frac{\partial x}{\partial x'}\right) = \begin{pmatrix} 1+\omega^2(x'^2+y'^2) & -\omega y' & \omega x' & 0 \\ -\omega y' & 1 & 0 & 0 \\ \omega x' & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = g'$$

However, if we calculate with the correct special relativity metrics:

$$g_{\mu\nu} = \begin{pmatrix} -c^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = g$$

and

$$g'_{\alpha\beta} = \frac{\partial x^\mu}{\partial x'^\alpha}\frac{\partial x^\nu}{\partial x'^\beta} g_{\mu\nu} = \left(\frac{\partial x}{\partial x'}\right)^T g \left(\frac{\partial x}{\partial x'}\right) = \begin{pmatrix} -c^2+\omega^2(x'^2+y'^2) & -\omega y' & \omega x' & 0 \\ -\omega y' & 1 & 0 & 0 \\ \omega x' & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = g'$$

We get the same Christoffel symbols as with the $\mathrm{diag}(1,1,1,1)$ metrics, because only the derivatives of the metrics are important. Then the only nonzero Christoffel symbols are

$$\Gamma^1_{00} = -x'\omega^2$$

$$\Gamma^1_{02} = \Gamma^1_{20} = -\omega$$

$$\Gamma^2_{00} = -y'\omega^2$$

$$\Gamma^2_{01} = \Gamma^2_{10} = \omega$$

If we want to avoid dealing with metrics, it is possible to start with the Christoffel symbols in the x^μ system:

$$\Gamma^\sigma_{\mu\nu} = 0$$

and then transforming them to the x'^μ system using the change of variable formula:

$$\Gamma'^\alpha{}_{\beta\gamma} = \frac{\partial x^\mu}{\partial x'^\beta}\frac{\partial x^\nu}{\partial x'^\gamma}\Gamma^\sigma{}_{\mu\nu}\frac{\partial x'^\alpha}{\partial x^\sigma} + \frac{\partial x'^\alpha}{\partial x^\sigma}\frac{\partial^2 x^\sigma}{\partial x'^\beta \partial x'^\gamma} = \frac{\partial x'^\alpha}{\partial x^\sigma}\frac{\partial^2 x^\sigma}{\partial x'^\beta \partial x'^\gamma}$$

As an example, let's calculate the coefficients above:

$$\Gamma'^2{}_{00} = \frac{\partial x'^2}{\partial x^\sigma}\frac{\partial^2 x^\sigma}{\partial x'^0 \partial x'^0} = \frac{\partial x'^2}{\partial x^\sigma}\frac{\partial}{\partial x'^0}\frac{\partial x^\sigma}{\partial x'^0} =$$

$$= \begin{pmatrix} -x\omega\cos\omega t - y\omega\sin\omega t & -\sin\omega t & \cos\omega t & 0 \end{pmatrix}\frac{\partial}{\partial t'}\begin{pmatrix} 1 \\ -x'\omega\sin\omega t' - y'\omega\cos\omega t' \\ x'\omega\cos\omega t' - y'\omega\sin\omega t' \\ 0 \end{pmatrix} =$$

$$= \begin{pmatrix} -x\omega\cos\omega t - y\omega\sin\omega t & -\sin\omega t & \cos\omega t & 0 \end{pmatrix}\begin{pmatrix} 0 \\ -x'\omega^2\cos\omega t' + y'\omega^2\sin\omega t' \\ -x'\omega^2\sin\omega t' - y'\omega^2\cos\omega t' \\ 0 \end{pmatrix} = -y'\omega^2$$

$$\Gamma'^1{}_{00} = -x'\omega^2$$

$$\Gamma'^2{}_{01} = \Gamma'^2{}_{10} = \frac{\partial x'^2}{\partial x^\sigma}\frac{\partial^2 x^\sigma}{\partial x'^0 \partial x'^1} = \frac{\partial x'^2}{\partial x^\sigma}\frac{\partial}{\partial x'^0}\frac{\partial x^\sigma}{\partial x'^1} =$$

$$= \begin{pmatrix} -x\omega\cos\omega t - y\omega\sin\omega t & -\sin\omega t & \cos\omega t & 0 \end{pmatrix}\frac{\partial}{\partial t'}\begin{pmatrix} 0 \\ \cos\omega t' \\ \sin\omega t' \\ 0 \end{pmatrix} =$$

$$= \begin{pmatrix} -x\omega\cos\omega t - y\omega\sin\omega t & -\sin\omega t & \cos\omega t & 0 \end{pmatrix}\begin{pmatrix} 0 \\ -\omega\sin\omega t' \\ \omega\cos\omega t' \\ 0 \end{pmatrix} = \omega$$

$$\Gamma'^1{}_{02} = \Gamma'^1{}_{20} = -\omega$$

So we got the same results.

Now let's see what we have got. Later we'll show, that the Γ^i_{00} coefficients are just $\partial_i\phi$ in the Newtonian theory. E.g. in our case we have:

$$\Gamma'^1_{00} = -x'\omega^2 = \partial'_x\phi$$

$$\Gamma'^2_{00} = -y'\omega^2 = \partial'_y\phi$$

$$\Gamma'^3_{00} = 0 = \partial'_z \phi$$

from which:

$$\phi(t, x, y, z) = -\tfrac{1}{2}(x'^2 + y'^2)\omega^2 + C(t)$$

and the force acting on a test particle is then:

$$\mathbf{F} = -m\nabla\phi = m\left(x', y', 0\right)\omega^2 = m\mathbf{r}'\omega^2$$

where we have defined $\mathbf{r}' = (x', y', 0)$. This is just the centrifugal force. Also observe, that we could have read ϕ directly from the metrics itself — just compare it to the Lorentzian metrics (with gravitation) in the next chapter.

The other two terms (Γ'^1_{02}, Γ'^2_{01} and the symmetric ones) don't behave as a gravitational force, but rather only act when we are differentiating (e.g. only act on moving bodies). Below we show this is just the $-2\boldsymbol{\omega} \times \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}$ term (responsible for the Coriolis acceleration).

Let's write the full equations of geodesics:

$$\frac{\mathrm{d}^2 x^0}{\mathrm{d}\lambda^2} = 0$$

$$\frac{\mathrm{d}^2 x^1}{\mathrm{d}\lambda^2} + \Gamma^1_{00}\left(\frac{\mathrm{d}x^0}{\mathrm{d}\lambda}\right)^2 + 2\Gamma^1_{20}\frac{\mathrm{d}x^2}{\mathrm{d}\lambda}\frac{\mathrm{d}x^0}{\mathrm{d}\lambda} = 0$$

$$\frac{\mathrm{d}^2 x^2}{\mathrm{d}\lambda^2} + \Gamma^2_{00}\left(\frac{\mathrm{d}x^0}{\mathrm{d}\lambda}\right)^2 + 2\Gamma^2_{10}\frac{\mathrm{d}x^1}{\mathrm{d}\lambda}\frac{\mathrm{d}x^0}{\mathrm{d}\lambda} = 0$$

$$\frac{\mathrm{d}^2 x^3}{\mathrm{d}\lambda^2} = 0$$

This becomes:

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = x\omega^2 + 2\omega\frac{\mathrm{d}y}{\mathrm{d}t}$$

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} = y\omega^2 - 2\omega\frac{\mathrm{d}x}{\mathrm{d}t}$$

$$\frac{\mathrm{d}^2 z}{\mathrm{d}t^2} = 0$$

we can define $\mathbf{r} = (x, y, 0)$ and $\boldsymbol{\omega} = (0, 0, \omega)$. Then the above equations can be rewritten as:

$$\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} = \mathbf{r}\omega^2 - 2\boldsymbol{\omega} \times \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}$$

So we get two fictituous forces, the centrifugal force and the Coriolis force.

Now imagine a static vector in the x^μ system along the x axis, i.e.

$$V^\mu = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = V$$

then

$$V'^\mu = \frac{\partial x'^\mu}{\partial x^\alpha}V^\alpha = \frac{\partial x'}{\partial x}V = \begin{pmatrix} 1 \\ -x\omega\sin\omega t + y\omega\cos\omega t + \cos\omega t \\ -x\omega\cos\omega t - y\omega\sin\omega t - \sin\omega t \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ y'\omega + \cos\omega t' \\ -x'\omega - \sin\omega t' \\ 0 \end{pmatrix} = V'$$

In the last equality we transformed from x^μ to x'^μ using the relation between frames.

Differentiating any vector in the x^μ coordinates is easy – it's just a partial derivative (due to the Euclidean metrics). Let's differentiate any vector in the x'^μ coordinates with respect to time (since $t = t'$, the time is the same in both coordinate systems):

$$\nabla_0 V'^\mu = \partial_0 V'^\mu + \Gamma^\mu_{0\alpha} V'^\alpha$$

$$\nabla_0 \begin{pmatrix} V'^0 \\ V'^1 \\ V'^2 \\ V'^3 \end{pmatrix} = \begin{pmatrix} \partial_0 V'^0 \\ \partial_0 V'^1 + \Gamma^1_{00} V'^0 + \Gamma^1_{02} V'^2 \\ \partial_0 V'^2 + \Gamma^2_{00} V'^0 + \Gamma^2_{01} V'^1 \\ \partial_0 V'^3 \end{pmatrix} = \begin{pmatrix} \partial_0 V'^0 \\ \partial_0 V'^1 - x'\omega^2 V'^0 - \omega V'^2 \\ \partial_0 V'^2 - y'\omega^2 V'^0 + \omega V'^1 \\ \partial_0 V'^3 \end{pmatrix} =$$

$$= \partial_0 \begin{pmatrix} V'^0 \\ V'^1 \\ V'^2 \\ V'^3 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ -x'\omega^2 & 0 & -\omega & 0 \\ -y'\omega^2 & \omega & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} V'^0 \\ V'^1 \\ V'^2 \\ V'^3 \end{pmatrix} \tag{3.40.2.3}$$

For our particular (static) vector this yields:

$$\nabla_0 \begin{pmatrix} 1 \\ y'\omega + \cos\omega t' \\ -x'\omega - \sin\omega t' \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

as expected, because it was at rest in the x^μ system. Let's imagine a static vector in the x'^μ system along the x' axis, i.e.

$$W'^\mu = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

$$W^\mu = \frac{\partial x^\mu}{\partial x'^\alpha} W'^\alpha = \begin{pmatrix} 1 \\ -x'\omega\sin\omega t' - y'\omega\cos\omega t' + \cos\omega t' \\ x'\omega\cos\omega t' - y'\omega\sin\omega t' + \sin\omega t' \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -y\omega + \cos\omega t \\ x\omega + \sin\omega t \\ 0 \end{pmatrix}$$

then

$$\nabla_0 W'^\mu = \nabla_0 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -x'\omega^2 \\ -y'\omega^2 + \omega \\ 0 \end{pmatrix}$$

$$\nabla_0 W^\mu = \partial_0 \begin{pmatrix} 1 \\ -y\omega + \cos\omega t \\ x\omega + \sin\omega t \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -\omega\sin\omega t \\ \omega\cos\omega t \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -\omega & 0 \\ 0 & \omega & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \cos\omega t \\ \sin\omega t \\ 0 \end{pmatrix} = \boldsymbol{\omega} \times \mathbf{W}$$

Similarly

$$\nabla_0 \nabla_0 W'^\mu == \begin{pmatrix} 0 \\ -y'\omega^3 - \omega^2 \\ -x'\omega^3 \\ 0 \end{pmatrix}$$

$$\nabla_0 \nabla_0 W^\mu == \begin{pmatrix} 0 \\ -\omega^2\cos\omega t \\ -\omega^2\sin\omega t \\ 0 \end{pmatrix}$$

How can one prove the relation:

$$\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t} = \boldsymbol{\omega} \times \mathbf{A} + \frac{\mathrm{d}'\mathbf{A}}{\mathrm{d}t} \tag{3.40.2.4}$$

that is used for example to derive the Coriolis acceleration etc.? We need to write it components to understand what it really means:

$$\nabla_0 \begin{pmatrix} A'^0 \\ A'^1 \\ A'^2 \\ A'^3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -\omega & 0 \\ 0 & \omega & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} A'^0 \\ A'^1 \\ A'^2 \\ A'^3 \end{pmatrix} + \partial_0 \begin{pmatrix} A'^0 \\ A'^1 \\ A'^2 \\ A'^3 \end{pmatrix}$$

Comparing to the covariant derivative above, it's clear that they are equal (provided that $x' = 0$ and $y' = 0$, i.e. we are at the center of rotation).

Let's show the derivation by Goldstein. The change in a time dt of a general vector $\mathbf{G}$ as seen by an observer in the body system of axes will differ from the corresponding change as seen by an observer in the space system:

$$(d\mathbf{G})_{\text{space}} = (d\mathbf{G})_{\text{body}} + (d\mathbf{G})_{\text{rot}}$$

Now consider a vector fixed in the rigid body. Then $(d\mathbf{G})_{\text{body}} = 0$ and

$$(d\mathbf{G})_{\text{rot}} = (d\mathbf{G})_{\text{space}} = d\boldsymbol{\Omega} \times \mathbf{G}$$

For an arbitrary vector, the change relative to the space axes is the sum of the two effects:

$$(d\mathbf{G})_{\text{space}} = (d\mathbf{G})_{\text{body}} + d\boldsymbol{\Omega} \times \mathbf{G}$$

A more rigorous derivation of the last equation follows from:

$$G_i = a_{ji} G'_j$$

$$dG_i = a_{ji} dG'_j + da_{ji} G'_j$$

Let's make the space and body instantaneously coincident at time t, then $a_{ji} = \delta_{ji}$ and $da_{ji} = -\epsilon_{ijk} d\Omega_k = \epsilon_{ikj} d\Omega_k$, so we get the same equation as earlier:

$$dG_i = dG'_i + \epsilon_{ikj} d\Omega_k G'_j$$

Anyhow, introducing $\boldsymbol{\omega}$ by:

$$\boldsymbol{\omega} = \frac{d\boldsymbol{\Omega}}{dt}$$

we get

$$\left(\frac{d\mathbf{G}}{dt}\right)_{\text{space}} = \left(\frac{d\mathbf{G}}{dt}\right)_{\text{body}} + \boldsymbol{\omega} \times \mathbf{G}$$

Linear Elasticity Equations in Cylindrical Coordinates

Authors: Pavel Solin & Lenka Dubcova

In this paper we derive the weak formulation of linear elasticity equations suitable for the finite element discretization of axisymmetric 3D problems.

Original equations in Cartesian coordinates

Let's start with some notations: By $\boldsymbol{u} = (u_1, u_2, u_3)^T$ we denote the displacement vector in 3D Cartesian coordinates, and by ϵ the tensor of small deformations,

$$\epsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \quad 1 \le i, j \le 3.$$

The stress tensor σ has the form

$$\sigma_{ij} = \lambda\delta_{ij}\mathrm{div}\boldsymbol{u} + 2\mu\epsilon_{ij}, \quad 1 \le i, j \le 3, \tag{3.40.2.5}$$

where

$$\mathrm{div}\boldsymbol{u} = \sum_{k=1}^{3}\frac{\partial u_k}{\partial x_k} = \sum_{k=1}^{3}\epsilon_{kk} = \mathrm{Tr}(\epsilon).$$

The symbols λ and μ are the Lam'e constants and δ_{ij} is the Kronecker symbol ($\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise). The equilibrium equations have the form

$$\sum_{j=1}^{3}\frac{\partial \sigma_{ij}}{\partial x_j} + f_i = 0, \quad 1 \le i \le 3, \tag{3.40.2.6}$$

where $(f_1, f_2, f_3)^T$ is the vector of internal forces (such as gravity).

The boundary conditions for linear elasticity are given by

$$\begin{aligned} u_i &= \hat{u}_i \quad \text{on } \Gamma_1 \\ \sum_{j=1}^{3}\sigma_{ij}n_j &= g_i \quad \text{on } \Gamma_2, \end{aligned}$$

where g_i are surface forces.

Weak formulation

Multiplying by test functions and integrating over the domain Ω we obtain

$$-\int_\Omega \sum_{j=1}^{3}\frac{\partial \sigma_{ij}}{\partial x_j}v_i = \int_\Omega f_i\, v_i, \quad 1 \le i \le 3. \tag{3.40.2.7}$$

Using Green's theorem and the boundary conditions

$$\int_\Omega \sum_{j=1}^{3}\sigma_{ij}\frac{\partial v_i}{\partial x_j} - \int_{\partial\Omega}\sum_{j=1}^{3}\sigma_{ij}n_j v_i = \int_\Omega f_i\, v_i, \quad 1 \le i \le 3.$$

Thus

$$\int_\Omega \sum_{j=1}^{3}\sigma_{ij}\frac{\partial v_i}{\partial x_j} - \int_{\Gamma_2} g_i v_i = \int_\Omega f_i\, v_i, \quad 1 \le i \le 3. \tag{3.40.2.8}$$

Let us write the equations (3.40.2.8) in detail using relation (3.40.2.5)

$$\begin{aligned}
\int_\Omega \left[\lambda\mathrm{div}u + 2\mu\frac{\partial u_1}{\partial x_1}\right]\frac{\partial v_1}{\partial x_1} + \mu\left(\frac{\partial u_1}{\partial x_2} + \frac{\partial u_2}{\partial x_1}\right)\frac{\partial v_1}{\partial x_2} + \mu\left(\frac{\partial u_1}{\partial x_3} + \frac{\partial u_3}{\partial x_1}\right)\frac{\partial v_1}{\partial x_3} - \int_{\Gamma_2} g_1 v_1 &= \int_\Omega f_1\, v_1, \\
\int_\Omega \mu\left(\frac{\partial u_1}{\partial x_2} + \frac{\partial u_2}{\partial x_1}\right)\frac{\partial v_2}{\partial x_1} + \left[\lambda\mathrm{div}u + 2\mu\frac{\partial u_2}{\partial x_2}\right]\frac{\partial v_2}{\partial x_2} + \mu\left(\frac{\partial u_2}{\partial x_3} + \frac{\partial u_3}{\partial x_2}\right)\frac{\partial v_2}{\partial x_3} - \int_{\Gamma_2} g_2 v_2 &= \int_\Omega f_2\, v_2, \\
\int_\Omega \mu\left(\frac{\partial u_1}{\partial x_3} + \frac{\partial u_3}{\partial x_1}\right)\frac{\partial v_3}{\partial x_1} + \mu\left(\frac{\partial u_2}{\partial x_3} + \frac{\partial u_3}{\partial x_2}\right)\frac{\partial v_3}{\partial x_2} + \left[\lambda\mathrm{div}u + 2\mu\frac{\partial u_3}{\partial x_3}\right]\frac{\partial v_3}{\partial x_3} - \int_{\Gamma_2} g_3 v_3 &= \int_\Omega f_3\, v_3.
\end{aligned}$$

Elementary transformation relations

First let us show how the partial derivatives of a scalar function g are transformed from Cartesian coordinates x_1, x_2, x_3 to cylindrical coordinates r, ϕ, z. Note that

$$x_1(r,\phi) = r\cos\phi, \quad x_2(r,\phi) = r\sin\phi, \quad x_3(z) = z.$$

Since

$$g(x_1, x_2, x_3) = g(x_1(r,\phi), x_2(r,\phi), x_3(z)),$$

it is

$$\begin{aligned}
\frac{\partial g}{\partial r} &= \frac{\partial g}{\partial x_1}\cos\phi + \frac{\partial g}{\partial x_2}\sin\phi, \\
\frac{\partial g}{\partial \phi} &= \frac{\partial g}{\partial x_1}(-r\sin\phi) + \frac{\partial g}{\partial x_2} r\cos\phi, \\
\frac{\partial g}{\partial z} &= \frac{\partial g}{\partial x_3}.
\end{aligned}$$

From here we obtain

$$\begin{aligned}
\frac{\partial g}{\partial x_1} &= \frac{\partial g}{\partial r}\cos\phi - \frac{1}{r}\frac{\partial g}{\partial \phi}\sin\phi, \\
\frac{\partial g}{\partial x_2} &= \frac{\partial g}{\partial r}\sin\phi + \frac{1}{r}\frac{\partial g}{\partial \phi}\cos\phi, \\
\frac{\partial g}{\partial x_3} &= \frac{\partial g}{\partial z}.
\end{aligned}$$

The relations between displacement components in Cartesian and cylindrical coordinates are

$$\begin{aligned}
u_1 &= u_r\cos\phi, \\
u_2 &= u_r\sin\phi, \\
u_3 &= u_z.
\end{aligned}$$

The same relations hold for surface forces g_i and volume forces f_i.

Applying (3.40.2) to u_1, we obtain

$$\begin{aligned}
\frac{\partial u_1}{\partial x_1} &= \frac{\partial u_1}{\partial r}\cos\phi - \frac{1}{r}\frac{\partial u_1}{\partial \phi}\sin\phi, \\
\frac{\partial u_1}{\partial x_2} &= \frac{\partial u_1}{\partial r}\sin\phi + \frac{1}{r}\frac{\partial u_1}{\partial \phi}\cos\phi, \\
\frac{\partial u_1}{\partial x_3} &= \frac{\partial u_1}{\partial z}.
\end{aligned}$$

Using (3.40.2) and the fact that u_r does not depend on ϕ, this yields

$$\begin{aligned}
\frac{\partial u_1}{\partial x_1} &= \frac{\partial u_r}{\partial r}\cos^2\phi + \frac{1}{r}u_r\sin^2\phi, \\
\frac{\partial u_1}{\partial x_2} &= \frac{\partial u_r}{\partial r}\cos\phi\sin\phi - \frac{1}{r}u_r\cos\phi\sin\phi, \\
\frac{\partial u_1}{\partial x_3} &= \frac{\partial u_r}{\partial z}\cos\phi.
\end{aligned}$$

Analogously, for u_2 we calculate

$$\begin{aligned}\frac{\partial u_2}{\partial x_1} &= \frac{\partial u_r}{\partial r}\cos\phi\sin\phi - \frac{1}{r}u_r\cos\phi\sin\phi,\\ \frac{\partial u_2}{\partial x_2} &= \frac{\partial u_r}{\partial r}\sin^2\phi + \frac{1}{r}u_r\cos^2\phi,\\ \frac{\partial u_2}{\partial x_3} &= \frac{\partial u_r}{\partial z}\sin\phi.\end{aligned}$$

For u_3, using that it does not depend on ϕ, we have

$$\begin{aligned}\frac{\partial u_3}{\partial x_1} &= \frac{\partial u_z}{\partial r}\cos\phi,\\ \frac{\partial u_3}{\partial x_2} &= \frac{\partial u_z}{\partial r}\sin\phi,\\ \frac{\partial u_3}{\partial x_3} &= \frac{\partial u_z}{\partial z}.\end{aligned}$$

For further reference, transform also div$u$ into cylindrical coordinates

$$\begin{aligned}\text{div}u &= \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} =\\ &= \frac{\partial u_r}{\partial r}\cos^2\phi + \frac{1}{r}u_r\sin^2\phi + \frac{\partial u_r}{\partial r}\sin^2\phi + \frac{1}{r}u_r\cos^2\phi + \frac{\partial u_z}{\partial z} =\\ &= \frac{\partial u_r}{\partial r} + \frac{1}{r}u_r + \frac{\partial u_z}{\partial z}\end{aligned}$$

Axisymmetric formulation

Assuming that the domain Ω is axisymmetric, we can begin to transform the integrals in (3.40.2) to cylindrical coordinates. Recall that the Jacobian of the transformation is $J(r,\phi,z) = r$. The first equation in (3.40.2) has the form:

$$\begin{aligned}&\int_\Omega r\left[\lambda(\frac{\partial u_r}{\partial r} + \frac{1}{r}u_r + \frac{\partial u_z}{\partial z}) + 2\mu(\frac{\partial u_r}{\partial r}\cos^2\phi + \frac{1}{r}u_r\sin^2\phi)\right](\frac{\partial v_r}{\partial r}\cos^2\phi + \frac{1}{r}v_r\sin^2\phi) +\\ &r2\mu\left(\frac{\partial u_r}{\partial r}\cos\phi\sin\phi - \frac{1}{r}u_r\cos\phi\sin\phi\right)\left(\frac{\partial v_r}{\partial r}\cos\phi\sin\phi - \frac{1}{r}v_r\cos\phi\sin\phi\right) +\\ &r\mu\left(\frac{\partial u_r}{\partial z}\cos\phi + \frac{\partial u_z}{\partial r}\cos\phi\right)\frac{\partial v_r}{\partial z}\cos\phi - \int_{\Gamma_2} r g_r v_r\cos^2\phi = \int_\Omega r f_r\, v_r\cos^2\phi,\end{aligned}$$

The second equation in (3.40.2) has the form:

$$\begin{aligned}&\int_\Omega r2\mu\left(\frac{\partial u_r}{\partial r}\cos\phi\sin\phi - \frac{1}{r}u_r\cos\phi\sin\phi\right)\left(\frac{\partial v_r}{\partial r}\cos\phi\sin\phi - \frac{1}{r}v_r\cos\phi\sin\phi\right) +\\ &r\left[\lambda(\frac{\partial u_r}{\partial r} + \frac{1}{r}u_r + \frac{\partial u_z}{\partial z}) + 2\mu(\frac{\partial u_r}{\partial r}\sin^2\phi + \frac{1}{r}u_r\cos^2\phi)\right](\frac{\partial v_r}{\partial r}\sin^2\phi + \frac{1}{r}v_r\cos^2\phi) +\\ &r\mu\left(\frac{\partial u_r}{\partial z}\sin\phi + \frac{\partial u_z}{\partial r}\sin\phi\right)(\frac{\partial v_r}{\partial z}\sin\phi) - \int_{\Gamma_2} r g_r v_r\sin^2\phi = \int_\Omega r f_r\, v_r\sin^2\phi,\end{aligned}$$

Adding these two equations together we get

$$\int_\Omega r\lambda(\frac{\partial u_r}{\partial r}+\frac{1}{r}u_r+\frac{\partial u_z}{\partial z})(\frac{\partial v_r}{\partial r}+\frac{1}{r}v_r)+$$
$$\int_\Omega r\mu\left[2\left(\frac{\partial u_r}{\partial r}\frac{\partial v_r}{\partial r}\cos^4\phi+\frac{1}{r}u_r\frac{\partial v_r}{\partial r}\sin^2\phi\cos^2\phi+\frac{1}{r}\frac{\partial u_r}{\partial r}v_r\sin^2\phi\cos^2\phi+\frac{1}{r^2}u_rv_r\sin^4\phi\right)+\right.$$
$$2\left(\frac{\partial u_r}{\partial r}\frac{\partial v_r}{\partial r}\sin^4\phi+\frac{1}{r}u_r\frac{\partial v_r}{\partial r}\sin^2\phi\cos^2\phi+\frac{1}{r}\frac{\partial u_r}{\partial r}v_r\sin^2\phi\cos^2\phi+\frac{1}{r^2}u_rv_r\cos^4\phi\right)+$$
$$4\left(\frac{\partial u_r}{\partial r}\frac{\partial v_r}{\partial r}\cos^2\phi\sin^2\phi-\frac{1}{r}u_r\frac{\partial v_r}{\partial r}\cos^2\phi\sin^2\phi-\frac{1}{r}\frac{\partial u_r}{\partial r}v_r\cos^2\phi\sin^2\phi+\frac{1}{r^2}u_rv_r\cos^2\phi\sin^2\phi\right)+$$
$$\left.\left(\frac{\partial u_r}{\partial z}\frac{\partial v_r}{\partial z}+\frac{\partial u_z}{\partial r}\frac{\partial v_r}{\partial z}\right)\right]-\int_{\Gamma_2}g_rv_rr=\int_\Omega f_r\,v_rr$$

This can be simplified to

$$\int_\Omega r\lambda(\frac{\partial u_r}{\partial r}+\frac{1}{r}u_r+\frac{\partial u_z}{\partial z})(\frac{\partial v_r}{\partial r}+\frac{1}{r}v_r)+\int_\Omega r\mu\left[2\left(\frac{\partial u_r}{\partial r}\frac{\partial v_r}{\partial r}+\frac{1}{r^2}u_rv_r\right)+\left(\frac{\partial u_r}{\partial z}\frac{\partial v_r}{\partial z}+\frac{\partial u_z}{\partial r}\frac{\partial v_r}{\partial z}\right)\right]$$
$$-\int_{\Gamma_2}g_rv_rr=\int_\Omega f_r\,v_rr$$

Finally, the third equation in (3.40.2) has the form

$$\int_\Omega r\mu\left(\frac{\partial u_r}{\partial z}\cos\phi+\frac{\partial u_z}{\partial r}\cos\phi\right)\frac{\partial v_z}{\partial r}\cos\phi+r\mu\left(\frac{\partial u_r}{\partial z}\sin\phi+\frac{\partial u_z}{\partial r}\sin\phi\right)\frac{\partial v_z}{\partial r}\sin\phi+$$
$$r\left[\lambda(\frac{\partial u_r}{\partial r}+\frac{1}{r}u_r+\frac{\partial u_z}{\partial z})+2\mu\frac{\partial u_z}{\partial z}\right]\frac{\partial v_z}{\partial z}-\int_{\Gamma_2}g_zv_zr=\int_\Omega f_z\,v_zr.$$

This gives us

$$\int_\Omega r\mu\left(\frac{\partial u_r}{\partial z}\frac{\partial v_z}{\partial r}+\frac{\partial u_z}{\partial r}\frac{\partial v_z}{\partial r}+2\frac{\partial u_z}{\partial z}\frac{\partial v_z}{\partial z}\right)+r\lambda\left(\frac{\partial u_r}{\partial r}+\frac{1}{r}u_r+\frac{\partial u_z}{\partial z}\right)\frac{\partial v_z}{\partial z}-\int_{\Gamma_2}g_zv_zr=\int_\Omega f_z\,v_zr.$$

Since the integrands do not depend on ϕ, we can simplify this to integral over Ω_0, where Ω_0 is the intersection of the domain Ω with the $x_1^+x_3$ half-plane. Dividing both equations by 2π we get

$$\int_{\Omega_0} r\lambda(\frac{\partial u_r}{\partial r}+\frac{1}{r}u_r+\frac{\partial u_z}{\partial z})(\frac{\partial v_r}{\partial r}+\frac{1}{r}v_r)+\int_{\Omega_0} r\mu\left[2\left(\frac{\partial u_r}{\partial r}\frac{\partial v_r}{\partial r}+\frac{1}{r^2}u_rv_r\right)+\left(\frac{\partial u_r}{\partial z}\frac{\partial v_r}{\partial z}+\frac{\partial u_z}{\partial r}\frac{\partial v_r}{\partial z}\right)\right]$$
$$-\int_{\Gamma_2}g_rv_rr=\int_{\Omega_0} f_r\,v_rr$$
$$\int_{\Omega_0} r\mu\left(\frac{\partial u_r}{\partial z}\frac{\partial v_z}{\partial r}+\frac{\partial u_z}{\partial r}\frac{\partial v_z}{\partial r}+2\frac{\partial u_z}{\partial z}\frac{\partial v_z}{\partial z}\right)+r\lambda\left(\frac{\partial u_r}{\partial r}+\frac{1}{r}u_r+\frac{\partial u_z}{\partial z}\right)\frac{\partial v_z}{\partial z}-\int_{\Gamma_2}g_zv_zr=\int_{\Omega_0} f_z\,v_zr.$$

Coordinate Independent Way

Let's write the elasticity equations in the cartesian coordinates again:

$$\sigma_{ij}=\lambda\delta_{ij}\partial_ku^k+\mu(\partial_ju_i+\partial_iu_j)$$
$$\partial_j\sigma^{ij}+f^i=0$$

Those only work in the cartesian coordinates, so we first write them in a coordinate independent way:

$$\sigma^{ij}=\lambda g^{ij}\nabla_ku^k+\mu(\nabla^ju^i+\nabla^iu^j)$$
$$\nabla_j\sigma^{ij}+f^i=0$$

so:

$$\nabla_j \left(\lambda g^{ij} \nabla_k u^k + \mu(\nabla^j u^i + \nabla^i u^j)\right) + f^i = 0$$

The weak formulation is then (do not sum over i):

$$-\int \nabla_j \left(\lambda g^{ij} \nabla_k u^k + \mu(\nabla^j u^i + \nabla^i u^j)\right) v^i \sqrt{|g|} \mathrm{d}^3x = \int f^i v^i \sqrt{|g|} \mathrm{d}^3x$$

We apply the integration by parts:

$$\int \left(\lambda g^{ij} \nabla_k u^k + \mu(\nabla^j u^i + \nabla^i u^j)\right) \nabla_j v^i \sqrt{|g|} \mathrm{d}^3x = \int f^i v^i \sqrt{|g|} \mathrm{d}^3x$$

This is the weak formulation valid in any coordinates. Using the cylindrical coordinates (see above) we get:

$$\begin{aligned}
\mathbf{x} &= (\rho, \phi, z) \\
\mathrm{d}^3x &= \mathrm{d}\rho\, \mathrm{d}\phi\, \mathrm{d}z \\
g^{ij} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\rho^2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
\sqrt{|g|} &= \sqrt{|\det g_{ij}|} = \rho \\
\nabla_k u^k &= \frac{1}{\sqrt{|g|}} \partial_k (\sqrt{|g|} u^k) = \frac{1}{\rho} \partial_k (\rho u^k) = \\
&= \frac{1}{\rho} u^\rho + \partial_\rho u^\rho + \partial_\phi u^\phi + \partial_z u^z \\
(\nabla^j u^z + \nabla^z u^j) \nabla_j v^z = (g^{jk} \nabla_k u^z + g^{zk} \nabla_k u^j) \nabla_j v^z &= (\partial_\rho u^z + \partial_z u^\rho) \partial_\rho v^z + (\partial_z u^z + \partial_z u^z) \partial_z v^z = \\
&= (\partial_\rho u^z + \partial_z u^\rho) \partial_\rho v^z + 2 \partial_z u^z \partial_z v^z \\
g^{\rho j} \nabla_j v^\rho = g^{\rho\rho} \nabla_\rho v^\rho &= \partial_\rho v^\rho + \Gamma^\rho_{k\rho} v^k = \partial_\rho v^\rho + \frac{1}{\rho} v^\phi \\
g^{\phi j} \nabla_j v^\phi = g^{\phi\phi} \nabla_\phi v^\phi &= \frac{1}{\rho^2} (\partial_\phi v^\phi + \Gamma^\phi_{k\phi} v^k) = \frac{1}{\rho^2} (\partial_\phi v^\phi + \frac{1}{\rho} v^\rho) = \\
g^{zj} \nabla_j v^z = g^{zz} \nabla_z v^z &= \partial_z v^z + \Gamma^z_{kz} v^k = \partial_z v^z \\
\int \left(\lambda g^{ij} \left(\frac{1}{\rho} u^\rho + \partial_\rho u^\rho + \partial_\phi u^\phi + \partial_z u^z\right) + \mu(\nabla^j u^i + \nabla^i u^j)\right) \nabla_j v^i \rho\, \mathrm{d}\rho\, \mathrm{d}\phi\, \mathrm{d}z &= \int f^i v^i \rho\, \mathrm{d}\rho\, \mathrm{d}\phi\, \mathrm{d}z
\end{aligned}$$

for $i = 1, 2, 3$ we get:

$$\begin{aligned}
\int \lambda \left(\frac{1}{\rho} u^\rho + \partial_\rho u^\rho + \partial_\phi u^\phi + \partial_z u^z\right) \left(\partial_\rho v^\rho + \frac{1}{\rho} v^\phi\right) \rho + \mu \left(2 \partial_\rho u^\rho \partial_\rho v^\rho + (\partial_z u^\rho + \partial_\rho u^z) \partial_z v^\rho\right) \rho\, \mathrm{d}\rho\, \mathrm{d}\phi\, \mathrm{d}z &= \int f^\rho v^\rho \rho\, \mathrm{d}\rho\, \mathrm{d}\phi\, \mathrm{d}z \\
\int \lambda \left(\frac{1}{\rho} u^\rho + \partial_\rho u^\rho + \partial_\phi u^\phi + \partial_z u^z\right) \frac{1}{\rho^2} \left(\partial_\phi v^\phi + \frac{1}{\rho} v^\rho\right) \rho + \mu \left(2 \partial_\rho u^\rho \partial_\rho v^\rho + (\partial_z u^\rho + \partial_\rho u^z) \partial_z v^\rho\right) \rho\, \mathrm{d}\rho\, \mathrm{d}\phi\, \mathrm{d}z &= \int f^\phi v^\phi \rho\, \mathrm{d}\rho\, \mathrm{d}\phi\, \mathrm{d}z \\
\int \lambda \left(\frac{1}{\rho} u^\rho + \partial_\rho u^\rho + \partial_\phi u^\phi + \partial_z u^z\right) \partial_z v^z \rho + \mu \left((\partial_\rho u^z + \partial_z u^\rho) \partial_\rho v^z + 2 \partial_z u^z \partial_z v^z\right) \rho\, \mathrm{d}\rho\, \mathrm{d}\phi\, \mathrm{d}z &= \int f^z v^z \rho\, \mathrm{d}\rho\, \mathrm{d}\phi\, \mathrm{d}z
\end{aligned}$$

3.41 Operators

3.41.1 Introduction

The domain of the operator A is $D(A)$, a subspace of the Hilbert space $\mathscr{H}$. Linear operator is:

$$A(\alpha |u\rangle + \beta |v\rangle) = \alpha A |u\rangle + \beta A |v\rangle$$

for all $|u\rangle, |v\rangle \in D(A)$. Symmetric operator is:

$$\langle u|Av\rangle = \langle Au|v\rangle$$

for all $|u\rangle, |v\rangle \in D(A)$ dense in $\mathscr{H}$. If $D(A)$ is dense in $\mathscr{H}$, then the adjoint operator $A^\dagger$ is defined by

$$\langle u|A^\dagger v\rangle = \langle Au|v\rangle$$

for all $|u\rangle \in D(A)$. The domain $D(A^\dagger)$ is given by all $|v\rangle$ for which the above relation holds. It can be shown that $D(A) \subset D(A^\dagger)$.

Operator A is self-adjoint if $A = A^\dagger$. Symmetric operator is self-adjoint only if $D(A) = D(A^\dagger)$. (Bounded symmetric operator is always self-adjoint.) Hermitean operator is a bounded symmetric operator.

Hermitian implies self-adjoint implies symmetric, but all converse implications are false. Below, we need the operator to be self-adjoint (we assume unbounded by default).

3.41.2 Spectrum

To obtain a spectrum of the operator A, we need to solve the following problem:

$$A|\lambda\rangle = \lambda|\lambda\rangle$$

Those values of λ for which the solution $|\lambda\rangle \in \mathscr{H}$ belong to the discrete part of the spectrum. λ are called eigenvalues and $|\lambda\rangle$ eigenvectors. Those values of λ for which $|\lambda\rangle$ can be normalized to a delta function:

$$\langle\lambda|\kappa\rangle = \delta(\lambda - \kappa)$$

belong to the continuous part of the spectrum (note that in this case $|\lambda\rangle \notin \mathscr{H}$).

Eigenvectors belonging to the continous part of the spectrum obey the completeness relation:

$$\int |\lambda\rangle\langle\lambda| \,\mathrm{d}\lambda = \mathbb{1}$$

Eigenvectors belonging to the discrete part obey the following completeness relation:

$$\sum_\lambda |\lambda\rangle\langle\lambda| \,\mathrm{d}\lambda = \mathbb{1}$$

The sum or integral runs over the whole spectrum (if the spectrum contains both discrete and continous part, we simply combine sums and integrals).

Spectrum of a self-adjoint operator is real, because

$$\langle\lambda|A|\lambda\rangle = \lambda\langle\lambda|\lambda\rangle = \lambda^*\langle\lambda|\lambda\rangle$$

The eigenvectors are orthogonal:

$$\begin{aligned}\langle\lambda|A|\kappa\rangle = \kappa\langle\lambda|\kappa\rangle &= \lambda\langle\lambda|\kappa\rangle \\ (\kappa - \lambda)\langle\lambda|\kappa\rangle &= 0\end{aligned}$$

So for $\kappa \neq \lambda$ we get $\langle\lambda|\kappa\rangle = 0$, for $\kappa = \lambda$ the $\langle\lambda|\lambda\rangle$ is equal to 1 if λ belongs to the discrete spectrum and we get:

$$\langle\lambda|\kappa\rangle = \delta_{\lambda\kappa}$$

or it is normalized as a delta function if it belongs to the continous part:

$$\langle\lambda|\kappa\rangle = \delta(\lambda - \kappa)$$

As such, eigenvectors of a self-adjoint operator are complete and orthogonal in the above sense. Thus any function from the space can then be expanded into the series:

$$f(x) = \langle x|f\rangle = \sum_{\lambda} \langle x|\lambda\rangle \langle \lambda|f\rangle$$

where $\langle x|\lambda\rangle$ are the eigenvectors and the coefficients $\langle \lambda|f\rangle$ are given by:

$$\langle \lambda|f\rangle = \int \langle \lambda|x\rangle \langle x|f\rangle \,\mathrm{d}x = \int \langle \lambda|x\rangle \, f(x)\mathrm{d}x$$

The sum over λ runs over the whole spectrum (i.e. it becomes an integral over the continuos parts). Also the coefficients $\langle \lambda|f\rangle$ are either discrete or continous depending on the part of the spectrum. The series converges in the norm, i.e. the following norm goes to zero as we sum over λ:

$$\left\| f(x) - \sum_{\lambda} \langle x|\lambda\rangle \langle \lambda|f\rangle \right\| \to 0$$

3.41.3 Derivative Operator

We have the eigenvalue problem

$$Au = \lambda u$$

where

$$A = -i\frac{\mathrm{d}}{\mathrm{d}x}$$

The operator A is unbounded. A is self-adjoint if:

$$\int_a^b u^*(x)Av(x)\mathrm{d}x = \int_a^b (Au(x))^* v(x)\mathrm{d}x$$

So

$$\begin{aligned}
\int_a^b u^*(x)Av(x)\mathrm{d}x &= \int_a^b u^*(x)\left(-i\frac{\mathrm{d}}{\mathrm{d}x}\right)v(x)\mathrm{d}x = \\
&= \int_a^b \left(i\frac{\mathrm{d}}{\mathrm{d}x}u^*(x)\right)v(x)\mathrm{d}x - i[u^*(x)v(x)]_a^b = \\
&= \int_a^b \left(-i\frac{\mathrm{d}}{\mathrm{d}x}u(x)\right)^* v(x)\mathrm{d}x - i[u^*(x)v(x)]_a^b = \\
&= \int_a^b (Au(x))^* v(x)\mathrm{d}x - i[u^*(x)v(x)]_a^b
\end{aligned}$$

The operator is self-adjoint if and only if $[u^*(x)v(x)]_a^b = 0$. Few boundary conditions that satisfy this condition:

- Dirichlet boundary conditions

$$u(a) = 0, \quad u(b) = 0$$

- Periodic boundary conditions

$$u(a) = u(b)$$

- Antiperiodic boundary conditions

$$u(a) = -u(b)$$

Solving the eigenproblem:

$$\begin{aligned} Au &= \lambda u \\ -i\frac{\mathrm{d}}{\mathrm{d}x}u &= \lambda u \\ u(x) &= e^{i\lambda x} \end{aligned}$$

Fourier Series

We restrict our space to periodic functions. Applying the periodic boundary condition:

$$u(a) = e^{i\lambda a} = u(b) = e^{i\lambda b}$$

so

$$\begin{aligned} e^{i\lambda(b-a)} &= 1 \\ \lambda = \frac{2\pi n}{b-a} \qquad & \text{for } n = 0, \pm 1, \pm 2, \ldots \end{aligned}$$

The normalized eigenvectors are:

$$u_n(x) = \frac{1}{\sqrt{b-a}} e^{i\frac{2\pi n}{b-a}x}$$

These eigenvectors belong to our space and as such all $\lambda = \frac{2\pi n}{b-a}$ form a discrete spectrum. Other solutions do not satisfy the periodic boundary condition and so there is no continous part in the spectrum.

The eigenvectors must be orthogonal, as we can check:

$$\begin{aligned} \int_a^b u_n^*(x) u_m(x) \mathrm{d}x = \\ = \int_a^b \frac{1}{\sqrt{b-a}} e^{-i\frac{2\pi n}{b-a}x} \frac{1}{\sqrt{b-a}} e^{i\frac{2\pi m}{b-a}x} \mathrm{d}x = \\ = \frac{1}{b-a} \int_a^b e^{i\frac{2\pi(m-n)}{b-a}x} \mathrm{d}x = \\ = \begin{cases} \frac{1}{b-a}\int_a^b e^0 \mathrm{d}x & \text{for } m = n \\ \frac{1}{i2\pi(m-n)} \left[e^{i\frac{2\pi(m-n)}{b-a}x} \right]_a^b & \text{for } m \neq n \end{cases} = \\ = \begin{cases} 1 & \text{for } m = n \\ \frac{1}{i2\pi(m-n)} \left(e^{i\frac{2\pi(m-n)}{b-a}b} - e^{i\frac{2\pi(m-n)}{b-a}a} \right) & \text{for } m \neq n \end{cases} = \\ = \begin{cases} 1 & \text{for } m = n \\ \frac{e^{i\frac{2\pi(m-n)}{b-a}a}}{i2\pi(m-n)} \left(e^{i\frac{2\pi(m-n)}{b-a}(b-a)} - 1 \right) & \text{for } m \neq n \end{cases} = \\ = \begin{cases} 1 & \text{for } m = n \\ 0 & \text{for } m \neq n \end{cases} = \delta_{mn} \end{aligned}$$

The eigenvectors must be complete:

$$\sum_{n=-\infty}^{\infty} |n\rangle \langle n| = \mathbb{1}$$

$$\sum_{n=-\infty}^{\infty} \langle x|n\rangle \langle n|x'\rangle = \langle x|x'\rangle$$

$$\sum_{n=-\infty}^{\infty} u_n(x) u_n^*(x') \mathrm{d}x = \delta(x - x')$$

Any function $f(x)$ can then be expanded on the interval $[a, b]$ into the Fourier series:

$$f(x) = \langle x|f\rangle = \sum_{n=-\infty}^{\infty} \langle x|n\rangle \langle n|f\rangle = \sum_{n=-\infty}^{\infty} c'_n u_n(x) = \sum_{n=-\infty}^{\infty} c'_n \frac{1}{\sqrt{b-a}} e^{i\frac{2\pi n}{b-a}x} = \sum_{n=-\infty}^{\infty} c_n e^{i\frac{2\pi n}{b-a}x}$$

$$c_n = c'_n \frac{1}{\sqrt{b-a}} = \langle n|f\rangle \frac{1}{\sqrt{b-a}} = \frac{1}{\sqrt{b-a}} \int_a^b \langle n|x\rangle \langle x|f\rangle \,\mathrm{d}x = \frac{1}{\sqrt{b-a}} \int_a^b u_n^*(x) f(x) \mathrm{d}x = \frac{1}{b-a} \int_a^b e^{-i\frac{2\pi n}{b-a}x} f(x) \mathrm{d}x$$

Equivalently, this can be written using $\sin$ and $\cos$ directly:

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{i\frac{2\pi n}{b-a}x} =$$

$$= \sum_{n=-\infty}^{\infty} c_n \cos\left(\frac{2\pi n}{b-a}x\right) + \sum_{n=-\infty}^{\infty} i c_n \sin\left(\frac{2\pi n}{b-a}x\right) =$$

$$= c_0 + \sum_{n=1}^{\infty} (c_n + c_{-n}) \cos\left(\frac{2\pi n}{b-a}x\right) + \sum_{n=1}^{\infty} i(c_n - c_{-n}) \sin\left(\frac{2\pi n}{b-a}x\right)$$

By introducing the coefficients a_n and b_n:

$$a_n = c_n + c_{-n} \qquad \text{for } n = 0, 1, 2, \ldots$$
$$b_n = i(c_n - c_{-n}) \qquad \text{for } n = 1, 2, \ldots$$

we can write the series as:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{2\pi n}{b-a}x\right) + \sum_{n=1}^{\infty} b_n \sin\left(\frac{2\pi n}{b-a}x\right)$$

we get:

$$a_n = c_n + c_{-n} = \frac{1}{b-a} \int_a^b \left(e^{-i\frac{2\pi n}{b-a}x} + e^{i\frac{2\pi n}{b-a}x}\right) f(x) \mathrm{d}x = \frac{2}{b-a} \int_a^b \cos\left(\frac{2\pi n}{b-a}x\right) f(x) \mathrm{d}x$$

$$b_n = i(c_n - c_{-n}) = \frac{i}{b-a} \int_a^b \left(e^{-i\frac{2\pi n}{b-a}x} - e^{i\frac{2\pi n}{b-a}x}\right) f(x) \mathrm{d}x = \frac{2}{b-a} \int_a^b \sin\left(\frac{2\pi n}{b-a}x\right) f(x) \mathrm{d}x$$

Conceptually, we are taking the complex orthonormal basis $u_n(x) = \frac{1}{\sqrt{b-a}} e^{i\frac{2\pi n}{b-a}x}$ and creating a real orthonormal basis $v_n(x)$ composed of $u_0, \operatorname{Re} u_1, \operatorname{Re} u_2, \ldots, \operatorname{Im} u_1, \operatorname{Im} u_2, \ldots$ as follows:

$$v_n(x) = \begin{cases} \sqrt{2} \operatorname{Re} u_n = \frac{\sqrt{2}}{\sqrt{b-a}} \cos \frac{2\pi}{b-a} n x & \text{for } n > 0 \\ u_0 = \frac{1}{\sqrt{b-a}} & \text{for } n = 0 \\ \sqrt{2} \operatorname{Im} u_{|n|} = \frac{\sqrt{2}}{\sqrt{b-a}} \sin \frac{2\pi}{b-a} |n| x & \text{for } n < 0 \end{cases}$$

We are only summing over the positive arguments in $\sin$ and $\cos$, thus the absolute value for $n < 0$. The basis v_n is orthonormal:

$$\int_a^b v_n(x) v_m(x) \mathrm{d}x = \delta_{nm}$$

and complete:

$$\sum_{n=-\infty}^{\infty} v_n(x)v_n(x')\mathrm{d}x = \delta(x-x')$$

This is not the only way to create the real orthonormal basis. In general:

$$\begin{aligned} u_n(x) &= \langle x|n\rangle \\ v_n(x) &= \langle x|n\rangle_R \\ |n\rangle_R &= \sum_m U_{nm}\,|m\rangle \end{aligned}$$

We require the new basis $|n\rangle_R$ to be orthonormal:

$$\begin{aligned} \langle n|m\rangle_R &= \delta_{nm} \\ \sum_{kl} \langle k|U^*_{nk}U_{ml}|l\rangle &= \delta_{nm} \\ \sum_{kl} U^*_{nk}U_{ml}\delta_{kl} &= \delta_{nm} \\ \sum_k U^*_{nk}U_{mk} &= \delta_{nm} \end{aligned}$$

This restricts the U_{nm} matrices to be unitary ($U^{-1} = U^\dagger$), because:

$$\begin{aligned} UU^\dagger &= \mathbb{1} \\ (UU^\dagger)_{mn} &= (\mathbb{1})_{mn} = \delta_{mn} \\ \sum_k (U)_{mk}(U^\dagger)_{kn} &= \delta_{mn} \\ \sum_k U_{mk}U^*_{nk} &= \delta_{mn} \end{aligned}$$

The unitarity condition also makes sure, that the real basis is complete:

$$\sum_n |n\rangle_R \langle n|_R = \sum_n \sum_{kl} U_{nk}\,|k\rangle \langle l|\, U^*_{nl} = \sum_{kl} \delta_{kl}\,|k\rangle \langle l| = \sum_k |k\rangle \langle k| = \mathbb{1}$$

Requiring $|n\rangle_R$ to be real and using $|m\rangle^* = |-m\rangle$ we get:

$$\begin{aligned} |n\rangle^*_R &= |n\rangle_R \\ \sum_m U^*_{nm}\,|m\rangle^* &= \sum_m U_{nm}\,|m\rangle \\ \sum_m U^*_{nm}\,|-m\rangle &= \sum_m U_{nm}\,|m\rangle \\ \sum_m U^*_{n,-m}\,|m\rangle &= \sum_m U_{nm}\,|m\rangle \\ \sum_m (U_{nm} - U^*_{n,-m})\,|m\rangle &= 0 \\ U_{nm} &= U^*_{n,-m} \\ U^*_{nm} &= U_{n,-m} \end{aligned}$$

Because the basis $|m\rangle$ is complete. So the only conditions on the matrices U_{mn} are:

$$\begin{aligned} U^{-1} &= U^\dagger \\ U^*_{nm} &= U_{n,-m} \end{aligned}$$

They imply that the new basis will be real, orthonormal and complete. Our final restriction is that we want each real basis element to correspond to the same frequency $\pm m$ (possible sign change is ok): this means that we can only mix the same frequencies, i.e.:

$$U_{nm} = 0 \quad \text{for } |n| \neq |m|$$

and also that the nonzero matrix elements can only be of the form $Re^{i\frac{\pi}{2}n}$ for $n = 0, 1, 2, 3$ (i.e. $\pm R$ or $\pm iR$ for some positive R).

Up to possible sign changes and permutations, this determines the matrix uniquely. Our choice above is:

$$U_{nm} = \begin{cases} \frac{\delta_{nm}+\delta_{n,-m}}{\sqrt{2}} & \text{for } n > 0 \\ \delta_{0m} & \text{for } n = 0 \\ \frac{\delta_{nm}-\delta_{n,-m}}{i\sqrt{2}} & \text{for } n < 0 \end{cases}$$

In other words, we get (except that the matrix is infinite):

$$\begin{pmatrix} v_3 \\ v_2 \\ v_1 \\ v_0 \\ v_{-1} \\ v_{-2} \\ v_{-3} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & & & & & & \frac{1}{\sqrt{2}} \\ & \frac{1}{\sqrt{2}} & & & & \frac{1}{\sqrt{2}} & \\ & & \frac{1}{\sqrt{2}} & & \frac{1}{\sqrt{2}} & & \\ & & & 1 & & & \\ & & \frac{1}{i\sqrt{2}} & & -\frac{1}{i\sqrt{2}} & & \\ & \frac{1}{i\sqrt{2}} & & & & -\frac{1}{i\sqrt{2}} & \\ \frac{1}{i\sqrt{2}} & & & & & & -\frac{1}{i\sqrt{2}} \end{pmatrix} \begin{pmatrix} u_3 \\ u_2 \\ u_1 \\ u_0 \\ u_{-1} \\ u_{-2} \\ u_{-3} \end{pmatrix}$$

Fourier Transform

Our domain is $(-\infty, \infty)$, so the solution of the eigen problem is:

$$\begin{aligned} Au &= \lambda u \\ -i\frac{\mathrm{d}}{\mathrm{d}x}u &= \lambda u \\ u(x) &= e^{i\lambda x} \end{aligned}$$

The normalized eigenfunctions are:

$$u_\lambda(x) = \frac{1}{\sqrt{2\pi}}e^{i\lambda x}$$

We calculate the normalization:

$$\begin{aligned} &\int_{-\infty}^{\infty} u_\lambda^*(x)u_\kappa(x)\mathrm{d}x = \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-i\lambda x}\frac{1}{\sqrt{2\pi}}e^{i\kappa x}\mathrm{d}x = \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i(\kappa-\lambda)x}\mathrm{d}x = \\ &= \delta(\kappa - \lambda) \end{aligned}$$

So the spectrum is continous. The eigenvectors must be complete:

$$\begin{aligned} \int_{-\infty}^{\infty} |\lambda\rangle \langle\lambda| \,\mathrm{d}\lambda &= \mathbb{1} \\ \int_{-\infty}^{\infty} \langle x|\lambda\rangle \langle\lambda|x'\rangle \,\mathrm{d}\lambda &= \langle x|x'\rangle \\ \int_{-\infty}^{\infty} u_\lambda(x)u_\lambda^*(x')\mathrm{d}\lambda &= \delta(x - x') \end{aligned}$$

Any function $f(x)$ can then be written as:

$$f(x) = \langle x|f\rangle = \int_{-\infty}^{\infty} \langle x|\lambda\rangle \langle \lambda|f\rangle \,\mathrm{d}\lambda = \int_{-\infty}^{\infty} u_\lambda(x)\hat{f}(\lambda)\mathrm{d}\lambda = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i\lambda x}\hat{f}(\lambda)\mathrm{d}\lambda$$

where $\hat{f}(\lambda)$ is called the Fourier transform of $f(x)$:

$$\hat{f}(\lambda) = \langle \lambda|f\rangle = \int_{-\infty}^{\infty} \langle \lambda|x\rangle \langle x|f\rangle \,\mathrm{d}x = \int_{-\infty}^{\infty} u_\lambda^*(x)f(x)\mathrm{d}x = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-i\lambda x}f(x)\mathrm{d}x$$

Note that both for Fourier series and Fourier transform, the sign convention in the exponentials ($e^{\pm i\lambda x}$) follows from choosing the sign in $A = -i\frac{\mathrm{d}}{\mathrm{d}x}$ and as such it is arbitrary. We can also choose $A' = i\frac{\mathrm{d}}{\mathrm{d}x}$ and then the sign will be flipped.

3.41.4 Sturm–Liouville Operator

The Sturm-Liouville operator L is:

$$Lu(x) = \frac{1}{w(x)}\left(-\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}u(x)}{dx}\right) + q(x)u(x)\right)$$

Everything is real. The scalar product is weighted by $w(x)$. The operator is self-adjoint if:

$$\int_a^b u(x)Lv(x)w(x)\mathrm{d}x = \int_a^b (Lu(x))v(x)w(x)\mathrm{d}x$$

so

$$\begin{aligned}
\int_a^b u(x)Lv(x)w(x)\mathrm{d}x &= \\
&= \int_a^b u(x)\frac{1}{w(x)}\left(-\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}v(x)}{dx}\right) + q(x)v(x)\right)w(x)\mathrm{d}x = \\
&= \int_a^b \left(-u(x)\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}v(x)}{dx}\right) + u(x)q(x)v(x)\right)\mathrm{d}x = \\
&= \int_a^b \left(\frac{\mathrm{d}u(x)}{\mathrm{d}x}p(x)\frac{\mathrm{d}v(x)}{dx} + u(x)q(x)v(x)\right)\mathrm{d}x - \left[u(x)p(x)\frac{\mathrm{d}v(x)}{dx}\right]_a^b = \\
&= \int_a^b \left(-\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}u(x)}{\mathrm{d}x}\right)v(x) + u(x)q(x)v(x)\right)\mathrm{d}x - \left[u(x)p(x)\frac{\mathrm{d}v(x)}{dx} - \frac{\mathrm{d}u(x)}{dx}p(x)v(x)\right]_a^b = \\
&= \int_a^b (Lu(x))\,v(x)w(x)\mathrm{d}x - \left[u(x)p(x)\frac{\mathrm{d}v(x)}{dx} - \frac{\mathrm{d}u(x)}{dx}p(x)v(x)\right]_a^b
\end{aligned}$$

And the operator L is self-adjoint if and only if:

$$[u(x)p(x)v'(x) - u'(x)p(x)v(x)]_a^b = 0$$

This condition can be satisfied by various boundary conditions. For example:

- Dirichlet boundary conditions

$$u(a) = 0, \quad u(b) = 0$$

- Neumann boundary conditions

$$u'(a) = 0, \quad u'(b) = 0$$

- Periodic boundary conditions

$$u(a) = u(b)$$
$$u'(a) = u'(b)$$

- Antiperiodic boundary conditions

$$u(a) = -u(b)$$
$$u'(a) = -u'(b)$$

or mixtures of these, e.g. Dirichlet at $x = a$ and Neumann at $x = b$.

Legendre Polynomials

Legendre polynomials $P_n(x)$ are solutions of the Sturm–Liouville problem on the interval $[-1, 1]$ with $p(x) = 1-x^2$, $q(x) = 0$, $w(x) = 1$ and $\lambda = n(n+1)$:

$$Lu(x) = n(n+1)u(x)$$
$$Lu(x) = -\frac{\mathrm{d}}{\mathrm{d}x}\left((1-x^2)\frac{\mathrm{d}u(x)}{dx}\right)$$

The operator L is self-adjoint due to vanishing $p(x)$ at the endpoints:

$$[(u(x)v'(x) - u'(x)v(x))p(x)]_{-1}^{1} = \left[(u(x)v'(x) - u'(x)v(x))(1-x^2)\right]_{-1}^{1} = 0$$

We restrict our space to bounded functions. The solutions of the eigenvalue problem for integer n are Legendre polynomials $P_n(x)$, the normalized eigenvectors $u_n(x)$ are:

$$u_n(x) = \sqrt{\frac{2n+1}{2}}P_n(x)$$

Solutions for non integer n are Legendre functions that are singular at the end points and as such are not solutions that we want. As such, the spectrum is discrete and the Legendre polynomials form a complete orthogonal basis for functions on the interval $[-1, 1]$:

$$\int_{-1}^{1} u_n(x)u_m(x) = \frac{2n+1}{2}\int_{-1}^{1} P_n(x)P_m(x) = \delta_{nm}$$
$$\sum_{n=0}^{\infty} u_n(x)u_n(x') = \frac{2n+1}{2}\sum_{n=0}^{\infty} P_n(x)P_n(x') = \delta(x-x')$$

any function $f(x)$ on the interval $[-1, 1]$ can be expanded as:

$$f(x) = \sum_{n=0}^{\infty} f'_n u_n(x) = \sum_{n=0}^{\infty} f'_n\sqrt{\frac{2n+1}{2}}P_n(x) = \sum_{n=0}^{\infty} f_n P_n(x)$$
$$f_n = f'_n\sqrt{\frac{2n+1}{2}} = \sqrt{\frac{2n+1}{2}}\int_{-1}^{1} u_n(x)f(x) = \frac{2n+1}{2}\int_{-1}^{1} P_n(x)f(x)$$

3.41.5 Angular Momentum Operator

The angular momentum operators L_1, L_2 and L_3 are given by:

$$L_j = -i\epsilon_{jkl}x_k\partial_l$$

in spherical coordinates:

$$\begin{aligned} L_1 &= i\left(\sin\phi\,\partial_\theta + \cot\theta\cos\phi\,\partial_\phi\right) \\ L_2 &= i\left(-\cos\phi\,\partial_\theta + \cot\theta\sin\phi\,\partial_\phi\right) \\ L_3 &= -i\partial_\phi \end{aligned}$$

and

$$L^2 = L_1^2 + L_2^2 + L_3^2 = -\left(\frac{1}{\sin\theta}\partial_\theta\left(\sin\theta\,\partial_\theta\right) + \frac{1}{\sin^2\theta}\partial_\phi^2\right)$$

The eigenproblem is:

$$\begin{aligned} L^2\left|lm\right\rangle &= l(l+1)\left|lm\right\rangle \\ L_3\left|lm\right\rangle &= m\left|lm\right\rangle \end{aligned} \tag{3.41.5.1}$$

Using Condon & Shortley phase convention, it can be shown that:

$$\left(L_1 \pm iL_2\right)\left|l,m\right\rangle = \sqrt{(l\mp m)(l\pm m+1)}\left|l,m\pm 1\right\rangle \tag{3.41.5.2}$$

and by repeated application:

$$\begin{aligned} &\left(L_1 \pm iL_2\right)^k\left|l,m\right\rangle = \\ &= \sqrt{(l\mp m)(l\mp m-1)\cdots(l\mp m-k+1)(l\pm m+1)(l\pm m+2)\cdots(l\mp m+k)}\left|l,m\pm k\right\rangle = \\ &= \sqrt{\frac{(l\mp m)!\,(l\pm m+k)!}{(l\pm m)!\,(l\mp m-k)!}}\left|l,m\pm k\right\rangle \end{aligned}$$

where

$$\begin{aligned} L_1 + iL_2 &= i\sin\phi\,\partial_\theta + i\cot\theta\cos\phi\,\partial_\phi \pm \left(\cos\phi\,\partial_\theta - \cot\theta\sin\phi\,\partial_\phi\right) = \\ &= e^{\pm i\phi}\left(\pm\partial_\theta + i\cot\theta\partial_\phi\right) \end{aligned}$$

The solution of (3.41.5.1) is of the form:

$$\langle\theta\phi|lm\rangle = Y_{lm}(\theta,\phi) = \Theta_{lm}(\theta)\Phi_m(\phi) \tag{3.41.5.3}$$

and we get from (3.41.5.1):

$$-i\frac{\mathrm{d}}{\mathrm{d}\phi}\Phi_m(\phi) = m\Phi_m(\phi)$$

on the interval $[0, 2\pi]$ with the boundary condition $\Phi_m(0) = \Phi_m(2\pi)$. From *Derivative Operator* the eigenvalues are all integer m and the normalized eigenvector is:

$$\Phi_m(\phi) = \frac{1}{\sqrt{2\pi}}e^{im\phi} \tag{3.41.5.4}$$

Substituting (3.41.5.4) into (3.41.5.3) we get from (3.41.5.1) an ordinary second order differential equation for $\Theta_{lm}(\theta)$:

$$\begin{aligned} L^2\left|lm\right\rangle &= l(l+1)\left|lm\right\rangle \\ -\left(\frac{1}{\sin\theta}\partial_\theta\left(\sin\theta\,\partial_\theta\right) + \frac{1}{\sin^2\theta}\partial_\phi^2\right)\frac{1}{\sqrt{2\pi}}e^{im\phi}\Theta_{lm} &= l(l+1)\frac{1}{\sqrt{2\pi}}e^{im\phi}\Theta_{lm} \\ \frac{1}{\sin\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\sin\theta\frac{\mathrm{d}}{\mathrm{d}\theta}\Theta_{lm}\right) + \left(l(l+1) - \frac{m^2}{\sin^2\theta}\right)\Theta_{lm} &= 0 \\ \frac{\mathrm{d}}{\mathrm{d}\cos\theta}\left((1-\cos^2\theta)\frac{\mathrm{d}}{\mathrm{d}\cos\theta}\Theta_{lm}\right) + \left(l(l+1) - \frac{m^2}{1-\cos^2\theta}\right)\Theta_{lm} &= 0 \\ \frac{\mathrm{d}}{\mathrm{d}z}\left((1-z^2)\frac{\mathrm{d}\Theta_{lm}}{\mathrm{d}z}\right) + \left(l(l+1) - \frac{m^2}{1-z^2}\right)\Theta_{lm} &= 0 \end{aligned}$$

where

$$z = \cos\theta$$

This equation can be solved using the following approach. From (3.41.5.2) we get:

$$\begin{aligned}
(L_1 \pm iL_2)Y_{lm}(\theta,\phi) = (L_1 \pm iL_2)\Theta_{lm}(\theta)\Phi_m(\phi) = \\
= e^{\pm i\phi}\left(\pm\partial_\theta + i\cot\theta\partial_\phi\right)\Theta_{lm}(\theta)\frac{1}{\sqrt{2\pi}}e^{im\phi} = \\
= \frac{1}{\sqrt{2\pi}}e^{i(m\pm1)\phi}\left(\pm\frac{\mathrm{d}}{\mathrm{d}\theta} - m\cot\theta\right)\Theta_{lm}(\theta) = \\
= \mp\frac{1}{\sqrt{2\pi}}e^{i(m\pm1)\phi}\left(\sin\theta\frac{\mathrm{d}}{\mathrm{d}\cos\theta} \mp m\frac{\mathrm{d}\sin\theta}{\mathrm{d}\cos\theta}\right)\Theta_{lm}(\theta) = \\
= \mp\frac{1}{\sqrt{2\pi}}e^{i(m\pm1)\phi}\sin^{1\pm m}\theta\left(\frac{\mathrm{d}}{\mathrm{d}\cos\theta}\sin^{\mp m}\theta\,\Theta_{lm}(\theta)\right) = \\
= \mp\Phi_{m\pm1}(\phi)\sin^{1\pm m}\theta\left(\frac{\mathrm{d}}{\mathrm{d}\cos\theta}\sin^{\mp m}\theta\,\Theta_{lm}(\theta)\right)
\end{aligned}$$

and by repeated application we get:

$$\begin{aligned}
(L_1 \pm iL_2)^k Y_{lm}(\theta,\phi) = (\mp1)^k\Phi_{m\pm k}(\phi)\sin^{k\pm m}\theta\left(\frac{\mathrm{d}^k}{(\mathrm{d}\cos\theta)^k}\sin^{\mp m}\theta\,\Theta_{lm}(\theta)\right) = \\
= \sqrt{\frac{(l\mp m)!}{(l\pm m)!}\frac{(l\pm m+k)!}{(l\mp m-k)!}}\Phi_{m\pm k}(\phi)\Theta_{l,m\pm k}(\theta)
\end{aligned}$$

from which we obtain:

$$\Theta_{l,m\pm k}(\theta) = \sqrt{\frac{(l\pm m)!}{(l\mp m)!}\frac{(l\mp m-k)!}{(l\pm m+k)!}}(\mp1)^k\sin^{k\pm m}\theta\left(\frac{\mathrm{d}^k}{(\mathrm{d}\cos\theta)^k}\sin^{\mp m}\theta\,\Theta_{lm}(\theta)\right) \tag{3.41.5.5}$$

As a special case for $m = 0$ and $k = m > 0$ we get:

$$\Theta_{l,\pm m}(\theta) = (\mp1)^m\sqrt{\frac{(l-m)!}{(l+m)!}}\sin^m\theta\left(\frac{\mathrm{d}^m}{(\mathrm{d}\cos\theta)^m}\Theta_{l0}(\theta)\right) \tag{3.41.5.6}$$

and for $m = l$ and $k = l - m$ we get (we only use the $\Theta_{l,m-k}$ branch):

$$\begin{aligned}
\Theta_{lm}(\theta) = \Theta_{l,l-(l-m)}(\theta) = \\
= \sqrt{\frac{(l-l)!}{(l+l)!}\frac{(l+l-(l-m))!}{(l-l+l-m)!}}(+1)^{l-m}\sin^{l-m-l}\theta\left(\frac{\mathrm{d}^{l-m}}{(\mathrm{d}\cos\theta)^{l-m}}\sin^{+l}\theta\,\Theta_{ll}(\theta)\right) = \\
= \sqrt{\frac{1}{(2l)!}\frac{(l+m)!}{(l-m)!}}\frac{1}{\sin^m\theta}\left(\frac{\mathrm{d}^{l-m}}{(\mathrm{d}\cos\theta)^{l-m}}\sin^l\theta\,\Theta_{ll}(\theta)\right)
\end{aligned} \tag{3.41.5.7}$$

From (3.41.5.2) we get:

$$(L_1 + iL_2)Y_{ll} = \sqrt{(l-l)(l+l+1)}Y_{ll} = 0$$

Using (3.41.5.4) this gives us a first order differential equation:

$$\begin{aligned}
(L_1 + iL_2)\Theta_{ll}\Phi_l = 0 \\
e^{i\phi}\left(\partial_\theta + i\cot\theta\partial_\phi\right)\Theta_{ll}\frac{1}{\sqrt{2\pi}}e^{il\phi} = 0 \\
\frac{\partial\Theta_{ll}}{\partial\theta} - l\cot\theta\,\Theta_{ll} = 0
\end{aligned}$$

from which

$$\Theta_{ll}(\theta) = (-1)^l \sqrt{\frac{(2l+1)!}{2}} \frac{1}{2^l l!} \sin^l \theta \tag{3.41.5.8}$$

It is normalized as:

$$\int_0^\pi \Theta_{ll}^2 \sin\theta \, \mathrm{d}\theta = 1$$

We used the value of the integral:

$$\int_0^\pi \sin^{2l+1}\theta \, \mathrm{d}\theta = \frac{\sqrt{\pi}\,\Gamma(l+1)}{\Gamma(l+\frac{3}{2})} = \frac{\sqrt{\pi}\,2^{l+1}l!}{(2l+1)!!\sqrt{\pi}} = \frac{2^{l+1}l!}{(2l+1)!!} = \frac{2^{2l+2}(l+1)!l!}{(2l+2)!} =$$
$$= \frac{(2^{l+1}l!)^2(l+1)}{(2l+2)!} = \frac{4(2^l l!)^2(l+1)}{(2l+1)!2(l+1)} = \frac{2(2^l l!)^2}{(2l+1)!}$$

Using (3.41.5.8) in (3.41.5.7) we get:

$$\Theta_{lm}(\theta) = (-1)^l \sqrt{\frac{2l+1}{2}\frac{(l+m)!}{(l-m)!}} \frac{1}{2^l l!} \frac{1}{\sin^m\theta} \frac{\mathrm{d}^{l-m}}{(\mathrm{d}\cos\theta)^{l-m}} \sin^{2l}\theta$$

for $m = 0$ we obtain:

$$\Theta_{l0}(\theta) = (-1)^l \sqrt{\frac{2l+1}{2}} \frac{1}{2^l l!} \frac{\mathrm{d}^l}{(\mathrm{d}\cos\theta)^l} \sin^{2l}\theta =$$
$$= \sqrt{\frac{2l+1}{2}} \frac{1}{2^l l!} \frac{\mathrm{d}^l}{(\mathrm{d}\cos\theta)^l} (\cos^2\theta - 1)^l =$$
$$= \sqrt{\frac{2l+1}{2}} P_l(\cos\theta)$$

where

$$P_l(z) = \frac{1}{2^l l!} \frac{\mathrm{d}^l}{\mathrm{d}z^l}(z^2-1)^l$$

is the Rodrigues' formula for Legendre polynomials. We substitute Θ_{l0} into (3.41.5.6) and get:

$$\Theta_{l,\pm m}(\theta) = (\mp 1)^m \sqrt{\frac{2l+1}{2}\frac{(l-m)!}{(l+m)!}} \sin^m\theta \left(\frac{\mathrm{d}^m}{(\mathrm{d}\cos\theta)^m} P_l(\cos\theta) \right) \tag{3.41.5.9}$$

Hence $\Theta_{lm} = (-1)^m \Theta_{l,-m}$. Using associated Legendre polynomials, we can write:

$$\Theta_{lm}(\theta) = \sqrt{\frac{2l+1}{2}\frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta) \tag{3.41.5.10}$$

where (for all m):

$$P_l^m(\cos\theta) = (-1)^l \frac{(l+m)!}{(l-m)!} \frac{1}{2^l l!} \frac{1}{\sin^m\theta} \frac{\mathrm{d}^{l-m}}{(\mathrm{d}\cos\theta)^{l-m}} \sin^{2l}\theta =$$
$$= \frac{(l+m)!}{(l-m)!} \frac{1}{2^l l!} \frac{1}{\sin^m\theta} \frac{\mathrm{d}^{l-m}}{(\mathrm{d}\cos\theta)^{l-m}} (\cos^2\theta - 1)^l =$$
$$= (-1)^m \frac{1}{2^l l!} \frac{(1-\cos^2)^m\theta}{\sin^m\theta} \frac{\mathrm{d}^{l+m}}{(\mathrm{d}\cos\theta)^{l+m}} (\cos^2\theta - 1)^l =$$
$$= (-1)^m \frac{1}{2^l l!} \sin^m\theta \frac{\mathrm{d}^{l+m}}{(\mathrm{d}\cos\theta)^{l+m}} (\cos^2\theta - 1)^l =$$
$$= (-1)^m \frac{1}{2^l l!} (1-z^2)^{\frac{m}{2}} \frac{\mathrm{d}^{l+m}}{\mathrm{d}z^{l+m}} (z^2-1)^l$$

hence (comparing the second and fourth equation above):

$$P_l^{-m}(z) = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(z) \tag{3.41.5.11}$$

This is valid for all m (positive or negative). For $m \geq 0$ we get from (3.41.5.9) and (3.41.5.11):

$$P_l^m(\cos\theta) = (-1)^m \sin^m\theta \frac{\mathrm{d}^m}{(\mathrm{d}\cos\theta)^m} P_l(\cos\theta)$$
$$P_l^{-m}(\cos\theta) = \frac{(l-m)!}{(l+m)!} \sin^m\theta \frac{\mathrm{d}^m}{(\mathrm{d}\cos\theta)^m} P_l(\cos\theta)$$
$$P_l^m(z) = (-1)^m (1-z^2)^{\frac{m}{2}} \frac{\mathrm{d}^m}{\mathrm{d}z^m} P_l(z)$$
$$P_l^{-m}(z) = \frac{(l-m)!}{(l+m)!} (1-z^2)^{\frac{m}{2}} \frac{\mathrm{d}^m}{\mathrm{d}z^m} P_l(z)$$

This is usually used as the definition of the associated Legendre polynomials. They include the Condon & Shortley phase factor $(-1)^m$ (only for positive m). Some authors omit it (then it needs to be included in the equation (3.41.5.10)). Note that (3.41.5.10) for $m < 0$ can be also written as:

$$\Theta_{lm}(\theta) = \sqrt{\frac{2l+1}{2}\frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta) = \sqrt{\frac{2l+1}{2}\frac{(l+m)!}{(l-m)!}\frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta) =$$
$$= (-1)^m \sqrt{\frac{2l+1}{2}\frac{(l+m)!}{(l-m)!}} P_l^{-m}(\cos\theta) = (-1)^m \sqrt{\frac{2l+1}{2}\frac{(l-|m|)!}{(l+|m|)!}} P_l^{|m|}(\cos\theta)$$

Thanks to

$$i^{m-|m|} = \begin{cases} 1 & \text{for } m \geq 0 \\ (-1)^m & \text{for } m < 0 \end{cases}$$

we can write for all m:

$$\Theta_{lm}(\theta) = \sqrt{\frac{2l+1}{2}\frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta) = i^{m-|m|} \sqrt{\frac{2l+1}{2}\frac{(l-|m|)!}{(l+|m|)!}} P_l^{|m|}(\cos\theta)$$

The normalization of associated Legendre polynomials is:

$$\int_{-1}^{1} \Theta_{lm}(\theta)\Theta_{l'm}(\theta) \sin\theta \mathrm{d}\theta = \delta_{ll'}$$
$$\int_{-1}^{1} P_l^m(x) P_{l'}^m(x) \mathrm{d}x = \frac{2}{2l+1}\frac{(l+m)!}{(l-m)!}\delta_{ll'}$$

Finally, we get (for all m):

$$Y_{lm}(\theta,\phi) = \Theta_{lm}(\theta)\Phi_m(\phi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta) e^{im\phi} =$$
$$= i^{m-|m|} \sqrt{\frac{2l+1}{4\pi}\frac{(l-|m|)!}{(l+|m|)!}} P_l^{|m|}(\cos\theta) e^{im\phi} =$$
$$= (-1)^m i^{m+|m|} \sqrt{\frac{2l+1}{4\pi}\frac{(l-|m|)!}{(l+|m|)!}} P_l^{|m|}(\cos\theta) e^{im\phi}$$

Any function on the sphere can be expanded as:

$$f(\theta,\phi) = \langle\theta\phi|f\rangle = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} \langle\theta\phi|lm\rangle \langle lm|f\rangle = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} Y_{lm}(\theta,\phi) f_{lm}$$

$$f_{lm} = \langle lm|f\rangle = \int \langle lm|\Omega\rangle \langle\Omega|f\rangle \,\mathrm{d}\Omega =$$

$$= \int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \mathrm{d}\theta \,\langle lm|\theta\phi\rangle \langle\theta\phi|f\rangle \sin\theta = \int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \mathrm{d}\theta \, Y_{lm}^*(\theta,\phi) f(\theta,\phi) \sin\theta$$

Real Spherical Harmonics

The most obvious approach is to use a similar way as for Fourier series. We rearrange the sum:

$$f(\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} Y_{lm}(\theta,\phi) f_{lm} = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} \Theta_{lm}(\theta,\phi) \frac{1}{\sqrt{2\pi}} e^{im\phi} f_{lm} =$$

$$= \frac{1}{\sqrt{2\pi}} \sum_{l=0}^{\infty}\sum_{m=-l}^{l} \left(\Theta_{lm}(\theta,\phi)\cos m\phi f_{lm} + \Theta_{lm}(\theta,\phi) i \sin m\phi f_{lm}\right) =$$

$$= \frac{1}{\sqrt{2\pi}} \sum_{l=0}^{\infty} \left(\Theta_{l0}(\theta,\phi) f_{l0} + \sum_{m=1}^{l} \left((\Theta_{lm}(\theta,\phi) f_{lm} + \Theta_{l,-m}(\theta,\phi) f_{l,-m}) \cos m\phi + i(\Theta_{lm}(\theta,\phi) f_{lm} - \Theta_{l,-m}(\theta,\phi) f_{l,-m}) \sin m\phi \right) \right) =$$

$$= \frac{1}{\sqrt{2\pi}} \sum_{l=0}^{\infty} \left(\Theta_{l0}(\theta,\phi) f_{l0} + \sum_{m=1}^{l} \left(\Theta_{lm}(\theta,\phi)(f_{lm} + (-1)^m f_{l,-m}) \cos m\phi + \Theta_{lm}(\theta,\phi) i (f_{lm} - (-1)^m f_{l,-m}) \sin m\phi \right) \right) =$$

$$= \frac{1}{\sqrt{2\pi}} \sum_{l=0}^{\infty} \left(\Theta_{l0}(\theta,\phi) f_{l0} + \sum_{m=1}^{l} \left(\Theta_{lm}(\theta,\phi) \frac{f_{lm} + (-1)^m f_{l,-m}}{\sqrt{2}} \sqrt{2} \cos m\phi + \Theta_{lm}(\theta,\phi) i \frac{f_{lm} - (-1)^m f_{l,-m}}{\sqrt{2}} \sqrt{2} \sin m\phi \right) \right) =$$

$$= \frac{1}{\sqrt{2\pi}} \sum_{l=0}^{\infty} \left(\Theta_{l0}(\theta,\phi) \tilde{f}_{l0} + \sum_{m=1}^{l} \left(\Theta_{lm}(\theta,\phi) \tilde{f}_{lm} \sqrt{2} \cos m\phi + \Theta_{lm}(\theta,\phi) \tilde{f}_{l,-m} \sqrt{2} \sin m\phi \right) \right) =$$

$$= \frac{1}{\sqrt{2\pi}} \sum_{l=0}^{\infty} \left(\Theta_{l0}(\theta,\phi) \tilde{f}_{l0} + \sum_{m=1}^{l} \Theta_{lm}(\theta,\phi) \tilde{f}_{lm} \sqrt{2} \cos m\phi + \sum_{m=-l}^{-1} \Theta_{l|m|}(\theta,\phi) \tilde{f}_{lm} \sqrt{2} \sin |m|\phi \right) =$$

$$= \sum_{l=0}^{\infty}\sum_{m=-l}^{l} Z_{lm}(\theta,\phi) \tilde{f}_{lm}$$

Where the real spherical harmonics Z_{lm} are:

$$Z_{lm}(\theta,\phi) = \begin{cases} \sqrt{2}\frac{\Theta_{lm}}{\sqrt{2\pi}} \cos m\phi & \text{for } m > 0 \\ \frac{\Theta_{l0}}{\sqrt{2\pi}} & \text{for } m = 0 \\ \sqrt{2}\frac{\Theta_{l|m|}}{\sqrt{2\pi}} \sin |m|\phi & \text{for } m < 0 \end{cases} = \begin{cases} \sqrt{2}\,\mathrm{Re}(Y_{lm}(\theta,\phi)) \\ Y_{l0}(\theta,\phi) \\ \sqrt{2}\,\mathrm{Im}(Y_{l|m|}(\theta,\phi)) \end{cases} = \begin{cases} \frac{1}{\sqrt{2}}(Y_{lm}(\theta,\phi) + Y_{lm}^*(\theta,\phi)) \\ Y_{l0}(\theta,\phi) \\ \frac{1}{i\sqrt{2}}(Y_{l|m|}(\theta,\phi) - Y_{l|m|}^*(\theta,\phi)) \end{cases}$$

and the coefficients $\tilde{f}_{lm}$ are:

$$\tilde{f}_{lm} = \begin{cases} \frac{f_{lm} + (-1)^m f_{l,-m}}{\sqrt{2}} & \text{for } m > 0 \\ f_{l0} & \text{for } m = 0 \\ i\frac{f_{l,-m} - (-1)^m f_{lm}}{\sqrt{2}} & \text{for } m < 0 \end{cases} = \begin{cases} \int \frac{Y_{lm}^* + (-1)^m Y_{l,-m}^*}{\sqrt{2}} f \mathrm{d}\Omega \\ \int Y_{l0}^* f \mathrm{d}\Omega \\ \int i\frac{Y_{l,-m}^* - (-1)^m Y_{lm}^*}{\sqrt{2}} f \mathrm{d}\Omega \end{cases}$$

$$= \begin{cases} \int \frac{Y_{lm}^* + Y_{lm}}{\sqrt{2}} f \mathrm{d}\Omega \\ \int Y_{l0} f \mathrm{d}\Omega \\ \int \frac{Y_{l,-m} - Y_{l,-m}^*}{i\sqrt{2}} f \mathrm{d}\Omega \end{cases} = \begin{cases} \int \frac{Y_{lm}^* + Y_{lm}}{\sqrt{2}} f \mathrm{d}\Omega \\ \int Y_{l0} f \mathrm{d}\Omega \\ \int \frac{Y_{l|m|} - Y_{l|m|}^*}{i\sqrt{2}} f \mathrm{d}\Omega \end{cases} = \int Z_{lm} f \mathrm{d}\Omega$$

The factor $\sqrt{2}$ in the definition makes the real spherical harmonics properly normalized:

$$\int Z_{lm}(\theta,\phi)Z_{l'm'}(\theta,\phi)\mathrm{d}\Omega = \delta_{ll'}\delta_{mm'}$$

From the above derivation, it is not immediately clear how to obtain other parametrizations of real spherical harmonics. And also what identities they obey. More systematic approach is to use the transformation matrices just like for the Fourier series:

$$Z_{l\mu}(\theta,\phi) = \langle\theta\phi|l\mu\rangle_R = \sum_{m=-l}^{l} U^l_{\mu m} Y_{lm}(\theta,\phi) = \sum_{m=-l}^{l} U^l_{\mu m} \langle\theta\phi|lm\rangle$$

$$|l\mu\rangle_R = \sum_{m=-l}^{l} U^l_{\mu m} |lm\rangle$$

We require orthonormality:

$$\langle l\mu|l\mu'\rangle_R = \delta_{\mu\mu'}$$

This implies unitarity of the U^l matrices for the given l. Requiring $|l\mu\rangle_R$ to be real and using $|lm\rangle^* = (-1)^m |l,-m\rangle$ we get:

$$\begin{aligned}
|l\mu\rangle^*_R &= |l\mu\rangle_R \\
\sum_m (U^l_{\mu m})^* |lm\rangle^* &= \sum_m U^l_{\mu m} |lm\rangle \\
\sum_m (U^l_{\mu m})^* (-1)^m |l,-m\rangle &= \sum_m U^l_{\mu m} |lm\rangle \\
\sum_m (U^l_{\mu,-m})^* (-1)^m |lm\rangle &= \sum_m U^l_{\mu m} |lm\rangle \\
\sum_m (U^l_{\mu m} - (U^l_{\mu,-m})^*)(-1)^m) |lm\rangle &= 0 \\
U^l_{\mu m} &= (-1)^m (U^l_{\mu,-m})^* \\
(U^l_{\mu m})^* &= (-1)^m U^l_{\mu,-m}
\end{aligned}$$

As for Fourier series, we require not to mix frequencies and phases, so we get:

$$U^l_{nm} = 0 \quad \text{for } |n| \neq |m|$$

and also that the nonzero matrix elements can only be of the form $Re^{i\frac{\pi}{2}n}$ for $n = 0, 1, 2, 3$ (i.e. $\pm R$ or $\pm iR$ for some positive R). Up to signs and permutations, this determines the matrices uniquely. As for Fourier series, this implies orthonormality and completeness of the real spherical harmonics:

$$\begin{aligned}
\langle l'm'|lm\rangle_R &= \delta_{ll'}\delta_{mm'} \\
\sum_{l=0}^{\infty} \sum_{m=-l}^{l} |lm\rangle_R \langle lm|_R &= \mathbb{1}
\end{aligned}$$

Also, thanks to unitarity we get:

$$\sum_{m=-l}^{l} Z_{lm}(\Omega)Z_{lm}(\Omega') = \sum_{m=-l}^{l} \sum_{m'm''} U^l_{mm'} Y_{lm'}(\Omega)(U^l_{mm''})^* Y^*_{lm''}(\Omega') = \sum_{m'm''} \delta_{m'm''} Y_{lm'}(\Omega) Y^*_{lm''}(\Omega') =$$

$$= \sum_{m=-l}^{l} Y_{lm}(\Omega) Y^*_{lm}(\Omega') = \frac{2l+1}{4\pi} P_l(\cos\gamma)$$

and

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|} = \sum_{l=0}^{\infty} \frac{r_<^l}{r_>^{l+1}} P_l(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}') = \sum_{lm} \frac{r_<^l}{r_>^{l+1}} \frac{4\pi}{2l+1} Y_{lm}(\hat{\mathbf{r}}) Y_{lm}^*(\hat{\mathbf{r}}') = \sum_{lm} \frac{r_<^l}{r_>^{l+1}} \frac{4\pi}{2l+1} Z_{lm}(\hat{\mathbf{r}}) Z_{lm}(\hat{\mathbf{r}}')$$

Following the Fourier series, the most natural way to choose the signs in the U^l matrices is such so as to keep $\sin$ and $\cos$ in the basis with positive frequencies (thus the absolute value for $m<0$):

$$Z_{lm}(\theta,\phi) = \begin{cases} \sqrt{2}\,\mathrm{Re}(Y_{lm}(\theta,\phi)) & \text{for } m>0 \\ Y_{l0}(\theta,\phi) & \text{for } m=0 \\ \sqrt{2}\,\mathrm{Im}(Y_{l|m|}(\theta,\phi)) & \text{for } m<0 \end{cases} = \begin{cases} \frac{1}{\sqrt{2}}(Y_{lm}(\theta,\phi)+Y_{lm}^*(\theta,\phi)) \\ Y_{l0}(\theta,\phi) \\ \frac{1}{i\sqrt{2}}(Y_{l|m|}(\theta,\phi)-Y_{l|m|}^*(\theta,\phi)) \end{cases} =$$

$$= \begin{cases} \frac{1}{\sqrt{2}}(Y_{lm}(\theta,\phi)+Y_{lm}^*(\theta,\phi)) \\ Y_{l0}(\theta,\phi) \\ \frac{-(-1)^m}{i\sqrt{2}}(Y_{lm}(\theta,\phi)-Y_{lm}^*(\theta,\phi)) \end{cases} = \begin{cases} \frac{1}{\sqrt{2}}(Y_{lm}(\theta,\phi)+(-1)^m Y_{l,-m}(\theta,\phi)) \\ Y_{l0}(\theta,\phi) \\ \frac{1}{i\sqrt{2}}(Y_{l,-m}(\theta,\phi)-(-1)^m Y_{lm}(\theta,\phi)) \end{cases} =$$

$$= \begin{cases} \sqrt{\frac{2l+1}{2\pi}\frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta)\cos m\phi \\ \sqrt{\frac{2l+1}{4\pi}} P_l(\cos\theta) \\ \sqrt{\frac{2l+1}{2\pi}\frac{(l-|m|)!}{(l+|m|)!}} P_l^{|m|}(\cos\theta)\sin|m|\phi \end{cases}$$

This gives:

$$U_{\mu m}^l = \begin{cases} \frac{\delta_{\mu m}+(-1)^m\delta_{\mu,-m}}{\sqrt{2}} & \text{for } \mu>0 \\ \delta_{0m} & \text{for } \mu=0 \\ \frac{\delta_{\mu,-m}-(-1)^m\delta_{\mu m}}{i\sqrt{2}} & \text{for } \mu<0 \end{cases}$$

Other convention

Some people use the following convention:

$$Z_{lm}(\theta,\phi) = \begin{cases} (-1)^m\sqrt{2}\,\mathrm{Re}(Y_{lm}(\theta,\phi)) & \text{for } m>0 \\ Y_{l0}(\theta,\phi) & \text{for } m=0 \\ (-1)^m\sqrt{2}\,\mathrm{Im}(Y_{l|m|}(\theta,\phi)) & \text{for } m<0 \end{cases} = \begin{cases} \frac{1}{\sqrt{2}}((-1)^m Y_{lm}(\theta,\phi)+Y_{l,-m}(\theta,\phi)) \\ Y_{l0}(\theta,\phi) \\ \frac{1}{i\sqrt{2}}((-1)^m Y_{l,-m}(\theta,\phi)-Y_{lm}(\theta,\phi)) \end{cases} =$$

$$= \begin{cases} (-1)^m\sqrt{\frac{2l+1}{2\pi}\frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta)\cos m\phi \\ \sqrt{\frac{2l+1}{4\pi}} P_l(\cos\theta) \\ (-1)^m\sqrt{\frac{2l+1}{2\pi}\frac{(l-|m|)!}{(l+|m|)!}} P_l^{|m|}(\cos\theta)\sin|m|\phi \end{cases}$$

It has the advantage that there are no minus signs in the final expressions using $\sin$, $\cos$ or using x, y, z. However, we will not use this convention.

Tables

Spherical harmonics:

$$Y_{0,0}(\theta,\phi) = \frac{1}{2\sqrt{\pi}}$$

$$Y_{1,-1}(\theta,\phi) = \frac{\sqrt{6}e^{-\imath\phi}\sin(\theta)}{4\sqrt{\pi}}$$

$$Y_{1,0}(\theta,\phi) = \frac{\sqrt{3}\cos(\theta)}{2\sqrt{\pi}}$$

$$Y_{1,1}(\theta,\phi) = -\frac{\sqrt{6}e^{\imath\phi}\sin(\theta)}{4\sqrt{\pi}}$$

$$Y_{2,-2}(\theta,\phi) = \frac{\sqrt{30}e^{-2\imath\phi}\sin^2(\theta)}{8\sqrt{\pi}}$$

$$Y_{2,-1}(\theta,\phi) = \frac{\sqrt{30}e^{-\imath\phi}\sin(\theta)\cos(\theta)}{4\sqrt{\pi}}$$

$$Y_{2,0}(\theta,\phi) = \frac{\sqrt{5}\left(\frac{3}{2}\cos^2(\theta) - \frac{1}{2}\right)}{2\sqrt{\pi}}$$

$$Y_{2,1}(\theta,\phi) = -\frac{\sqrt{30}e^{\imath\phi}\sin(\theta)\cos(\theta)}{4\sqrt{\pi}}$$

$$Y_{2,2}(\theta,\phi) = \frac{\sqrt{30}e^{2\imath\phi}\sin^2(\theta)}{8\sqrt{\pi}}$$

$$Y_{3,-3}(\theta,\phi) = \frac{\sqrt{35}e^{-3\imath\phi}\sin^3(\theta)}{8\sqrt{\pi}}$$

$$Y_{3,-2}(\theta,\phi) = \frac{\sqrt{210}e^{-2\imath\phi}\sin^2(\theta)\cos(\theta)}{8\sqrt{\pi}}$$

$$Y_{3,-1}(\theta,\phi) = -\frac{\sqrt{21}\left(6\sin^4(\theta) - 24\sin^2(\theta)\cos^2(\theta)\right)e^{-\imath\phi}}{48\sqrt{\pi}\sin(\theta)}$$

$$Y_{3,0}(\theta,\phi) = \frac{\sqrt{7}\left(-\frac{3}{2}\sin^2(\theta)\cos(\theta) + \cos^3(\theta)\right)}{2\sqrt{\pi}}$$

$$Y_{3,1}(\theta,\phi) = -\frac{\sqrt{21}\left(360\cos^2(\theta) - 72\right)e^{\imath\phi}\sin(\theta)}{576\sqrt{\pi}}$$

$$Y_{3,2}(\theta,\phi) = \frac{\sqrt{210}e^{2\imath\phi}\sin^2(\theta)\cos(\theta)}{8\sqrt{\pi}}$$

$$Y_{3,3}(\theta,\phi) = -\frac{\sqrt{35}e^{3\imath\phi}\sin^3(\theta)}{8\sqrt{\pi}}$$

Real spherical harmonics:

$$Z_{0,0}(\theta,\phi) = \frac{1}{2\sqrt{\pi}}$$

$$Z_{1,-1}(\theta,\phi) = -\frac{\sqrt{3}\sin(\phi)\sin(\theta)}{2\sqrt{\pi}}$$

$$Z_{1,0}(\theta,\phi) = \frac{\sqrt{3}\cos(\theta)}{2\sqrt{\pi}}$$

$$Z_{1,1}(\theta,\phi) = -\frac{\sqrt{3}\sin(\theta)\cos(\phi)}{2\sqrt{\pi}}$$

$$Z_{2,-2}(\theta,\phi) = \frac{\sqrt{15}\sin(2\phi)\sin^2(\theta)}{4\sqrt{\pi}}$$

$$Z_{2,-1}(\theta,\phi) = -\frac{\sqrt{15}\sin(\phi)\sin(\theta)\cos(\theta)}{2\sqrt{\pi}}$$

$$Z_{2,0}(\theta,\phi) = \frac{\sqrt{5}\left(\frac{3}{2}\cos^2(\theta) - \frac{1}{2}\right)}{2\sqrt{\pi}}$$

$$Z_{2,1}(\theta,\phi) = -\frac{\sqrt{15}\sin(\theta)\cos(\phi)\cos(\theta)}{2\sqrt{\pi}}$$

$$Z_{2,2}(\theta,\phi) = \frac{\sqrt{15}\sin^2(\theta)\cos(2\phi)}{4\sqrt{\pi}}$$

$$Z_{3,-3}(\theta,\phi) = -\frac{\sqrt{70}\sin(3\phi)\sin^3(\theta)}{8\sqrt{\pi}}$$

$$Z_{3,-2}(\theta,\phi) = \frac{\sqrt{105}\sin(2\phi)\sin^2(\theta)\cos(\theta)}{4\sqrt{\pi}}$$

$$Z_{3,-1}(\theta,\phi) = -\frac{\sqrt{42}\left(360\cos^2(\theta) - 72\right)\sin(\phi)\sin(\theta)}{576\sqrt{\pi}}$$

$$Z_{3,0}(\theta,\phi) = \frac{\sqrt{7}\left(-\frac{3}{2}\sin^2(\theta)\cos(\theta) + \cos^3(\theta)\right)}{2\sqrt{\pi}}$$

$$Z_{3,1}(\theta,\phi) = -\frac{\sqrt{42}\left(360\cos^2(\theta) - 72\right)\sin(\theta)\cos(\phi)}{576\sqrt{\pi}}$$

$$Z_{3,2}(\theta,\phi) = \frac{\sqrt{105}\sin^2(\theta)\cos(2\phi)\cos(\theta)}{4\sqrt{\pi}}$$

$$Z_{3,3}(\theta,\phi) = -\frac{\sqrt{70}\sin^3(\theta)\cos(3\phi)}{8\sqrt{\pi}}$$

Real spherical harmonics (using x, y and z, assuming $x^2 + y^2 + z^2 = 1$):

$$Z_{0,0}(x, y, z) = \frac{1}{2\sqrt{\pi}}$$

$$Z_{1,-1}(x, y, z) = -\frac{\sqrt{3}y}{2\sqrt{\pi}}$$

$$Z_{1,0}(x, y, z) = \frac{\sqrt{3}z}{2\sqrt{\pi}}$$

$$Z_{1,1}(x, y, z) = -\frac{\sqrt{3}x}{2\sqrt{\pi}}$$

$$Z_{2,-2}(x, y, z) = \frac{\sqrt{15}xy}{2\sqrt{\pi}}$$

$$Z_{2,-1}(x, y, z) = -\frac{\sqrt{15}yz}{2\sqrt{\pi}}$$

$$Z_{2,0}(x, y, z) = \frac{\sqrt{5}\left(3z^2 - 1\right)}{4\sqrt{\pi}}$$

$$Z_{2,1}(x, y, z) = -\frac{\sqrt{15}xz}{2\sqrt{\pi}}$$

$$Z_{2,2}(x, y, z) = \frac{\sqrt{15}\left(x^2 - y^2\right)}{4\sqrt{\pi}}$$

$$Z_{3,-3}(x, y, z) = \frac{\sqrt{70}y\left(-3x^2 + y^2\right)}{8\sqrt{\pi}}$$

$$Z_{3,-2}(x, y, z) = \frac{\sqrt{105}xyz}{2\sqrt{\pi}}$$

$$Z_{3,-1}(x, y, z) = \frac{\sqrt{42}y\left(-5z^2 + 1\right)}{8\sqrt{\pi}}$$

$$Z_{3,0}(x, y, z) = \frac{\sqrt{7}z\left(5z^2 - 3\right)}{4\sqrt{\pi}}$$

$$Z_{3,1}(x, y, z) = \frac{\sqrt{42}x\left(-5z^2 + 1\right)}{8\sqrt{\pi}}$$

$$Z_{3,2}(x, y, z) = \frac{\sqrt{105}z\left(x^2 - y^2\right)}{4\sqrt{\pi}}$$

$$Z_{3,3}(x, y, z) = \frac{\sqrt{70}x\left(-x^2 + 3y^2\right)}{8\sqrt{\pi}}$$

These tables were generated using `spherical_harmonics.py`:

```
from sympy import (sympify, factorial, var, cos, S, sin, Dummy, sqrt, pi, exp,
        I, latex, symbols)

def Plm(l, m, z):
    """
    Returns the associated Legendre polynomial P_{lm}(z).

    The Condon & Shortley (-1)^m factor is included.
    """
    l = sympify(l)
    m = sympify(m)
    z = sympify(z)
```

(continues on next page)

(continued from previous page)

```
    if m >= 0:
        r = ((z**2-1)**l).diff(z, l+m)
        return (-1)**m * (1-z**2)**(m/2) * r / (2**l * factorial(l))
    else:
        m = -m
        r = ((z**2-1)**l).diff(z, l+m)
        return factorial(l-m)/factorial(l+m) * (1-z**2)**(m/2) * r / (2**l *␣
↪factorial(l))


def Plm_cos(l, m, theta):
    """
    Returns the associated Legendre polynomial P_{lm}(cos(theta)).

    The Condon & Shortley (-1)^m factor is included.
    """
    l = sympify(l)
    m = sympify(m)
    theta = sympify(theta)
    z = Dummy("z")
    r = ((z**2-1)**l).diff(z, l+m).subs(z**2-1, -sin(theta)**2).subs(z, cos(theta))
    return (-1)**m * sin(theta)**m * r / (2**l * factorial(l))

def Ylm(l, m, theta, phi):
    """
    Returns the spherical harmonics Y_{lm}(theta, phi) using the Condon & Shortley␣
↪convention.
    """
    l, m, theta, phi = sympify(l), sympify(m), sympify(theta), sympify(phi)
    return sqrt((2*l+1)/(4*pi) * factorial(l-m)/factorial(l+m)) * Plm_cos(l, m,␣
↪theta) * exp(I*m*phi)

def Zlm(l, m, theta, phi):
    """
    Returns the real spherical harmonics Z_{lm}(theta, phi).
    """
    l, m, theta, phi = sympify(l), sympify(m), sympify(theta), sympify(phi)
    if m > 0:
        return sqrt((2*l+1)/(2*pi) * factorial(l-m)/factorial(l+m)) * Plm_cos(l, m,␣
↪theta) * cos(m*phi)
    elif m < 0:
        m = -m
        return sqrt((2*l+1)/(2*pi) * factorial(l-m)/factorial(l+m)) * Plm_cos(l, m,␣
↪theta) * sin(m*phi)
    elif m == 0:
        return sqrt((2*l+1)/(4*pi)) * Plm_cos(l, 0, theta)
    else:
        raise ValueError("Invalid m.")

def Zlm_xyz(l, m, x, y, z):
    """
    Returns the real spherical harmonics Z_{lm}(x, y, z).

    It is assumed x**2 + y**2 + z**2 == 1.
    """
    l, m, x, y, z = sympify(l), sympify(m), sympify(x), sympify(y), sympify(z)
    if m > 0:
```

(continues on next page)

(continued from previous page)

```
        r = (x+I*y)**m
        r = r.as_real_imag()[0]
        return sqrt((2*l+1)/(2*pi) * factorial(l-m)/factorial(l+m)) * Plm(l, m, z) *␣
↪r / sqrt(1-z**2)**m
    elif m < 0:
        m = -m
        r = (x+I*y)**m
        r = r.as_real_imag()[1]
        return sqrt((2*l+1)/(2*pi) * factorial(l-m)/factorial(l+m)) * Plm(l, m, z) *␣
↪r / sqrt(1-z**2)**m
    elif m == 0:
        return sqrt((2*l+1)/(4*pi)) * Plm(l, 0, z)
    else:
        raise ValueError("Invalid m.")


var("theta phi")
x, y, z = symbols("x y z", real=True)
print "Spherical harmonics:"
print
print ".. math::"
print
for l in range(4):
    for m in range(-l, l+1):
        print r"    Y_{%d,%d}(\theta, \phi) =" % (l, m), \
            latex(Ylm(l, m, theta, phi))
        print

print
print "Real spherical harmonics:"
print
print ".. math::"
print
for l in range(4):
    for m in range(-l, l+1):
        print r"    Z_{%d,%d}(\theta, \phi) =" % (l, m), \
            latex(Zlm(l, m, theta, phi))
        print

print
print "Real spherical harmonics (using $x$, $y$ and $z$, assuming $x^2 + y^2 + z^2 = 1
↪$):"
print
print ".. math::"
print
for l in range(4):
    for m in range(-l, l+1):
        print r"    Z_{%d,%d}(x, y, z) =" % (l, m), \
            latex(Zlm_xyz(l, m, x, y, z).simplify())
        print
```

3.42 Variational Formulation of PDEs

Not every equation allows a variational formulation (e.g., Navier-Stokes or Euler equations do not have such a formulation), but many equations have one, and we explain how it works on several examples.

3.42.1 Poisson Equation

The Lagrangian for Poisson equation is:

$$L[u] = \int_a^b \left[\tfrac{1}{2}u'^2(x) - f(x)u(x)\right] \mathrm{d}x - [g(x)u(x)]_a^b \,. \tag{3.42.1.1}$$

Important note: technically, as we will see below, this imposes the Neumann boundary condition and 1D Poisson equation with two Neumann boundary conditions does not have a unique solution. At least one Dirichlet boundary condition is needed for a unique solution. For example with $u(a) = u_0$ and $u'(b) = g$ the boundary term becomes just $-gu(b)$. However, for simplicity, we will show the derivation with two Neumann boundary conditions first and we will discuss how to impose the Dirichlet boundary condition later.

The variational formulation is:

$$\delta L = 0 \,,$$

which yields:

$$\begin{aligned}
\delta L &= \int_a^b [u'(x)\delta u'(x) - f(x)\delta u(x)] \,\mathrm{d}x - [g(x)\delta u(x)]_a^b \\
&= \int_a^b [-u''(x)\delta u(x) - f(x)\delta u(x)] \,\mathrm{d}x + [u'(x)\delta u(x)]_a^b - [g(x)\delta u(x)]_a^b \\
&= \int_a^b [-u''(x) - f(x)] \,\delta u(x) \mathrm{d}x + [(u'(x) - g(x))\delta u(x)]_a^b \\
&= 0 \,,
\end{aligned} \tag{3.42.1.2}$$

where we applied integration by parts. This equation holds for any $\delta u(x)$, and in particular it holds for $\delta u(x) = 0$ at the boundary (i.e., for $\delta u(a) = 0$ and $\delta u(b) = 0$). Then the boundary term in (3.42.1.2) vanishes and we obtain:

$$\int_a^b [-u''(x) - f(x)] \,\delta u(x) \mathrm{d}x = 0 \,, \tag{3.42.1.3}$$

This equation holds for any $\delta u(x)$ that is zero at the boundary, and thus it implies:

$$u''(x) + f(x) = 0 \,. \tag{3.42.1.4}$$

Now we substitute (3.42.1.4) into (3.42.1.2) and obtain:

$$[(u'(x) - g(x))\delta u(x)]_a^b = 0 \,. \tag{3.42.1.5}$$

Thus (3.42.1.2) implies both (3.42.1.4) and (3.42.1.5). The equation (3.42.1.5) holds for any $\delta u(x)$ (generally not zero at the boundary) and thus it implies:

$$u'(x) - g(x) = 0 \tag{3.42.1.6}$$

at the boundary. Thus $g(x)$ imposes the Neumann boundary condition, i.e., the value of the derivative $u'(x) = g(x)$ at the boundary. This condition is imposed variationally.

To impose a Dirichlet boundary condition, we want to impose the value of $u(x) = u_0(x)$ at the boundary for some constant $u_0(x)$. As such, $u(x)$ is not allowed to vary at that part of the boundary, which means that the variation $\delta u(x) = 0$ at the boundary. So we restrict the variation $\delta u(x)$ to be zero at the Dirichlet part of the boundary in (3.42.1.2) and thus also in (3.42.1.5). This implies that (3.42.1.6) does not hold at the Dirichlet part of the boundary and we have to set the value $u(x)$ there directly.

Example

As a particular example, let $u(a) = u_0$ and $u'(b) = g$. Then the Lagrangian (3.42.1.1) becomes:

$$L[u] = \int_a^b \left[\tfrac{1}{2}u'^2(x) - f(x)u(x)\right] \mathrm{d}x - gu(b)\,. \tag{3.42.1.7}$$

We can explicitly define the space U of all trial functions $u \in U$ that one can choose (admissible) and substitute in (3.42.1.7) as follows. We have to impose the Dirichlet condition $u(a) = u_0$ on the space itself, and we also have to choose how smooth functions we want. For finite element applications one typically chooses H^1 (i.e., values and first derivatives are from L^2) and we obtain:

$$U := \{u : u \in H^1(a,b), u(a) = u_0\} \tag{3.42.1.8}$$

Now we derive what space the variation $\delta u(x)$ belongs to. Let u_{min} be the solution (the extremum of the functional (3.42.1.7)). Then from calculus of variations:

$$u = u_{\text{min}} + \varepsilon\delta u(x) \tag{3.42.1.9}$$

Here u is called the trial function and $\delta u(x)$ is called the test function. Both u and u_{min} are from the space U. Thus we can compute:

$$\delta u(a) = \frac{u(a) - u_{\text{min}}(a)}{\varepsilon} = \frac{u_0 - u_0}{\varepsilon} = 0\,.$$

In addition, both $u, u_{\text{min}} \in H^1(a,b)$, so also their difference $u(x) - u_{\text{min}}(x)$ and thus also $\delta u(x) = \frac{u(x)-u_{\text{min}}(x)}{\varepsilon}$ is from $H^1(a,b)$. There are no other conditions ($u(b)$ and $u_{\text{min}}(b)$ are generally different, so in general $\delta u(b) \neq 0$) and so $\delta u(x) \in U_0$ where the space U_0 is:

$$U_0 := \{w : w \in H^1(a,b), w(a) = 0\}\,. \tag{3.42.1.10}$$

The definition of the space U_0 in (3.42.1.10) is derived from the definition of the space U in (3.42.1.8).

To compute the variation of L, we substitute (3.42.1.9) into (3.42.1.7), differentiate with respect to ε and then set $\varepsilon = 0$ using (3.17.3.2):

$$\delta L[u] = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} L[u_{\text{min}} + \varepsilon\delta u]\bigg|_{\varepsilon=0}$$

as was done in (3.42.1.2) and one obtains the weak form (below we drop the label min from u_{min} and just use u):

$$\delta L[u] = \int_a^b [u'(x)\delta u'(x) - f(x)\delta u(x)]\,\mathrm{d}x - g\delta u(b) = 0\,. \tag{3.42.1.11}$$

The task is to find such function $u \in U$ so that (3.42.1.11) holds for all $\delta u \in U_0$. From (3.42.1.11) one obtains (as in (3.42.1.2)):

$$\int_a^b [-u''(x) - f(x)]\,\delta u(x)\mathrm{d}x + (u'(b) - g)\delta u(b) = 0\,. \tag{3.42.1.12}$$

The governing equation (3.42.1.4) is the same:

$$u''(x) + f(x) = 0\,. \tag{3.42.1.13}$$

The boundary term (3.42.1.5) becomes (see (3.42.1.12)):

$$(u'(b) - g)\delta u(b) = 0\,.$$

Which implies $u'(b) = g$.

The Dirichlet boundary condition is part of the definition of the function space (3.42.1.8), so all trial functions u that one can choose (admissible) and substitute in $L[u]$ must lie in U. From the derivation of the space U_0 in (3.42.1.10) we can see that since the value of $u(a)$ is fixed, we always have $\delta u(a) = 0$; on the other hand, since $u(b)$ is not fixed, in general we have $\delta u(b) \neq 0$.

The Neumann boundary condition is imposed variationally due to the surface term in the weak form (3.42.1.11).

Summary

We have shown above that there are three equivalent formulations which fully and uniquely determine the solution and boundary conditions (both Dirichlet and Neumann):

1. Define the functional $L[u]$ in (3.42.1.7) and the space U for the trial functions $u \in U$ in (3.42.1.8).
2. Define the weak form (3.42.1.11) and the two spaces U and U_0, where $u \in U$ and $\delta u \in U_0$.
3. Define the strong form (3.42.1.13) and the boundary conditions $u(a) = u_0$ and $u'(b) = g$.

Let us write down the three formulations in detail.

Variational Formulation

The variational formulation is the formulation 1. above.

$$L[u] = \int_a^b \left[\tfrac{1}{2}u'^2(x) - f(x)u(x)\right] \mathrm{d}x - gu(b)\,.$$

The task is to find such $u \in U$ that extremizes this functional ($\delta L[u] = 0$), where:

$$U := \{u : u \in H^1(a,b), u(a) = u_0\}\,.$$

Weak Formulation

Weak formulation is the formulation 2. above, and it is customary to write $w(x) \equiv \delta u(x)$ in the weak form (3.42.1.11):

$$\int_a^b [u'(x)w'(x) - f(x)w(x)]\,\mathrm{d}x - gw(b) = 0\,. \tag{3.42.1.14}$$

The task is to find such $u \in U$ so that (3.42.1.14) holds for all $w \in U_0$, where

$$U := \{u : u \in H^1(a,b), u(a) = u_0\}\,,$$
$$U_0 := \{w : w \in H^1(a,b), w(a) = 0\}\,.$$

We can also define:

$$a(u,w) = \int_a^b u'(x)w'(x)\mathrm{d}x\,,$$
$$b(w) = \int_a^b f(x)w(x)\mathrm{d}x + gw(b)$$

and write (3.42.1.14) as:

$$a(u,w) = b(w)\,.$$

Strong Formulation

Strong formulation is the formulation 3. above. We are solving the equation:

$$u''(x) + f(x) = 0$$

subject to boundary conditions $u(a) = u_0$ and $u'(b) = g$.

3.42.2 Radial Schrödinger Equation

The derivation is similar as for the Poisson equation, except that we have $g(x) = 0$ based on physical reasoning (that we cannot set the derivative to a given value, or, alternatively, that we require the operator to be self-adjoint).

The Lagrangian for the radial Schrödinger equation is:

$$L[R] = \int_0^\infty \left[\tfrac{1}{2}R'^2(r) + \left(V(r) + \frac{l(l+1)}{2r^2}\right)R^2(r)\right] r^2 \,\mathrm{d}r\,. \tag{3.42.2.1}$$

We minimize the Lagrangian subject to the normalization condition $N[R] = \int_0^\infty R^2(r)r^2 \,\mathrm{d}r = 1$ as follows:

$$\begin{aligned} 0 &= \delta(L - \epsilon(N-1)) \\ &= \delta \int_0^\infty \left[\tfrac{1}{2}r^2R'^2 + (r^2V + \tfrac{1}{2}l(l+1))R^2 - \epsilon r^2R^2\right] \,\mathrm{d}r = \\ &= 2\int_0^\infty \left[\tfrac{1}{2}r^2R'(\delta R)' + (r^2V + \tfrac{1}{2}l(l+1))R\delta R - \epsilon r^2 R\delta R\right] \,\mathrm{d}r = \\ &= 2\int_0^\infty \left[-\tfrac{1}{2}(r^2R')' + (r^2V + \tfrac{1}{2}l(l+1))R - \epsilon r^2 R\right]\delta R \,\mathrm{d}r + [r^2R'\delta R]_0^\infty \end{aligned} \tag{3.42.2.2}$$

This equation holds for any $\delta R(r)$, and so it also holds when we restrict $\delta R(r) = 0$ on the boundary and the boundary term vanishes. Then it implies the radial Schrödinger equation:

$$-\tfrac{1}{2}(r^2R'(r))' + (r^2V(r) + \tfrac{1}{2}l(l+1))R(r) = \epsilon r^2 R(r) \tag{3.42.2.3}$$

Substituting (3.42.2.3) into (3.42.2.2) we obtain:

$$[r^2R'\delta R]_0^\infty = 0 \tag{3.42.2.4}$$

And we can see that (3.42.2.2) implies both the equation (3.42.2.3) and the boundary term (3.42.2.4). The boundary term is zero for $r = 0$, so it reduces to:

$$\lim_{r\to\infty} r^2R'(r)\delta R(r) = 0 \tag{3.42.2.5}$$

We can see that there is no natural condition at $r = 0$, and for $r = \infty$ we only have two possible options. Either we impose $\delta R(\infty) = 0$ and obtain the Dirichlet condition and the boundary term (3.42.2.5) vanishes. Or we allow $\delta R(\infty)$ to vary, and then (3.42.2.5) implies $R'(\infty) = 0$.

Unlike for the Poisson equation we are not allowed to set $R'(\infty)$ to anything other than zero, and that's why (3.42.2.1) has no surface term.

CHAPTER

FOUR

CLASSICAL MECHANICS, SPECIAL AND GENERAL RELATIVITY

4.1 Gravitation and Electromagnetism as a Field Theory

The action for macroscopic gravity, electromagnetism and (possibly) charged relativistic dust is:

$$S = S_H + S_M + S_{EM} + S_q$$

where:

$$\begin{aligned} S_H[g^{\mu\nu}] &= \frac{c^4}{16\pi G}\int R\sqrt{-g}\mathrm{d}^4x \\ S_M[g^{\mu\nu}, x^\mu] &= -c\int \rho\sqrt{v_\mu v^\mu}\sqrt{-g}\mathrm{d}^4x \\ S_{EM}[g^{\mu\nu}, A^\mu] &= -\frac{1}{4\mu_0}\int F_{\alpha\beta}F^{\alpha\beta}\sqrt{-g}\mathrm{d}^4x \\ S_q[x^\mu, A^\mu] &= -\int \rho_{EM} v^\mu A_\mu\sqrt{-g}\mathrm{d}^4x \end{aligned}$$

where x^μ is the field of the matter, A^μ is the electromagnetic field and $g^{\mu\nu}$ is the gravitational field. We vary with respect to each of them to obtain (interacting) equations of motion. c is the speed of light, G is the gravitational constant, μ_0 the permeability of vacuum. ρ is the mass density of the dust, ρ_{EM} is the charge density of the dust, $v^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}$ is 4-velocity of the dust, $F_{\alpha\beta} = \nabla_\alpha A_\beta - \nabla_\beta A_\alpha$ is the electromagnetic field tensor, R is the Ricci scalar.

4.1.1 Gravitation

We vary with respect to $g^{\mu\nu}$. By changing the metric, we also change the invariant volume element (thus also ρ), so we need to be careful to vary properly. We start with S_H:

$$\begin{aligned} \delta S_H &= \delta\frac{c^4}{16\pi G}\int R\sqrt{-g}\mathrm{d}^4x = \\ &= \frac{c^4}{16\pi G}\int (\delta g^{\mu\nu})R_{\mu\nu}\sqrt{-g} + g^{\mu\nu}(\delta R_{\mu\nu})\sqrt{-g} + R(\delta\sqrt{-g})\mathrm{d}^4x = \\ &= \frac{c^4}{16\pi G}\int (\delta g^{\mu\nu})R_{\mu\nu}\sqrt{-g} + g^{\mu\nu}\left(\nabla_\lambda(\delta\Gamma^\lambda_{\nu\mu}) - \nabla_\nu(\delta\Gamma^\lambda_{\lambda\mu})\right)\sqrt{-g} + R(-\tfrac{1}{2}\sqrt{-g}\, g_{\mu\nu}(\delta g^{\mu\nu}))\mathrm{d}^4x = \\ &= \frac{c^4}{16\pi G}\int (\delta g^{\mu\nu})R_{\mu\nu}\sqrt{-g} + \left(\nabla_\lambda g^{\mu\nu}(\delta\Gamma^\lambda_{\nu\mu}) - \nabla_\nu g^{\mu\nu}(\delta\Gamma^\lambda_{\lambda\mu})\right)\sqrt{-g} - \tfrac{1}{2}Rg_{\mu\nu}\sqrt{-g}\,(\delta g^{\mu\nu})\mathrm{d}^4x = \\ &= \frac{c^4}{16\pi G}\int (\delta g^{\mu\nu})R_{\mu\nu}\sqrt{-g} - \tfrac{1}{2}Rg_{\mu\nu}\sqrt{-g}\,(\delta g^{\mu\nu})\mathrm{d}^4x = \\ &= \frac{c^4}{16\pi G}\int \left(R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu}\right)(\delta g^{\mu\nu})\sqrt{-g}\mathrm{d}^4x \end{aligned}$$

Variation of S_M is:

$$\begin{aligned}
\delta S_M = -c\delta \int \rho\sqrt{v_\mu v^\mu}\sqrt{-g}\mathrm{d}^4x = \\
= -\delta \int c\sqrt{\mathfrak{p}_\mu \mathfrak{p}^\mu}\mathrm{d}^4x = \\
= -\int c\frac{\delta(g^{\mu\nu}\mathfrak{p}_\mu\mathfrak{p}_\nu)}{2\sqrt{\mathfrak{p}_\alpha\mathfrak{p}^\alpha}}\mathrm{d}^4x = \\
= -\int c\frac{\mathfrak{p}_\mu\mathfrak{p}_\nu}{2\sqrt{\mathfrak{p}_\alpha\mathfrak{p}^\alpha}}\delta(g^{\mu\nu})\mathrm{d}^4x = \\
= -\int c\frac{\rho v_\mu \rho v_\nu \sqrt{-g}^2}{2\rho c\sqrt{-g}}\delta(g^{\mu\nu})\mathrm{d}^4x = \\
= -\int \tfrac{1}{2}\rho v_\mu v_\nu \delta(g^{\mu\nu})\sqrt{-g}\mathrm{d}^4x
\end{aligned}$$

The variation of S_{EM} is:

$$\begin{aligned}
\delta S_{EM} = -\delta \int \frac{1}{4\mu_0}F_{\alpha\beta}F^{\alpha\beta}\sqrt{-g}\mathrm{d}^4x = \\
= -\delta \int \frac{1}{4\mu_0}g^{\alpha\lambda}g^{\beta\rho}F_{\alpha\beta}F_{\lambda\rho}\sqrt{-g}\mathrm{d}^4x = \\
= -\frac{1}{4\mu_0}\int \left(\delta(g^{\alpha\lambda}g^{\beta\rho})F_{\alpha\beta}F_{\lambda\rho}\sqrt{-g} + g^{\alpha\mu}g^{\beta\rho}F_{\alpha\beta}F_{\lambda\rho}\left(\delta\sqrt{-g}\right)\right)\mathrm{d}^4x = \\
= -\frac{1}{4\mu_0}\int \left(2(\delta g^{\alpha\lambda})g^{\beta\rho}F_{\alpha\beta}F_{\lambda\rho}\sqrt{-g} + g^{\alpha\lambda}g^{\beta\rho}F_{\alpha\beta}F_{\lambda\rho}\left(-\tfrac{1}{2}\sqrt{-g}g_{\mu\nu}(\delta g^{\mu\nu})\right)\right)\mathrm{d}^4x = \\
= -\frac{1}{4\mu_0}\int \left(2(\delta g^{\alpha\lambda})F_{\alpha\beta}F_\lambda{}^\beta - \tfrac{1}{2}F_{\alpha\beta}F^{\alpha\beta}g_{\mu\nu}(\delta g^{\mu\nu})\right)\sqrt{-g}\mathrm{d}^4x = \\
= -\frac{1}{2\mu_0}\int \left(F_{\mu\beta}F_\nu{}^\beta - \frac{1}{4}F_{\alpha\beta}F^{\alpha\beta}g_{\mu\nu}\right)(\delta g^{\mu\nu})\sqrt{-g}\mathrm{d}^4x
\end{aligned}$$

The variation of $\delta S_q = 0$.

The equations of motion are:

$$\frac{c^4}{16\pi G}\left(R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu}\right) - \tfrac{1}{2}\rho v_\mu v_\nu - \frac{1}{2\mu_0}\left(F_{\mu\beta}F_\nu{}^\beta - \frac{1}{4}F_{\alpha\beta}F^{\alpha\beta}g_{\mu\nu}\right) = 0$$

We rearrange:

$$R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu} = \frac{8\pi G}{c^4}\rho v_\mu v_\nu + \frac{8\pi G}{c^4}\frac{1}{\mu_0}\left(F_{\mu\beta}F_\nu{}^\beta - \frac{1}{4}F_{\alpha\beta}F^{\alpha\beta}g_{\mu\nu}\right)$$

We define the stress energy tensor as:

$$R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu} \tag{4.1.1.1}$$

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}}\frac{\delta(S_M + S_{EM} + S_q)}{\delta g^{\mu\nu}} \tag{4.1.1.2}$$

And we get:

$$\begin{aligned}
T_{\mu\nu} = T^M_{\mu\nu} + T^{EM}_{\mu\nu} \\
T^M_{\mu\nu} = \rho v_\mu v_\nu \\
T^{EM}_{\mu\nu} = \frac{1}{\mu_0}\left(F_{\mu\beta}F_\nu{}^\beta - \frac{1}{4}F_{\alpha\beta}F^{\alpha\beta}g_{\mu\nu}\right)
\end{aligned} \tag{4.1.1.3}$$

The equations (4.1.1.1) are called Einstein's equations and the equations (4.1.1.3) are stress energy tensors for the relativistic dust and electromagnetism. The equation (4.1.1.2) is the stress energy tensor corresponding to the given action. Sometimes it is not possible to write an action for more complex matter (perfect fluid, Navier-Stokes equations for fluid, ...) in which case we cannot use (4.1.1.2), but we can still specify the stress energy tensor directly and (4.1.1.1) are the equations of motion.

4.1.2 Electromagnetism

We vary with respect to A^μ. The variation of $\delta S_H = 0$. The variation of $\delta S_M = 0$. The variation of S_{EM} is:

$$\begin{aligned}
\delta S_{EM} = \\
= -\frac{1}{4\mu_0}\delta \int F_{\mu\nu}F^{\mu\nu}\sqrt{-g}\mathrm{d}^4x = \\
= -\frac{1}{2\mu_0}\int F^{\mu\nu}(\delta F_{\mu\nu})\sqrt{-g}\mathrm{d}^4x = \\
= -\frac{1}{\mu_0}\int F^{\mu\nu}(\delta\partial_\nu A_\mu)\sqrt{-g}\mathrm{d}^4x = \\
= -\frac{1}{\mu_0}\int F^{\mu\nu}\partial_\nu(\delta A_\mu)\sqrt{-g}\mathrm{d}^4x = \\
= \frac{1}{\mu_0}\int \partial_\nu(F^{\mu\nu}\sqrt{-g})(\delta A_\mu)\mathrm{d}^4x = \\
= \frac{1}{\mu_0}\int \left(\frac{1}{\sqrt{-g}}\partial_\nu(F^{\mu\nu}\sqrt{-g})\right)(\delta A_\mu)\sqrt{-g}\mathrm{d}^4x = \\
= \frac{1}{\mu_0}\int \nabla_\mu F^{\mu\nu}(\delta A_\nu)\sqrt{-g}\mathrm{d}^4x
\end{aligned}$$

The variation of S_q is:

$$\begin{aligned}
\delta S_q = \\
= -\delta\int \rho_{EM}v^\nu A_\nu\sqrt{-g}\mathrm{d}^4x = \\
= -\int \rho_{EM}v^\nu(\delta A_\nu)\sqrt{-g}\mathrm{d}^4x =
\end{aligned}$$

The equation of motion is:

$$\frac{1}{\mu_0}\nabla_\mu F^{\mu\nu} - \rho_{EM}v^\nu = 0$$

Rearranging:

$$\nabla_\mu F^{\mu\nu} = \mu_0\rho_{EM}v^\nu$$

4.1.3 Relativistic Dust

We vary the whole action with respect to x^μ. The variation of $\delta S_H = 0$. The variation of S_M is:

$$\begin{aligned}
\delta S_M &= -c\delta \int \rho \sqrt{v_\mu v^\mu} \sqrt{-g} \mathrm{d}^4 x = \\
&= -\delta \int c \sqrt{\mathfrak{p}_\mu \mathfrak{p}^\mu} \mathrm{d}^4 x = \\
&= -\int c \frac{\delta(g^{\mu\nu} \mathfrak{p}_\mu \mathfrak{p}_\nu)}{2\sqrt{\mathfrak{p}_\alpha \mathfrak{p}^\alpha}} \mathrm{d}^4 x = \\
&= -\int c \frac{2 g^{\mu\nu} \mathfrak{p}_\mu (\delta \mathfrak{p}_\nu)}{2\sqrt{\mathfrak{p}_\alpha \mathfrak{p}^\alpha}} \mathrm{d}^4 x = \\
&= -\int c \frac{\mathfrak{p}_\mu}{\sqrt{\mathfrak{p}_\alpha \mathfrak{p}^\alpha}} (\delta \mathfrak{p}^\mu) \mathrm{d}^4 x = \\
&= -\int c \frac{\mathfrak{p}_\mu}{\sqrt{\mathfrak{p}_\alpha \mathfrak{p}^\alpha}} \partial_\nu \left(\mathfrak{p}^\nu (\delta x^\mu) - \mathfrak{p}^\mu (\delta x^\nu) \right) \mathrm{d}^4 x = \\
&= \int c \partial_\nu \left(\frac{\mathfrak{p}_\mu}{\sqrt{\mathfrak{p}_\alpha \mathfrak{p}^\alpha}} \right) \left(\mathfrak{p}^\nu (\delta x^\mu) - \mathfrak{p}^\mu (\delta x^\nu) \right) \mathrm{d}^4 x = \\
&= \int c \left(\partial_\nu \left(\frac{\mathfrak{p}_\mu}{\sqrt{\mathfrak{p}_\alpha \mathfrak{p}^\alpha}} \right) - \partial_\mu \left(\frac{\mathfrak{p}_\nu}{\sqrt{\mathfrak{p}_\alpha \mathfrak{p}^\alpha}} \right) \right) \mathfrak{p}^\nu (\delta x^\mu) \mathrm{d}^4 x = \\
&= \int c \left(\nabla_\nu \left(\frac{\mathfrak{p}_\mu}{\sqrt{\mathfrak{p}_\alpha \mathfrak{p}^\alpha}} \right) - \nabla_\mu \left(\frac{\mathfrak{p}_\nu}{\sqrt{\mathfrak{p}_\alpha \mathfrak{p}^\alpha}} \right) \right) \mathfrak{p}^\nu (\delta x^\mu) \mathrm{d}^4 x = \\
&= \int \left(\nabla_\nu v_\mu - \nabla_\mu v_\nu \right) \rho v^\nu (\delta x^\mu) \sqrt{-g} \mathrm{d}^4 x = \\
&= \int \rho (\nabla_\nu v_\mu) v^\nu (\delta x^\mu) \sqrt{-g} \mathrm{d}^4 x
\end{aligned}$$

The variation of $\delta S_{EM} = 0$. The variation of S_q is:

$$\begin{aligned}
\delta S_q &= -\delta \int \rho_{EM} v^\mu A_\mu \sqrt{-g} \mathrm{d}^4 x = \\
&= -\delta \int \mathfrak{j}^\mu A_\mu \mathrm{d}^4 x = \\
&= -\int (\delta \mathfrak{j}^\mu) A_\mu \mathrm{d}^4 x = \\
&= -\int \partial_\nu \left(\mathfrak{j}^\nu (\delta x^\mu) - \mathfrak{j}^\mu (\delta x^\nu) \right) A_\mu \mathrm{d}^4 x = \\
&= \int \left(\mathfrak{j}^\nu (\delta x^\mu) - \mathfrak{j}^\mu (\delta x^\nu) \right) \partial_\nu A_\mu \mathrm{d}^4 x = \\
&= \int \mathfrak{j}^\nu (\delta x^\mu) (\partial_\nu A_\mu - \partial_\mu A_\nu) \mathrm{d}^4 x = \\
&= \int \rho_{EM} v^\nu (\nabla_\nu A_\mu - \nabla_\mu A_\nu) (\delta x^\mu) \sqrt{-g} \mathrm{d}^4 x = \\
&= -\int \rho_{EM} v^\nu F_{\mu\nu} (\delta x^\mu) \sqrt{-g} \mathrm{d}^4 x
\end{aligned}$$

The equation of motion is:

$$\rho (\nabla_\nu v_\mu) v^\nu - \rho_{EM} v^\nu F_{\mu\nu} = 0$$

Rearranging:

$$\rho (\nabla_\nu v_\mu) v^\nu = \rho_{EM} v^\nu F_{\mu\nu}$$

This is the geodesic equation with Lorentz force.

4.1.4 Equations of Motion

All together, the equations of motion are:

$$R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu} = \frac{8\pi G}{c^4}\rho v_\mu v_\nu + \frac{8\pi G}{c^4}\frac{1}{\mu_0}\left(F_{\mu\beta}F_\nu{}^\beta - \frac{1}{4}F_{\alpha\beta}F^{\alpha\beta}g_{\mu\nu}\right)$$
$$\nabla_\mu F^{\mu\nu} = \mu_0\rho_{EM}v^\nu$$
$$\rho(\nabla_\nu v_\mu)v^\nu = \rho_{EM}v^\nu F_{\mu\nu}$$

The first equation determines $g_{\mu\nu}$ from the given sources (the stress energy tensors) on the right hand side, that depend on ρ, v^μ, A^μ and $g_{\mu\nu}$. The second equation determines A^μ from the sources (ρ_{EM} and v^μ) and from $g_{\mu\nu}$ (through the covariant derivative). Finally, the last equation determines x^μ and v^μ from the given fields A^μ (through the electromagnetic field tensor) and $g_{\mu\nu}$ (through the covariant derivative).

Conservation

We apply covariant 4-divergence and use Bianci identities on the first equation:

$$0 = \nabla_\mu T^{\mu\nu} = \nabla_\mu(T_M^{\mu\nu} + T_{EM}^{\mu\nu})$$

So the total stress energy tensor is conserved. This fact makes the equations of motion (that follow from the action principle) not all independent. The third equation can be derived from the fist two as follows.

We calculate:

$$\nabla_\mu T_M^{\mu\nu} = \nabla_\mu(\rho v^\mu v^\nu)$$
$$\nabla_\mu T_{EM}^{\mu\nu} = F^{\alpha\nu}\rho_{EM}v_\alpha$$

and we get:

$$\nabla_\mu(\rho v^\mu v^\nu) + F^{\alpha\nu}\rho_{EM}v_\alpha = 0$$
$$\nabla_\mu(\rho v^\mu)v^\nu + \rho v^\mu\nabla_\mu v^\nu + F^{\alpha\nu}\rho_{EM}v_\alpha = 0$$

The first term vanishes, because:

$$v_\nu\nabla_\mu(\rho v^\mu)v^\nu + v_\nu\rho v^\mu\nabla_\mu v^\nu + v_\nu F^{\alpha\nu}\rho_{EM}v_\alpha = 0$$
$$v_\nu\nabla_\mu(\rho v^\mu)v^\nu + v_\nu F^{\alpha\nu}\rho_{EM}v_\alpha = 0$$
$$c^2\nabla_\mu(\rho v^\mu) + v_\nu F^{\alpha\nu}\rho_{EM}v_\alpha = 0$$
$$c^2\nabla_\mu(\rho v^\mu) = 0$$

where we used $v_\nu\nabla_\mu v^\nu = 0$ (follows from differentiating $c^2 = v_\nu v^\nu$) and $v_\nu F^{\alpha\nu}v_\alpha = 0$ (contracting symmetric and antisymmetric tensors). We are left with:

$$\rho v^\mu\nabla_\mu v^\nu + F^{\alpha\nu}\rho_{EM}v_\alpha = 0$$
$$\rho v^\mu\nabla_\mu v^\nu = -F^{\alpha\nu}\rho_{EM}v_\alpha$$
$$\rho v^\mu\nabla_\mu v^\nu = F^{\nu\alpha}\rho_{EM}v_\alpha$$

Which is the third equation.

4.2 Classical Mechanics

4.2.1 Rigid Body Rotation

In all the sections below, the rigid body is rotating around the $\boldsymbol{\omega}$ axis, so:

$$\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$$

Kinetic Energy

The kinetic energy is:

$$\begin{aligned} T &= \int \tfrac{1}{2}\rho(\mathbf{r})v^2 \mathrm{d}^3 r = \\ &= \int \tfrac{1}{2}\rho(\mathbf{r})\mathbf{v}\cdot\mathbf{v}\mathrm{d}^3 r = \\ &= \int \tfrac{1}{2}\rho(\mathbf{r})\mathbf{v}\cdot(\boldsymbol{\omega}\times\mathbf{r})\mathrm{d}^3 r = \\ &= \int \tfrac{1}{2}\rho(\mathbf{r})\boldsymbol{\omega}\cdot(\mathbf{r}\times\mathbf{v})\mathrm{d}^3 r = \\ &= \tfrac{1}{2}\boldsymbol{\omega}\cdot\int \rho(\mathbf{r})(\mathbf{r}\times\mathbf{v})\mathrm{d}^3 r = \\ &= \tfrac{1}{2}\boldsymbol{\omega}\cdot\mathbf{L} \end{aligned}$$

where $\mathbf{L}$ is the total angular momentum:

$$\mathbf{L} = \int \rho(\mathbf{r})(\mathbf{r}\times\mathbf{v})\mathrm{d}^3 r$$

Angular Momentum

Total angular momentum is:

$$\begin{aligned} \mathbf{L} &= \int \rho(\mathbf{r})(\mathbf{r}\times\mathbf{v})\mathrm{d}^3 r = \\ &= \int \rho(\mathbf{r})(\mathbf{r}\times(\boldsymbol{\omega}\times\mathbf{r}))\mathrm{d}^3 r = \\ &= \int \rho(\mathbf{r})(\boldsymbol{\omega}r^2 - \mathbf{r}(\mathbf{r}\cdot\boldsymbol{\omega}))\mathrm{d}^3 r = \\ &= \int \rho(\mathbf{r})(\mathbb{1}r^2 - \mathbf{r}\mathbf{r})\mathrm{d}^3 r\cdot\boldsymbol{\omega} = \\ &= \mathbf{I}\cdot\boldsymbol{\omega} \end{aligned}$$

Where $\mathbf{I}$ is the moment of inertia tensor:

$$\mathbf{I} = \int \rho(\mathbf{r})(\mathbb{1}r^2 - \mathbf{r}\mathbf{r})\mathrm{d}^3 r$$

Moment of Inertia

The moment of inertia tensor and its components are:

$$\mathbf{I} = \int \rho(\mathbf{r})(\mathbb{1}r^2 - \mathbf{r}\mathbf{r})\mathrm{d}^3r$$
$$I^{ij} = \int \rho(\mathbf{r})(\delta^{ij}r_k r^k - r^i r^j)\mathrm{d}^3r$$

Let's write $\boldsymbol{\omega} = \omega\mathbf{n}$ (where $\mathbf{n}$ is a unit vector), then the kinetic energy is:

$$T = \tfrac{1}{2}\boldsymbol{\omega}\cdot\mathbf{L} = \tfrac{1}{2}\boldsymbol{\omega}\cdot\mathbf{I}\cdot\boldsymbol{\omega} = \tfrac{1}{2}\mathbf{n}\cdot\mathbf{I}\cdot\mathbf{n}\,\omega^2 = \tfrac{1}{2}I\omega^2$$

where I is the moment of inertia about the axis of rotation:

$$I = \mathbf{n}\cdot\mathbf{I}\cdot\mathbf{n} =$$
$$= \mathbf{n}\cdot\int \rho(\mathbf{r})(\mathbb{1}r^2 - \mathbf{r}\mathbf{r})\mathrm{d}^3r\cdot\mathbf{n} =$$
$$= \int \rho(\mathbf{r})(r^2 - (\mathbf{r}\cdot\mathbf{n})^2)\mathrm{d}^3r$$

Cylinder

Solid cylinder of radius R, height h and mass m. We'll use cylindrical coordinates. First for rotation about the z axis:

$$V = \pi R^2 h$$
$$\mathbf{n} = (0,0,1)$$
$$\mathbf{r} = (\rho\cos\phi, \rho\sin\phi, z)$$
$$\mathbf{r}\cdot\mathbf{n} = z$$
$$r^2 = \rho^2 + z^2$$
$$I = \int \rho(\mathbf{r})(r^2 - (\mathbf{r}\cdot\mathbf{n})^2)\mathrm{d}^3r = \int \frac{m}{V}(\rho^2 + z^2 - z^2)\mathrm{d}^3r =$$
$$= \int \frac{m}{V}\rho^2\mathrm{d}^3r = \frac{m}{V}\int_0^{2\pi}\mathrm{d}\phi\int_0^R\mathrm{d}R\int_{-\frac{h}{2}}^{\frac{h}{2}}\mathrm{d}z\rho^2\rho =$$
$$= \frac{m}{V}2\pi\frac{R^4}{4}h = \frac{m}{\pi R^2 h}2\pi\frac{R^4}{4}h = \tfrac{1}{2}mR^2$$

Code:

```
>>> from sympy import var, integrate, pi
>>> var("m V R rho z phi h")
(m, V, R, rho, z, phi, h)
>>> I = m/V * integrate(rho**2 * rho, (rho, 0, R), (phi, 0, 2*pi), (z, -h/2, h/2))
>>> I.subs(V, pi * R**2 * h)
R**2*m/2
```

And about the x axis:

$$\begin{aligned}
\mathbf{n} &= (1, 0, 0) \\
\mathbf{r} &= (\rho\cos\phi, \rho\sin\phi, z) \\
\mathbf{r}\cdot\mathbf{n} &= \rho\cos\phi \\
r^2 &= \rho^2 + z^2 \\
I = \int \rho(\mathbf{r})(r^2 - (\mathbf{r}\cdot\mathbf{n})^2)\mathrm{d}^3 r &= \int \frac{m}{V}(\rho^2 + z^2 - \rho^2\cos^2\phi)\mathrm{d}^3 r = \\
&= \frac{m}{V}\int_0^{2\pi}\mathrm{d}\phi\int_0^R \mathrm{d}R\int_{-\frac{h}{2}}^{\frac{h}{2}}\mathrm{d}z(\rho^2 + z^2 - \rho^2\cos^2\phi)\rho = \\
&= \frac{m}{V}\left(\frac{\pi R^4 h}{2} + \frac{\pi R^2 h^3}{12} - \frac{\pi R^4 h}{4}\right) = \\
&= \frac{m}{\pi R^2 h}\left(\frac{\pi R^4 h}{2} + \frac{\pi R^2 h^3}{12} - \frac{\pi R^4 h}{4}\right) = \\
&= \frac{m}{12}(6R^2 + h^2 - 3R^2) = \\
&= \frac{m}{12}(3R^2 + h^2)
\end{aligned}$$

Code:

```
>>> from sympy import var, integrate, pi, cos
>>> var("m V R rho z phi h")
(m, V, R, rho, z, phi, h)
>>> I = m/V * integrate((rho**2+z**2-rho**2*cos(phi)**2) * rho, (rho, 0, R), (phi, 0,␣
↪2*pi), (z, -h/2, h/2))
>>> I.subs(V, pi * R**2 * h).simplify()
m*(3*R**2 + h**2)/12
```

Special cases are a rod of length h (set $R = 0$ above) and a thin solid disk of radius R and mass m (set $h = 0$ above).

Sphere

Solid sphere of radius R and mass m. We'll use spherical coordinates. All axes are equivalent, so we use rotation about the z axis:

$$\begin{aligned}
V &= \frac{4}{3}\pi R^3 \\
\mathbf{n} &= (0,0,1) \\
\mathbf{r} &= (\rho\cos\phi\sin\theta, \rho\sin\phi\sin\theta, \rho\cos\theta) \\
\mathbf{r}\cdot\mathbf{n} &= \rho\cos\theta \\
r^2 &= \rho^2 \\
I = \int \rho(\mathbf{r})(r^2 - (\mathbf{r}\cdot\mathbf{n})^2)\mathrm{d}^3 r &= \int \frac{m}{V}(\rho^2 - \rho^2\cos^2\theta)\mathrm{d}^3 r = \\
&= \frac{m}{V}\int_0^{2\pi}\mathrm{d}\phi\int_0^R \mathrm{d}R\int_0^\pi \mathrm{d}\theta\rho^2(1-\cos^2\theta)\rho^2\sin\theta = \\
&= \frac{m}{V}\int_0^{2\pi}\mathrm{d}\phi\int_0^R \mathrm{d}R\int_0^\pi \mathrm{d}\theta\rho^4\sin^3\theta = \\
&= \frac{m}{V}2\pi\frac{R^5}{5}\frac{4}{3} = \\
&= \frac{m}{V}\frac{8}{15}\pi R^5 = \frac{m}{\frac{4}{3}\pi R^3}\frac{8}{15}\pi R^5 = \frac{2}{5}mR^2
\end{aligned}$$

Code:

```
>>> from sympy import var, integrate, pi, sin
>>> var("m V R rho theta phi")
(m, V, R, rho, theta, phi)
>>> I = m/V * integrate(rho**4 * sin(theta)**3, (rho, 0, R), (phi, 0, 2*pi), (theta,␣
↪0, pi))
>>> I
8*pi*R**5*m/(15*V)
>>> I.subs(V, 4*pi*R**3/3)
2*R**2*m/5
```

4.3 Relativity

4.3.1 Introduction: Why Tensors

This section gives a brief introduction, and in the next sections we derive everything in detail. The Newton law is:

$$m\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}t^2} = \mathbf{F}$$

and using a potential for $\mathbf{F}$, we get:

$$\begin{aligned}
\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}t^2} &= -\nabla\phi \\
\frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} &= -\partial^i\phi
\end{aligned}$$

the last two equations are two different equivalent ways to write a tensor equation in 3D, which means that this equation has the exact same form (is valid) in any (spatial) coordinate system (rotated, translated, in cartesian coordinates,

spherical coordinates, ...). Each coordinate system has a different metric, but we can always locally transform into $g_{ij} = \mathrm{diag}(1, 1, 1)$.

However, if our coordinate transformation depends on time (e.g. a rotating disk), then the above tensor equation changes (e.g. for the rotating disk, we get the Coriolis acceleration term), that's because time is treated as a parameter, not as a coordinate.

To fix this, we need to work in 4D and treat time as a coordinate, so we introduce $x^0 = ct$ where c is any constant speed (it can be any speed, doesn't have to be the speed of light). Then in 4D, the above equations are not tensor equations anymore, because the operator $\frac{\mathrm{d}}{\mathrm{d}t} = c\partial_0$ is not a tensor. The 4D tensor formulation happens to be the geodesic equation:

$$\begin{aligned} \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda}\nabla_\beta\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda} &= 0 \\ R_{00} &= 4\pi G\rho \\ R_{ij} &= 0 \end{aligned}$$

Which (given that we know how to calculate the Ricci tensor in our coordinates) is valid in any coordinates, not only rotated, translated, cartesian, spherical, ..., but also with arbitrary time dependence, e.g. a rotating disk, accelerating disk, ...

After suitable local coordinate transformation, we can only get two possible metrics (that connect the time and spatial coordinates): $\mathrm{diag}(-1, 1, 1, 1)$ and $\mathrm{diag}(1, 1, 1, 1)$. Inertial systems have no fictitious forces, so the metrics is one of the two above (possibly with $c \to \infty$). Transformation between inertial systems is such a coordinate transformation that leaves the metric intact, e.g.:

$$g' = \Lambda^T g \Lambda$$

There is no coordinate transformation that turns the metric $\mathrm{diag}(-1, 1, 1, 1)$ into $\mathrm{diag}(1, 1, 1, 1)$, so we need to choose either one to describe one inertial system and then all other inertial systems will automatically have a metric with the same signature.

The Newton law is valid for small speeds compared to the speed of light, so when we want to extend the theory for all speeds, we only have 4 options: O(3, 1) with either $c \to \infty$ or c finite and O(4) with either $c \to \infty$ or c finite. If c is finite, it has to be large enough, so that we still recover the Newton law for small speeds with the given experimental precision. All 4 cases give the correct Newton law, but give different predictions for large speeds. All we need to do to decide which one is correct is to perform such large speeds (relativistic) experiments. It turns out that all such relativistic experiments are in agreement with the O(3, 1) case where c is the (finite) speed of light and with disagreement with the 3 other cases. For small speeds however (i.e. Newtonean physics), all 4 cases will work, as long as c is chosen large enough.

Given a tensor equation, we can easily determine, if it transforms correctly under the Galilean ($c \to \infty$) or Lorentz transformations (c is finite). All we have to do is to perform the limit $c \to \infty$. For example the Newton second law is recovered if we do the $c \to \infty$ limit, but Maxwell equations are only recovered if we choose c to be exactly the speed of light in the Maxwell equations.

The reason why we write equations as tensor equations in 4D is that we can then use any coordinates (including any time dependence), i.e. any observer, and the equations still have the exact same form. So specifying the metrics is enough to define the coordinates (observer) and since the equations has only one form, that is all we need. If we write equations only as tensors in 3D, we not only need to specify the (3D) metrics, but also how the observer accelerates with respect to some (usually inertial) frame where the equations (let's say Newton law) is defined and we then need to transform all the time derivatives correctly. By using tensors in 4D, all those transformations are taken care of by the standard tensor machinery and all we need to care about is exactly one observer, defined by its metric tensor.

By choosing the correct metrics and c (i.e. $\mathrm{diag}(-1, 1, 1, 1)$ and c the speed of light), all equations are then automatically Lorentz invariant. If we choose $c \to \infty$ (and any metric), we automatically get all equations Galilean invariant.

4.3.2 High School Formulation

The usual (high school) formulation is the second Newton's law:

$$m\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}t^2} = \mathbf{F}$$

for some particle of the mass m and position $\mathbf{x}$. To determine the force $\mathbf{F}$, we have at hand the Newton's law of gravitation:

$$|\mathbf{F}| = G\frac{m_1 m_2}{r^2}$$

that expresses the magnitude $|\mathbf{F}|$ of the force between two particles with masses m_1 and m_2 and we also know that the direction of the force is directly towards the other particle. We need to take into account all particles in the system, determine the direction and magnitude of the force due to each of them and sum it up.

4.3.3 College Formulation

Unfortunately, it is quite messy to keep track of the direction of the forces and all the masses involved, it quickly becomes cumbersome for more than 2 particles. For this reason, the better approach is to calculate the force (field) from the mass density function ρ:

$$\nabla \cdot \mathbf{F} = -4\pi G m\, \rho(t, x, y, z)$$

To see that both formulations are equivalent, integrate both sides inside some sphere:

$$\int \nabla \cdot \mathbf{F}\, \mathrm{d}x\mathrm{d}y\mathrm{d}z = -4\pi G m_2 \int \rho\, \mathrm{d}x\mathrm{d}y\mathrm{d}z$$

apply the Gauss theorem to the left hand side:

$$\int \nabla \cdot \mathbf{F}\, \mathrm{d}x\mathrm{d}y\mathrm{d}z = \int \mathbf{F} \cdot \mathbf{n}\, \mathrm{d}S = 4\pi r^2\, \mathbf{F} \cdot \mathbf{n}$$

where $\mathbf{n} = \frac{\mathbf{r}}{|\mathbf{r}|}$ and the right hand side is equal to $-4\pi G m_1 m_2$ and we get:

$$\mathbf{F} \cdot \mathbf{n} = -G\frac{m_1 m_2}{r^2}$$

now we multiply both sides with $\mathbf{n}$, use the fact that $(\mathbf{F} \cdot \mathbf{n})\mathbf{n} = \mathbf{F}$ (because $\mathbf{F}$ is spherically symmetric), and we get the traditional Newton's law of gravitation:

$$\mathbf{F} = -G\frac{m_1 m_2}{r^2}\mathbf{n}$$

It is useful to deal with a scalar field instead of a vector field (and also not to have the mass m of the test particle in our equations explicitly), so we define a gravitational potential by:

$$\mathbf{F} = -m\nabla\phi(t, x, y, z)$$

then the law of gravitation is

$$\nabla^2\phi = 4\pi G\rho \tag{4.3.3.1}$$

and the second law is:

$$m\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}t^2} = -m\nabla\phi(t, x, y, z)$$

Note about units:

$$[r] = [\mathbf{x}] = \mathrm{m}$$

$$[m] = \mathrm{kg}$$

$$[\rho] = \mathrm{kg\,m^{-3}}$$

$$[F] = \mathrm{kg\,m\,s^{-2}}$$

$$[G] = \mathrm{kg^{-1}\,m^3\,s^{-2}}$$

$$[\phi] = \mathrm{m^2\,s^{-2}}$$

Example

Calculate the force acting on a test particle inside an infinitely thin spherical shell of radius R and surface mass distribution $\sigma(\theta, \phi) = 1$. We need to solve

$$\nabla^2 \phi = 4\pi G \rho \tag{4.3.3.2}$$

with

$$\rho(x, y, z) = \sigma(\theta, \phi) \frac{\delta(R - r)}{r^2}$$
$$r = \sqrt{x^2 + y^2 + z^2}$$

the Green function of (4.3.3.2) is

$$G(\mathbf{x}, \mathbf{y}) = \frac{1}{|\mathbf{x} - \mathbf{y}|}$$

so the solution is:

$$\phi = \int G(\mathbf{x}, \mathbf{y}) 4\pi G \rho(\mathbf{y}) \mathrm{d}^3 y = 4\pi G \int \frac{\rho(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|} \mathrm{d}^3 y =$$
$$= 4\pi G \int \frac{\sigma(\theta, \phi) \frac{\delta(R-r)}{r^2} r^2 \sin\theta}{\sqrt{(x - r\sin\theta\cos\phi)^2 + (y - r\sin\theta\sin\phi)^2 + (z - r\cos\theta)^2}} \mathrm{d}\theta \mathrm{d}\phi \mathrm{d}r =$$
$$= 4\pi G \int \frac{\delta(R - r)\sin\theta}{\sqrt{(x - r\sin\theta\cos\phi)^2 + (y - r\sin\theta\sin\phi)^2 + (z - r\cos\theta)^2}} \mathrm{d}\theta \mathrm{d}\phi \mathrm{d}r =$$
$$= 4\pi G \int \frac{\sin\theta}{\sqrt{(x - R\sin\theta\cos\phi)^2 + (y - R\sin\theta\sin\phi)^2 + (z - R\cos\theta)^2}} \mathrm{d}\theta \mathrm{d}\phi =$$
$$= 4\pi G \int \frac{\sin\theta}{\sqrt{x^2 + y^2 + z^2 + R^2 - 2R(x\sin\theta\cos\phi + y\sin\theta\sin\phi + z\cos\theta)}} \mathrm{d}\theta \mathrm{d}\phi$$

for symmetry reasons we can set $x = 0$, $y = 0$ (it can also be done more exactly, as shown in *Example II*):

$$\begin{aligned}
\phi(0,0,z) = 4\pi G \int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \mathrm{d}\theta \frac{\sin\theta}{\sqrt{z^2+R^2-2Rz\cos\theta}} = \\
= 8\pi^2 G \int_0^{\pi} \mathrm{d}\theta \frac{\sin\theta}{\sqrt{z^2+R^2-2Rz\cos\theta}} = \\
= 8\pi^2 G \int_{-1}^{1} \frac{\mathrm{d}y}{\sqrt{z^2+R^2-2Rzy}} = \\
= -\frac{4\pi^2 G}{Rz} \int_{(R-z)^2}^{(R+z)^2} \frac{\mathrm{d}u}{\sqrt{u}} = \\
= -\frac{4\pi^2 G}{Rz} \Big[2\sqrt{u}\Big]_{(R-z)^2}^{(R+z)^2} = \\
= -\frac{4\pi^2 G}{Rz} \Big[2|R+z| - 2|R-z|\Big] = \\
= -\frac{4\pi^2 G}{Rz} \Big[4z\Big] = \\
= -\frac{16\pi^2 G}{R}
\end{aligned}$$

This must hold for all x and y (less than R), so:

$$\phi(x,y,z) = -\frac{16\pi^2 G}{R}$$

And the force is

$$\mathbf{F} = -m\nabla\phi(t,x,y,z) = -m\nabla\left(-\frac{16\pi^2 G}{R}\right) = 0$$

So the force acting on a test particle inside the shell is zero.

4.3.4 Differential Geometry Formulation

There are still problems with this formulation, because it is not immediatelly clear how to write those laws in other frames, for example rotating, or accelerating – one needs to employ nontrivial assumptions about the systems, space, relativity principle and it is often a source confusion. Fortunately there is a way out — differential geometry. By reformulating the above laws in the language of the differential geometry, everything will suddenly be very explicit and clear. As an added bonus, because the special and general relativity uses the same language, the real differences between all these three theories will become clear.

We write x, y, z and t as components of one 4-vector

$$x^\mu = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

In this section, you can imagine $c = 1$, but we'll need it later, so we put it in right now, so that we don't need to rederive all equations again. Now we need to connect the Newtonian equations to geometry. To do that, we reformulate the Newton's second law:

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} + \delta^{ij}\partial_j\phi = 0$$

by choosing a parameter λ such, that $\frac{\mathrm{d}^2\lambda}{\mathrm{d}t^2} = 0$, so in general

$$\lambda = at + b$$

and

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} = a^2 \frac{\mathrm{d}^2}{\mathrm{d}\lambda^2}$$

so

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}\lambda^2} + \frac{1}{a^2}\delta^{ij}\partial_j\phi = 0$$

and using the relation $\frac{\mathrm{d}\lambda}{\mathrm{d}a} = a$ we get

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}\lambda^2} + \delta^{ij}\partial_j\phi \left(\frac{\mathrm{d}t}{\mathrm{d}\lambda}\right)^2 = 0$$

So using x^0 instead of t, we endup with the following equations:

$$\begin{aligned} \frac{\mathrm{d}^2 x^0}{\mathrm{d}\lambda^2} &= 0 \\ \frac{\mathrm{d}^2 x^i}{\mathrm{d}\lambda^2} + \frac{1}{c^2}\delta^{ij}\partial_j\phi \left(\frac{\mathrm{d}x^0}{\mathrm{d}\lambda}\right)^2 &= 0 \end{aligned}$$

But this is exactly the geodesic equation for the following Christoffel symbols:

$$\Gamma^i_{00} = \frac{1}{c^2}\delta^{ij}\partial_j\phi \tag{4.3.4.1}$$

and all other components are zero.

In order to formulate the gravitation law, we now need to express $\nabla^2\phi$ in terms of geometric quantities like $\Gamma^\alpha_{\beta\gamma}$ or $R^\alpha{}_{\beta\gamma\delta}$. We get the only nonzero components of the Riemann tensor:

$$R^j{}_{0k0} = -R^j{}_{00k} = \frac{1}{c^2}\delta^{ji}\partial_i\partial_k\phi$$

we calculate the $R_{\alpha\beta}$ by contracting:

$$R_{00} = R^\mu{}_{0\mu0} = R^i{}_{0i0} = \frac{1}{c^2}\delta^{ij}\partial_i\partial_j\phi$$

$$R_{ij} = 0$$

comparing with (4.3.3.1) we see that the Newton gravitation law is

$$\begin{aligned} R_{00} &= \frac{4\pi G}{c^2}\rho \\ R_{ij} &= 0 \end{aligned}$$

Thus we have reformulated the Newton's laws in a frame invariant way — the matter curves the geometry using the equations:

$$\begin{aligned} R_{00} &= \frac{4\pi G}{c^2}\rho \\ R_{ij} &= 0 \end{aligned}$$

from which one can (for example) calculate the Christoffel symbols and other things. The particles then move on the geodesics:

$$\frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}\lambda^2} + \Gamma^\alpha_{\beta\gamma} \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\gamma}{\mathrm{d}\lambda} = 0$$

Both equations now have the same form in all coordinate systems (inertial or not) and it is clear how to transform them — only the Christoffel symbols (and Ricci tensor) change and we have a formula for their transformation.

Obviously this works for any value of c (as it cancels out in the final equations of motion) and at this level we don't really need it yet, so we can set $c = 1$ and forget about it. In the next section we will need some constant in the metric to send to infinity in order to obtain the correct Christoffel symbols, and we can conveniently just use c. Later on we introduce special relativity and we need to introduce a speed of light and it turns out that we can again just use c for that without any loss of generality.

4.3.5 Metrics

There is a slight problem with the metrics — it can be proven that there is no metrics, that generates the Christoffel symbols above. However, it turns out that if we introduce an invariant speed c in the metrics, then calculate the Christoffel symbols (thus they depend on c) and then do the limit $c \to \infty$, we can get the Christoffel symbols above.

In fact, it turns out that there are many such metrics that generate the right Christoffel symbols. Below we list several similar metrics and the corresponding Christoffel symbols (in the limit $c \to \infty$), so that we can get a better feeling what metrics work and what don't and why:

$$g_{\mu\nu} = \begin{pmatrix} -c^2 - 2\phi & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\Gamma^1_{00} = \partial_x \phi$$

$$\Gamma^2_{00} = -\partial_y \phi$$

$$\Gamma^3_{00} = \partial_z \phi$$

$$g_{\mu\nu} = \begin{pmatrix} -c^2 - 2\phi & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$\Gamma^1_{00} = \partial_x \phi$$

$$\Gamma^2_{00} = -\partial_y \phi$$

$$\Gamma^3_{00} = -\partial_z \phi$$

$$g_{\mu\nu} = \begin{pmatrix} -c^2 - 2\phi & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$\Gamma^1_{00} = -\partial_x \phi$$

$$\Gamma^2_{00} = -\partial_y \phi$$

$$\Gamma^3_{00} = -\partial_z \phi$$

$$g_{\mu\nu} = \begin{pmatrix} -c^2 + 45 - 2\phi & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\Gamma^1_{00} = \partial_x \phi$$

$$\Gamma^2_{00} = \partial_y \phi$$

$$\Gamma^3_{00} = \partial_z \phi$$

$$g_{\mu\nu} = \begin{pmatrix} -c^2 - 2\phi & 0 & 0 & 0 \\ 0 & 1 - \frac{2\phi}{c^2} & 0 & 0 \\ 0 & 0 & 1 - \frac{2\phi}{c^2} & 0 \\ 0 & 0 & 0 & 1 - \frac{2\phi}{c^2} \end{pmatrix}$$

$$\Gamma^1_{00} = \partial_x \phi$$

$$\Gamma^2_{00} = \partial_y \phi$$

$$\Gamma^3_{00} = \partial_z \phi$$

$$g_{\mu\nu} = \begin{pmatrix} -c^2 - 2\phi & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\Gamma^1_{00} = \partial_x \phi$$

$$\Gamma^2_{00} = \partial_y \phi$$

$$\Gamma^3_{00} = \partial_z \phi$$

$$g_{\mu\nu} = \begin{pmatrix} c^2 - 2\phi & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\Gamma^1_{00} = \partial_x \phi$$

$$\Gamma^2_{00} = \partial_y \phi$$

$$\Gamma^3_{00} = \partial_z \phi$$

$$g_{\mu\nu} = \begin{pmatrix} c^2 - 2\phi & 0 & 0 & 0 \\ 0 & c^2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\Gamma^2_{00} = \partial_y \phi$$

$$\Gamma^3_{00} = \partial_z \phi$$

$$g_{\mu\nu} = \begin{pmatrix} c^2 - 2\phi & 0 & 0 & 0 \\ 0 & 1 & 0 & \frac{2\phi}{c^2} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\Gamma^1_{00} = \partial_x \phi$$

$$\Gamma^2_{00} = \partial_y \phi$$

$$\Gamma^3_{00} = \partial_z \phi$$

$$g_{\mu\nu} = \begin{pmatrix} c^2 - 2\phi & 0 & 0 & 0 \\ 0 & 1 & 0 & c^2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\Gamma^1_{00} = -\infty$$

$$\Gamma^2_{00} = \partial_y \phi$$

$$\Gamma^3_{00} = \partial_z \phi$$

$$g_{\mu\nu} = \begin{pmatrix} c^2 - 2\phi & 0 & 0 & 0 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\Gamma^1_{00} = \partial_x \phi - 5\partial_z \phi$$

$$\Gamma^2_{00} = \partial_y \phi$$

$$\Gamma^3_{00} = \partial_z \phi$$

$$g_{\mu\nu} = \begin{pmatrix} c^2 - 2\phi & 0 & 5 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\Gamma^1_{00} = \partial_x \phi$$

$$\Gamma^2_{00} = \partial_y \phi$$

$$\Gamma^3_{00} = \partial_z \phi$$

If we do the limit $c \to \infty$ in the metrics itself, all the working metrics degenerate to:

$$g_{\mu\nu} = \begin{pmatrix} \pm\infty & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(possibly with nonzero but finite elements $g_{0i} = g_{i0} \neq 0$). So it seems like any metrics whose limit is $\mathrm{diag}(\pm\infty, 1, 1, 1)$, generates the correct Christoffel symbols:

$$\Gamma^1_{00} = \partial_x \phi$$

$$\Gamma^2_{00} = \partial_y \phi$$

$$\Gamma^3_{00} = \partial_z \phi$$

but this would have to be investigated further.

Let's take the metrics $\mathrm{diag}(-c^2-2\phi, 1-\frac{2\phi}{c^2}, 1-\frac{2\phi}{c^2}, 1-\frac{2\phi}{c^2})$ and calculate the Christoffel symbols (without the limit $c\to\infty$):

$$\Gamma^0_{\mu\nu} = \begin{pmatrix} -\frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{-2\phi(t,x,y,z)-c^2} & -\frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{-2\phi(t,x,y,z)-c^2} & -\frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{-2\phi(t,x,y,z)-c^2} & -\frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{-2\phi(t,x,y,z)-c^2} \\ -\frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{-2\phi(t,x,y,z)-c^2} & \frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{c^2(-2\phi(t,x,y,z)-c^2)} & 0 & 0 \\ -\frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{-2\phi(t,x,y,z)-c^2} & 0 & \frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{c^2(-2\phi(t,x,y,z)-c^2)} & 0 \\ -\frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{-2\phi(t,x,y,z)-c^2} & 0 & 0 & \frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{c^2(-2\phi(t,x,y,z)-c^2)} \end{pmatrix}$$

$$\Gamma^1_{\mu\nu} = \begin{pmatrix} \frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{1-2\frac{\phi(t,x,y,z)}{c^2}} & -\frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} & 0 & 0 \\ -\frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} & -\frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} & -\frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} & -\frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} \\ 0 & -\frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} & \frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} & 0 \\ 0 & -\frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} & 0 & \frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} \end{pmatrix}$$

$$\Gamma^2_{\mu\nu} = \begin{pmatrix} \frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{1-2\frac{\phi(t,x,y,z)}{c^2}} & 0 & -\frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} & 0 \\ 0 & \frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} & -\frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} & 0 \\ -\frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} & -\frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} & -\frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} & -\frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} \\ 0 & 0 & -\frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} & \frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} \end{pmatrix}$$

$$\Gamma^3_{\mu\nu} = \begin{pmatrix} \frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{1-2\frac{\phi(t,x,y,z)}{c^2}} & 0 & 0 & -\frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} \\ 0 & \frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} & 0 & -\frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} \\ 0 & 0 & \frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} & -\frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} \\ -\frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} & -\frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} & -\frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} & -\frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{c^2\left(1-2\frac{\phi(t,x,y,z)}{c^2}\right)} \end{pmatrix}$$

By taking the limit $c\to\infty$, the only nonzero Christoffel symbols are:

$$\Gamma^1_{00} = \partial_x\phi$$

$$\Gamma^2_{00} = \partial_y\phi$$

$$\Gamma^3_{00} = \partial_z\phi$$

or written compactly:

$$\Gamma^i_{00} = \delta^{ij}\partial_j\phi$$

So the geodesics equation

$$\frac{\mathrm{d}^2x^\alpha}{\mathrm{d}\lambda^2} + \Gamma^\alpha_{\beta\gamma}\frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\gamma}{\mathrm{d}\lambda} = 0$$

becomes

$$\frac{\mathrm{d}^2x^0}{\mathrm{d}\lambda^2} = 0$$

$$\frac{\mathrm{d}^2x^i}{\mathrm{d}\lambda^2} + \delta^{ij}\partial_j\phi\left(\frac{\mathrm{d}x^0}{\mathrm{d}\lambda}\right)^2 = 0$$

From the first equation we get $x^0 = a\lambda + b$, we substitute to the second equation:

$$\frac{1}{a^2}\frac{\mathrm{d}^2 x^i}{\mathrm{d}\lambda^2} + \delta^{ij}\partial_j\phi = 0$$

or

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}(x^0)^2} + \delta^{ij}\partial_j\phi = 0$$

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} = -\delta^{ij}\partial_j\phi$$

So the Newton's second law *is* the equation of geodesics.

In the above, we have set $c = 1$ in the Christoffel symbols themselves (see the last paragraph from the last section) and introduced another constant c in the metric itself. As we can see, the metric will become infinite with this approach in the limit $c \to \infty$. Another approach is to store this c in the x^μ vector itself, then the metric stays finite (in fact becomes a diagonal matrix $\mathrm{diag}(\pm 1, 1, 1, 1)$, thus it gives all the Christoffel symbols equal to zero, in the limit), but the vector becomes infinite in the limit.

Either way our formalism breaks down, and thus we need to keep c finite and only do the limit in the final equations (after we don't need differential geometry anymore). When needed, we can also carefully neglect higher terms in c, that will not appear in the final equations after doing the limit, but one needs to make sure that no mistake is made.

It is customary to put the constant c into the vector x^μ and so we will do so too from this point on.

4.3.6 Conclusion About Metric

We will use the convention to keep c in the 4-vector and the simplest metric that generates the correct Christoffel symbols is the following:

$$g_{\mu\nu} = \begin{pmatrix} \pm 1 - \frac{2\phi}{c^2} & 0 & 0 & 0 \\ 0 & 1 - \frac{2\phi}{c^2} & 0 & 0 \\ 0 & 0 & 1 - \frac{2\phi}{c^2} & 0 \\ 0 & 0 & 0 & 1 - \frac{2\phi}{c^2} \end{pmatrix}$$

In the limit $c \to \infty$ we get the following nonzero Christoffel symbols (for both signs in ± 1 above):

$$\Gamma^i_{00} = \frac{1}{c^2}\delta^{ij}\partial_j\phi$$

all other symbols contain higher powers of c and thus will not contribute in the limit $c \to \infty$. The remaining c^2 in Γ^i_{00} will cancel with the c in $x^0 = ct$ in the final equations.

As seen above, there is some freedom in which metric we can use in order to obtain the correct Christoffel symbols, but the above metric is the simplest, so we'll use it from now on.

4.3.7 Einstein's Equations

Einstein's equations are derived from the Hilbert action:

$$S_H = \frac{c^4}{16\pi G}\int R\sqrt{|\det g_{\mu\nu}|}\mathrm{d}^4 x = \frac{c^4}{16\pi G}\int g^{\mu\nu}R_{\mu\nu}\sqrt{|\det g|}\mathrm{d}^4 x$$

The Lagrangian density $R\sqrt{|\det g_{\mu\nu}|}$ has to be given, that's our assumption and everything else is derived from it. In principle it can have other terms, for example $\alpha_1 R^2 + \alpha_2 R_{\mu\nu}R^{\mu\nu} + \alpha_3 g^{\mu\nu}\nabla_\mu R\nabla_\nu R + \cdots$ and there are a lot of possibilities and ultimately the exact form of the Lagrangian has to be decided by experiment. The Hilbert action is

the simplest possible action and it already gives a theory which agrees with experiment, so that will be our starting point.

Varying it with respect to the metric $g^{\mu\nu}$ we get:

$$\begin{aligned}
\delta S_H = \delta\frac{c^4}{16\pi G}\int R\sqrt{|\det g|}\mathrm{d}^4x = \\
= \frac{c^4}{16\pi G}\int (\delta g^{\mu\nu})R_{\mu\nu}\sqrt{|\det g|} + g^{\mu\nu}(\delta R_{\mu\nu})\sqrt{|\det g|} + R(\delta\sqrt{|\det g|})\mathrm{d}^4x = \\
= \frac{c^4}{16\pi G}\int (\delta g^{\mu\nu})R_{\mu\nu}\sqrt{|\det g|} + g^{\mu\nu}\left(\nabla_\lambda(\delta\Gamma^\lambda_{\nu\mu}) - \nabla_\nu(\delta\Gamma^\lambda_{\lambda\mu})\right)\sqrt{|\det g|} + R(-\tfrac{1}{2}\sqrt{|\det g|}\,g_{\mu\nu}(\delta g^{\mu\nu}))\mathrm{d}^4x = \\
= \frac{c^4}{16\pi G}\int (\delta g^{\mu\nu})R_{\mu\nu}\sqrt{|\det g|} + \left(\nabla_\lambda g^{\mu\nu}(\delta\Gamma^\lambda_{\nu\mu}) - \nabla_\nu g^{\mu\nu}(\delta\Gamma^\lambda_{\lambda\mu})\right)\sqrt{|\det g|} - \tfrac{1}{2}Rg_{\mu\nu}\sqrt{|\det g|}\,(\delta g^{\mu\nu})\mathrm{d}^4x = \\
= \frac{c^4}{16\pi G}\int (\delta g^{\mu\nu})R_{\mu\nu}\sqrt{|\det g|} - \tfrac{1}{2}Rg_{\mu\nu}\sqrt{|\det g|}\,(\delta g^{\mu\nu})\mathrm{d}^4x = \\
= \frac{c^4}{16\pi G}\int \left(R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu}\right)(\delta g^{\mu\nu})\sqrt{|\det g|}\mathrm{d}^4x
\end{aligned}$$

Where we used the following identities:

$$\begin{aligned}
\delta\sqrt{|\det g|} &= -\tfrac{1}{2}\sqrt{|\det g|}\,g_{\mu\nu}(\delta g^{\mu\nu}) \\
\delta R^\rho{}_{\mu\lambda\nu} &= \nabla_\lambda(\delta\Gamma^\rho_{\nu\mu}) - \nabla_\nu(\delta\Gamma^\rho_{\lambda\mu}) \\
\delta R_{\mu\nu} = \delta R^\lambda{}_{\mu\lambda\nu} &= \nabla_\lambda(\delta\Gamma^\lambda_{\nu\mu}) - \nabla_\nu(\delta\Gamma^\lambda_{\lambda\mu})
\end{aligned}$$

and the fact that the four divergence doesn't contribute to the integral. By setting $\delta S_H = 0$, we get:

$$\frac{2}{\sqrt{|\det g|}}\frac{\delta S_H}{\delta g^{\mu\nu}} = \frac{c^4}{8\pi G}(R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu}) = 0$$

Combining the Hilbert action S_H with the action for matter S_M we get:

$$S = S_H + S_M$$

Varying this action as above we get:

$$\frac{2}{\sqrt{|\det g|}}\frac{\delta S}{\delta g^{\mu\nu}} = \frac{c^4}{8\pi G}\left(R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu}\right) + \frac{2}{\sqrt{|\det g|}}\frac{\delta S_M}{\delta g^{\mu\nu}} = 0$$

so:

$$R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu} = -\frac{8\pi G}{c^4}\frac{2}{\sqrt{|\det g|}}\frac{\delta S_M}{\delta g^{\mu\nu}} = \frac{8\pi G}{c^4}T_{\mu\nu}$$

Where we set:

$$T_{\mu\nu} = -\frac{2}{\sqrt{|\det g|}}\frac{\delta S_M}{\delta g^{\mu\nu}}$$

This is a definition of the stress energy tensor corresponding to the action $S_M = \int \mathcal{L}_M\sqrt{|\det g|}\mathrm{d}^4x$. We can also

write it in terms of the Lagrangian $\mathcal{L}_M$ directly as:

$$\begin{aligned}
T_{\mu\nu} &= -\frac{2}{\sqrt{|\det g|}}\frac{\delta S_M}{\delta g^{\mu\nu}} = \\
&= -\frac{2}{\sqrt{|\det g|}}\frac{\delta \int \mathcal{L}_M \sqrt{|\det g|}\mathrm{d}^4x}{\delta g^{\mu\nu}} = \\
&= -\frac{2}{\sqrt{|\det g|}}\frac{\int (\delta\mathcal{L}_M)\sqrt{|\det g|} + \mathcal{L}_M\left(\delta\sqrt{|\det g|}\right)\mathrm{d}^4x}{\delta g^{\mu\nu}} = \\
&= -\frac{2}{\sqrt{|\det g|}}\frac{\int \left(\frac{\delta\mathcal{L}_M}{\delta g^{\mu\nu}}(\delta g^{\mu\nu})\right)\sqrt{|\det g|} + \mathcal{L}_M\left(-\frac{1}{2}\sqrt{|\det g|}\, g_{\mu\nu}(\delta g^{\mu\nu})\right)\mathrm{d}^4x}{\delta g^{\mu\nu}} = \\
&= -\frac{2}{\sqrt{|\det g|}}\left(\frac{\delta\mathcal{L}_M}{\delta g^{\mu\nu}}\sqrt{|\det g|} - \frac{1}{2}\mathcal{L}_M\sqrt{|\det g|}\, g_{\mu\nu}\right) = \\
&= -2\frac{\delta\mathcal{L}_M}{\delta g^{\mu\nu}} + g_{\mu\nu}\mathcal{L}_M
\end{aligned}$$

If this action contains electromagnetic field, we get an electromagnetic stress energy tensor. For continous matter, we get the stress energy tensor for continous matter, see the next section. The right hand side of the Einstein's equations contains the sum of all stress energy tensors (for all fields in the Lagrangian).

4.3.8 Continuous Distribution of Matter

The action is:

$$S_M = -\int \rho c\sqrt{v_\mu v^\mu}\sqrt{|\det g|}\mathrm{d}^4x$$

But it isn't suitable for applying variations because ρ and v^μ are not independent quantities. So we write it in terms of a 4-momentum vector density $\mathfrak{p}^\mu$:

$$\begin{aligned}
p^\mu &= \rho v^\mu \\
\mathfrak{p}^\mu &= p^\mu\sqrt{|\det g|} = \rho v^\mu\sqrt{|\det g|} \\
\sqrt{\mathfrak{p}_\mu\mathfrak{p}^\mu} &= \sqrt{\rho v_\mu\sqrt{|\det g|}\rho v^\mu\sqrt{|\det g|}} = \rho\sqrt{v_\mu v^\mu}\sqrt{|\det g|}
\end{aligned}$$

and the action becomes:

$$S_M = -\int \rho c\sqrt{v_\mu v^\mu}\sqrt{|\det g|}\mathrm{d}^4x = -\int c\sqrt{\mathfrak{p}_\mu\mathfrak{p}^\mu}\mathrm{d}^4x$$

We vary S_M with respect to $g^{\mu\nu}$:

$$\begin{aligned}
\delta S_M &= -\delta\int c\sqrt{\mathfrak{p}_\mu\mathfrak{p}^\mu}\mathrm{d}^4x = \\
&= -\int c\frac{\delta(g^{\mu\nu}\mathfrak{p}_\mu\mathfrak{p}_\nu)}{2\sqrt{\mathfrak{p}_\alpha\mathfrak{p}^\alpha}}\mathrm{d}^4x = \\
&= -\int c\frac{\mathfrak{p}_\mu\mathfrak{p}_\nu}{2\sqrt{\mathfrak{p}_\alpha\mathfrak{p}^\alpha}}\delta(g^{\mu\nu})\mathrm{d}^4x = \\
&= -\int c\frac{\rho v_\mu\rho v_\nu\sqrt{|\det g|}^2}{2\rho c\sqrt{|\det g|}}\delta(g^{\mu\nu})\mathrm{d}^4x = \\
&= -\int \frac{1}{2}\rho v_\mu v_\nu\delta(g^{\mu\nu})\sqrt{|\det g|}\mathrm{d}^4x
\end{aligned}$$

And the stress energy tensor is:

$$\begin{aligned} T_{\mu\nu} &= -\frac{2}{\sqrt{|\det g|}}\frac{\delta S_M}{\delta g^{\mu\nu}} = \\ &= -\frac{2}{\sqrt{|\det g|}}\left(-\tfrac{1}{2}\rho v_\mu v_\nu\sqrt{|\det g|}\right) = \\ &= \rho v_\mu v_\nu \end{aligned}$$

Now we vary S_M with respect to x^μ:

$$\begin{aligned} \delta S_M &= -\delta\int c\sqrt{\mathfrak{p}_\mu\mathfrak{p}^\mu}\mathrm{d}^4x = \\ &= -\int c\frac{\delta(g^{\mu\nu}\mathfrak{p}_\mu\mathfrak{p}_\nu)}{2\sqrt{\mathfrak{p}_\alpha\mathfrak{p}^\alpha}}\mathrm{d}^4x = \\ &= -\int c\frac{2g^{\mu\nu}\mathfrak{p}_\mu(\delta\mathfrak{p}_\nu)}{2\sqrt{\mathfrak{p}_\alpha\mathfrak{p}^\alpha}}\mathrm{d}^4x = \\ &= -\int c\frac{\mathfrak{p}_\mu}{\sqrt{\mathfrak{p}_\alpha\mathfrak{p}^\alpha}}(\delta\mathfrak{p}^\mu)\mathrm{d}^4x = \\ &= -\int c\frac{\mathfrak{p}_\mu}{\sqrt{\mathfrak{p}_\alpha\mathfrak{p}^\alpha}}\partial_\nu\left(\mathfrak{p}^\nu(\delta x^\mu)-\mathfrak{p}^\mu(\delta x^\nu)\right)\mathrm{d}^4x = \\ &= \int c\partial_\nu\left(\frac{\mathfrak{p}_\mu}{\sqrt{\mathfrak{p}_\alpha\mathfrak{p}^\alpha}}\right)\left(\mathfrak{p}^\nu(\delta x^\mu)-\mathfrak{p}^\mu(\delta x^\nu)\right)\mathrm{d}^4x = \\ &= \int c\left(\partial_\nu\left(\frac{\mathfrak{p}_\mu}{\sqrt{\mathfrak{p}_\alpha\mathfrak{p}^\alpha}}\right)-\partial_\mu\left(\frac{\mathfrak{p}_\nu}{\sqrt{\mathfrak{p}_\alpha\mathfrak{p}^\alpha}}\right)\right)\mathfrak{p}^\nu(\delta x^\mu)\mathrm{d}^4x = \\ &= \int c\left(\nabla_\nu\left(\frac{\mathfrak{p}_\mu}{\sqrt{\mathfrak{p}_\alpha\mathfrak{p}^\alpha}}\right)-\nabla_\mu\left(\frac{\mathfrak{p}_\nu}{\sqrt{\mathfrak{p}_\alpha\mathfrak{p}^\alpha}}\right)\right)\mathfrak{p}^\nu(\delta x^\mu)\mathrm{d}^4x = \\ &= \int\left(\nabla_\nu v_\mu-\nabla_\mu v_\nu\right)\rho v^\nu(\delta x^\mu)\sqrt{|\det g|}\mathrm{d}^4x = \\ &= \int\rho(\nabla_\nu v_\mu)v^\nu(\delta x^\mu)\sqrt{|\det g|}\mathrm{d}^4x \end{aligned}$$

So the equation of motion is the geodesic:

$$\rho(\nabla_\nu v_\mu)v^\nu = 0$$

Charged matter has the interaction action:

$$S_q = -\int\rho_{EM}v^\mu A_\mu\sqrt{|\det g|}\mathrm{d}^4x = -\int j^\mu A_\mu\sqrt{|\det g|}\mathrm{d}^4x = -\int \mathfrak{j}^\mu A_\mu\mathrm{d}^4x$$

where we have introduced the 4-current j^μ and 4-current density $\mathfrak{j}^\mu$:

$$\begin{aligned} j^\mu &= \rho_{EM}v^\mu \\ \mathfrak{j}^\mu &= j^\mu\sqrt{|\det g|} = \rho_{EM}v^\mu\sqrt{|\det g|} \end{aligned}$$

We vary S_q with respect to x^μ:

$$\begin{aligned}
\delta S_q = -\delta \int \jmath^\mu A_\mu \mathrm{d}^4x = \\
= -\int (\delta\jmath^\mu) A_\mu \mathrm{d}^4x = \\
= -\int \partial_\nu \left(\jmath^\nu(\delta x^\mu) - \jmath^\mu(\delta x^\nu)\right) A_\mu \mathrm{d}^4x = \\
= \int \left(\jmath^\nu(\delta x^\mu) - \jmath^\mu(\delta x^\nu)\right) \partial_\nu A_\mu \mathrm{d}^4x = \\
= \int \jmath^\nu(\delta x^\mu)(\partial_\nu A_\mu - \partial_\mu A_\nu)\mathrm{d}^4x = \\
= \int \rho_{EM} v^\nu (\partial_\nu A_\mu - \partial_\mu A_\nu)(\delta x^\mu)\sqrt{|\det g|}\mathrm{d}^4x = \\
= -\int \rho_{EM} v^\nu F_{\mu\nu}(\delta x^\mu)\sqrt{|\det g|}\mathrm{d}^4x
\end{aligned}$$

So the combined action $S_M + S_q$ yields:

$$\rho(\nabla_\nu v_\mu)v^\nu - \rho_{EM} v^\nu F_{\mu\nu} = 0$$

Varying S_q with respect to A^μ yields the 4-current $j^\mu = \rho_{EM} v^\mu$ which ends up on the right hand side of the Maxwell's equations when varying the S_{EM} action.

4.3.9 Obsolete Section

This section is obsolete, ideas from it should be polished (sometimes corrected) and put to other sections.

The problem is, that in general, Christoffel symbols have 40 components and metrics only 10 and in our case, we cannot find such a metrics, that generates the Christoffel symbols above. In other words, the spacetime that describes the Newtonian theory is affine, but not a metric space. The metrics is singular, and we have one metrics $\mathrm{diag}(-1, 0, 0, 0)$ that describes the time coordinate and another metrics $\mathrm{diag}(0, 1, 1, 1)$ that describes the spatial coordinates. We know the affine connection coefficients $\Gamma^\alpha_{\beta\gamma}$, so that is enough to calculate geodesics and to differentiate vectors and do everything we need.

However, for me it is still not satisfactory, because I really want to have a metrics tensor, so that I can easily derive things in exactly the same way as in general relativity. To do that, we will have to work in the regime c is finite and only at the end do the limit $c \to \infty$.

We start with Einstein's equations:

$$R_{\alpha\beta} - \tfrac{1}{2}Rg_{\alpha\beta} = \frac{8\pi G}{c^4}T_{\alpha\beta}$$

or

$$R_{\alpha\beta} = \frac{8\pi G}{c^4}(T_{\alpha\beta} - \tfrac{1}{2}Tg_{\alpha\beta})$$

$$R^\alpha{}_\beta = \frac{8\pi G}{c^4}(T^\alpha{}_\beta - \tfrac{1}{2}T)$$

The energy-momentum tensor is

$$T^{\alpha\beta} = \rho U^\alpha U^\beta$$

in our approximation $U^i \sim 0$ and $U^0 \sim c$, so the only nonzero component is:

$$T^{00} = \rho c^2$$

$$T = \rho c^2$$

and

$$R^i{}_j = \frac{8\pi G}{c^4}(-\tfrac{1}{2}\rho c^2) = -\frac{4\pi G}{c^2}\rho$$

$$R^0{}_0 = \frac{8\pi G}{c^4}(\tfrac{1}{2}\rho c^2) = \frac{4\pi G}{c^2}\rho$$

We need to find such a metric tensor, that

$$R^0{}_0 = \frac{1}{c^2}\nabla^2\phi$$

then we get (4.3.3.1).

There are several ways to choose the metrics tensor. We start We can always find a coordinate transformation, that converts the metrics to a diagonal form with only 1, 0 and -1 on the diagonal. If we want nondegenerate metrics, we do not accept 0 (but as it turns out, the metrics for the Newtonian mechanics *is* degenerated). Also, it is equivalent if we add a minus to all diagonal elements, e.g. $\mathrm{diag}(1,1,1,1)$ and $\mathrm{diag}(-1,-1,-1,-1)$ are equivalent, so we are left with these options only: signature 4:

$$g_{\mu\nu} = \mathrm{diag}(1,1,1,1)$$

signature 2:

$$g_{\mu\nu} = \mathrm{diag}(-1,1,1,1)$$

$$g_{\mu\nu} = \mathrm{diag}(1,-1,1,1)$$

$$g_{\mu\nu} = \mathrm{diag}(1,1,-1,1)$$

$$g_{\mu\nu} = \mathrm{diag}(1,1,1,-1)$$

signature 0:

$$g_{\mu\nu} = \mathrm{diag}(-1,-1,1,1)$$

$$g_{\mu\nu} = \mathrm{diag}(-1,1,-1,1)$$

$$g_{\mu\nu} = \mathrm{diag}(-1,1,1,-1)$$

No other possibility exists (up to adding a minus to all elements). We can also quite easily find coordinate transformations that swap coordinates, i.e. we can always find a transformation so that we first have only -1 and then only 1 on the diagonal, so we are left with: signature 4:

$$g_{\mu\nu} = \mathrm{diag}(1,1,1,1)$$

signature 2:

$$g_{\mu\nu} = \mathrm{diag}(-1,1,1,1)$$

signature 0:

$$g_{\mu\nu} = \mathrm{diag}(-1,-1,1,1)$$

One possible physical interpretation of the signature 0 metrics is that we have 2 time coordinates and 2 spatial coordinates. In any case, this metrics doesn't describe our space (neither Newtonian nor general relativity), because we really need the spatial coordinates to have the metrics either $\mathrm{diag}(1,1,1)$ or $\mathrm{diag}(-1,-1,-1)$.

So we are left with either (this case will probably not work, but I want to have an explicit reason why it doesn't work):

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

or (this is the usual special relativity)

$$g_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

It turns out, that one option to turn on gravitation is to add the term $-\frac{2\phi}{c^2}\mathbb{1}$ to the metric tensor, in the first case:

$$g_{\mu\nu} = \begin{pmatrix} 1-\frac{2\phi}{c^2} & 0 & 0 & 0 \\ 0 & 1-\frac{2\phi}{c^2} & 0 & 0 \\ 0 & 0 & 1-\frac{2\phi}{c^2} & 0 \\ 0 & 0 & 0 & 1-\frac{2\phi}{c^2} \end{pmatrix}$$

and second case:

$$g_{\mu\nu} = \begin{pmatrix} -1-\frac{2\phi}{c^2} & 0 & 0 & 0 \\ 0 & 1-\frac{2\phi}{c^2} & 0 & 0 \\ 0 & 0 & 1-\frac{2\phi}{c^2} & 0 \\ 0 & 0 & 0 & 1-\frac{2\phi}{c^2} \end{pmatrix}$$

The second law is derived from the equation of geodesic:

$$\frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}\lambda^2} + \Gamma^\alpha_{\beta\gamma} \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\gamma}{\mathrm{d}\lambda} = 0$$

in an equivalent form

$$\frac{\mathrm{d}U^\alpha}{\mathrm{d}\tau} + \Gamma^\alpha_{\beta\gamma} U^\beta U^\gamma = 0$$

The only nonzero Christoffel symbols in the first case are (in the expressions for the Christoffel symbols below, we set

$c = 1$):

$$\Gamma^0_{\mu\nu} = \begin{pmatrix} -\frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & -\frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & -\frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & -\frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} \\ -\frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & \frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & 0 & 0 \\ -\frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & 0 & \frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & 0 \\ -\frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & 0 & 0 & \frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} \end{pmatrix}$$

$$\Gamma^1_{\mu\nu} = \begin{pmatrix} \frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & -\frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & 0 & 0 \\ -\frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & -\frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & -\frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & -\frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} \\ 0 & -\frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & \frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & 0 \\ 0 & -\frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & 0 & \frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} \end{pmatrix}$$

$$\Gamma^2_{\mu\nu} = \begin{pmatrix} \frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & 0 & -\frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & 0 \\ 0 & \frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & -\frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & 0 \\ -\frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & -\frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & -\frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & -\frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} \\ 0 & 0 & -\frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & \frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} \end{pmatrix}$$

$$\Gamma^3_{\mu\nu} = \begin{pmatrix} \frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & 0 & 0 & -\frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} \\ 0 & \frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & 0 & -\frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} \\ 0 & 0 & \frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & -\frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} \\ -\frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & -\frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & -\frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} & -\frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{1-2\phi(t,x,y,z)} \end{pmatrix}$$

and in the second case, only $\Gamma^0_{\mu\nu}$ is different:

$$\Gamma^0_{\mu\nu} = \begin{pmatrix} \frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{1+2\phi(t,x,y,z)} & \frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{1+2\phi(t,x,y,z)} & \frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{1+2\phi(t,x,y,z)} & \frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{1+2\phi(t,x,y,z)} \\ \frac{\frac{\partial}{\partial x}\phi(t,x,y,z)}{1+2\phi(t,x,y,z)} & -\frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{1+2\phi(t,x,y,z)} & 0 & 0 \\ \frac{\frac{\partial}{\partial y}\phi(t,x,y,z)}{1+2\phi(t,x,y,z)} & 0 & -\frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{1+2\phi(t,x,y,z)} & 0 \\ \frac{\frac{\partial}{\partial z}\phi(t,x,y,z)}{1+2\phi(t,x,y,z)} & 0 & 0 & -\frac{\frac{\partial}{\partial t}\phi(t,x,y,z)}{1+2\phi(t,x,y,z)} \end{pmatrix}$$

Now we assume that $\partial_\mu\phi \sim \phi \ll c^2$, so all $\Gamma^\alpha_{\beta\gamma}$ are of the same order. Also $|U^i| \ll |U^0|$ and $U^0 = c$, so the only nonnegligible term is

$$\frac{\mathrm{d}U^\alpha}{\mathrm{d}\tau} + \Gamma^\alpha_{00}(U^0)^2 = 0$$

Substituting for the Christoffel symbol we get

$$\frac{\mathrm{d}U^i}{\mathrm{d}\tau} = -\frac{\delta^{ij}\partial_j\frac{\phi}{c^2}}{1-\frac{2\phi}{c^2}}\,c^2 = -\delta^{ij}(\partial_j\phi)\,\left(1+O\left(\frac{\phi}{c^2}\right)\right) = -\delta^{ij}\partial_j\phi + O\left(\left(\frac{\phi}{c^2}\right)^2\right)$$

and multiplying both sides with m:

$$m\frac{\mathrm{d}U^i}{\mathrm{d}\tau} = -m\partial_j\phi\,\delta^{ij}$$

which is the second Newton's law. For the zeroth component we get (first case metric)

$$m\frac{\mathrm{d}U^0}{\mathrm{d}\tau} = m\frac{\mathrm{d}\phi}{\mathrm{d}\tau}$$

second case:

$$m\frac{\mathrm{d}U^0}{\mathrm{d}\tau} = -m\frac{\mathrm{d}\phi}{\mathrm{d}\tau}$$

Where $mU^0 = p^0$ is the energy of the particle (with respect to this frame only), this means the energy is conserved unless the gravitational field depends on time.

To summarize: the Christoffel symbols (4.3.4.1) that we get from the Newtonian theory contain c, which up to this point can be any speed, for example we can set $c = 1\,\mathrm{ms}^{-1}$. However, in order to have some metrics tensor that generates those Christoffel symbols, the only way to do that is by the metrics

$$\mathrm{diag}(-1,1,1,1) - \frac{2\phi}{c^2}\mathbb{1}$$

then calculating the Christoffel symbols. If we neglect the terms of the order $O\left(\left(\frac{\phi}{c^2}\right)^2\right)$ and higher, we get the Newtonian Christoffel symbols (4.3.4.1) that we want. It's clear that in order to neglect the terms, we must have $|\phi| \ll c^2$, so we must choose c large enough for this to work. To put it plainly, unless c is large, there is no metrics in our Newtonian spacetime. However for c large, everything is fine.

4.3.10 Inertial frames

What is an inertial frame? Inertial frame is such a frame that doesn't have any fictitious forces. What is a fictitious force? If we take covariant time derivative of any vector, then fictitious forces are all the terms with nonzero Christoffel symbols. In other words, nonzero Christoffel symbols mean that by (partially) differentiating with respect to time, we need to add additional terms in order to get a proper vector again – and those terms are called fictitious forces if we are differentiating the velocity vector.

Inertial frame is a frame without fictitious forces, i.e. with all Christoffel symbols zero in the whole frame. This is equivalent to all components of the Riemann tensor being zero:

$$R^{\alpha}{}_{\beta\gamma\delta} = 0$$

In general, if $R^{\alpha}{}_{\beta\gamma\delta} \neq 0$ in the whole universe, then no such frame exists, but one can always achieve that locally, because one can always find a coordinate transformation so that the Christoffel symbols are zero locally (e.g. at one point), but unless $R^{\alpha}{}_{\beta\gamma\delta} = 0$, the Christoffel symbols will *not* be zero in the whole frame. So the (local) inertial frame is such a frame that has zero Christoffel symbols (locally).

What is the metrics of the inertial frame? It is such a metrics, that $\Gamma^{\alpha}{}_{\beta\gamma} = 0$. The derivatives $\partial_\mu \Gamma^{\alpha}{}_{\beta\gamma}$ however doesn't have to be zero. We know that taking any of the metrics listed above with $\phi = const$ we get all the Christoffel symbols zero. So for example these two metrics (one with a plus sign, the other with a minus sign) have all the Christoffel symbols zero:

$$g_{\mu\nu} = \begin{pmatrix} \pm c^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Such a metrics corresponds to an inertial frame then.

What are the (coordinate) transformations, that transform from one inertial frame to another? Those are all transformations that start with an inertial frame metrics (an example of such a metrics is given above), transform it using the transformation matrix and the resulting metrics is also inertial. In particular, let x^μ be inertial, thus $g_{\mu\nu}$ is an inertial metrics, then transform to x'^μ and g':

$$g'_{\alpha\beta} = \frac{\partial x^\mu}{\partial x'^\alpha}\frac{\partial x^\nu}{\partial x'^\beta} g_{\mu\nu} = \left(\frac{\partial x}{\partial x'}\right)^T g \left(\frac{\partial x}{\partial x'}\right)$$

if we denote the transformation matrix by Λ:

$$\Lambda^{\mu}{}_{\alpha} = \frac{\partial x^{\mu}}{\partial x'^{\alpha}}$$

then the transformation law is:

$$g' = \Lambda^T g \Lambda$$

Now let's assume that $g' = g$, i.e. both inertial systems are given by the same matrix and let's assume this particular form:

$$g'_{\mu\nu} = g_{\mu\nu} = \begin{pmatrix} \pm c^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(e.g. this covers almost all possible Newtonian metrics tensors).

4.3.11 Lorentz Group

The Lorentz group is O(3,1), e.g. all matrices satisfying:

$$g = \Lambda^T g \Lambda \tag{4.3.11.1}$$

with $g = \mathrm{diag}(-c^2, 1, 1, 1)$. Taking the determinant of (4.3.11.1) we get $(\det \Lambda)^2 = 1$ or $\det \Lambda = \pm 1$. Writing the 00 component of (4.3.11.1) we get

$$-c^2 = -c^2 (A^0{}_0)^2 + (A^0{}_1)^2 + (A^0{}_2)^2 + (A^0{}_3)^2$$

or

$$(A^0{}_0)^2 = 1 + \frac{1}{c^2}\left((A^0{}_1)^2 + (A^0{}_2)^2 + (A^0{}_3)^2\right)$$

Thus we can see that either $A^0{}_0 \geq 1$ (the transformation preserves the direction of time, orthochronous) or $A^0{}_0 \leq -1$ (not orthochronous). Thus we can see that the O(3, 1) group consists of 4 continuous parts, that are not connected.

First case: elements with $\det \Lambda = 1$ and $A^0{}_0 \geq 1$. Transformations with $\det \Lambda = 1$ form a subgroup and are called SO(3, 1), if they also have $A^0{}_0 \geq 1$ (orthochronous), then they also form a subgroup and are called the proper Lorentz transformations and denoted by $\mathrm{SO}^+(3,1)$. They consists of Lorentz boosts, example in the x-direction:

$$\Lambda^{\mu}{}_{\nu} = \begin{pmatrix} \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & -\frac{\frac{v}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 \\ -\frac{v}{\sqrt{1-\frac{v^2}{c^2}}} & \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

which in the limit $c \to \infty$ gives

$$\Lambda^{\mu}{}_{\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -v & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and spatial rotations:

$$R_1(\phi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\phi & \sin\phi \\ 0 & 0 & -\sin\phi & \cos\phi \end{pmatrix}$$

$$R_2(\phi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\phi & 0 & \sin\phi \\ 0 & 0 & 1 & 0 \\ 0 & -\sin\phi & 0 & \cos\phi \end{pmatrix}$$

$$R_3(\phi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\phi & \sin\phi & 0 \\ 0 & -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(More rigorous derivation will be given in a moment.) It can be shown (see below), that all other elements (improper Lorentz transformations) of the O(3, 1) group can be written as products of an element from $\mathrm{SO}^+(3,1)$ and an element of the discrete group:

$$\{\mathbb{1},\ P,\ T,\ PT\}$$

where P is parity (also called space reflection or space inversion):

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

and T is time reversal (also called time inversion):

$$T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Second case: elements with $\det\Lambda = 1$ and ${A^0}_0 \leq -1$. An example of such an element is PT. In general, any product from $\mathrm{SO}^+(3,1)$ and PT belongs here.

Third case: elements with $\det\Lambda = -1$ and ${A^0}_0 \geq 1$. An example of such an element is P. In general, any product from $\mathrm{SO}^+(3,1)$ and P belongs here.

Fourth case: elements with $\det\Lambda = -1$ and ${A^0}_0 \leq -1$. An example of such an element is T. In general, any product from $\mathrm{SO}^+(3,1)$ and T belongs here.

Example: where does the reflection around a single spatial axis $(t, x, y, z) \to (t, -x, y, z)$ belong to? It is the third case, because the determinant is $\det\Lambda = -1$ and the 00 element is 1. Written in the matrix form:

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\pi & \sin\pi \\ 0 & 0 & -\sin\pi & \cos\pi \end{pmatrix} = PR_1(\pi)$$

So it is constructed using the R_1 element from $\mathrm{SO}^+(3,1)$ and P from the discrete group above.

We can now show why the decomposition $\mathrm{O}(3,1) = \mathrm{SO}^+(3,1) \times \{\mathbb{1},\ P,\ T,\ PT\}$ works. Note that $PT = -\mathbb{1}$. First we show that $\mathrm{SO}(3,1) = \mathrm{SO}^+(3,1) \times \{\mathbb{1}, -\mathbb{1}\}$. This follows from the fact, that all matrices with ${\Lambda^0}_0 \leq -1$ can be written using $-\mathbb{1}$ and a matrix with ${\Lambda^0}_0 \geq 1$. All matrices with $\det\Lambda = -1$ can be constructed from a matrix with $\det\Lambda = 1$ (i.e. SO(3, 1)) and a diagonal matrix with odd number of -1, below we list all of them together with their

construction using time reversal, parity and spatial rotations:

$$\begin{aligned}
\operatorname{diag}(-1,1,1,1) &= T \\
\operatorname{diag}(1,-1,1,1) &= PR_1(\pi) \\
\operatorname{diag}(1,1,-1,1) &= PR_2(\pi) \\
\operatorname{diag}(1,1,1,-1) &= PR_3(\pi) \\
\operatorname{diag}(1,-1,-1,-1) &= P \\
\operatorname{diag}(-1,1,-1,-1) &= TR_1(\pi) \\
\operatorname{diag}(-1,-1,1,-1) &= TR_2(\pi) \\
\operatorname{diag}(-1,-1,-1,1) &= TR_3(\pi)
\end{aligned}$$

But $R_i(\pi)$ belongs to $\mathrm{SO}^+(3,1)$, so we just need two extra elements, T and P to construct all matrices with $\det\Lambda = -1$ using matrices from SO(3, 1). So to recapitulate, if we start with $\mathrm{SO}^+(3,1)$ we need to add the element $PT = -\mathbb{1}$ to construct SO(3, 1) and then we need to add P and T to construct O(3, 1). Because all other combinations like $PPT = T$ reduce to just one of $\{\mathbb{1}, P, T, -\mathbb{1}\}$, we are done.

The elements from $\mathrm{SO}^+(3,1)$ are proper Lorentz transformations, all other elements are improper. Now we'd like to construct the proper Lorentz transformation matrix A explicitly. As said above, all improper transformations are just proper transformations multiplied by either P, T or PT, so it is sufficient to construct A.

We can always write $A = e^L$, then:

$$\det A = \det e^L = e^{\operatorname{Tr} L} = 1$$

so $\operatorname{Tr} L = 0$ and L is a real, traceless matrix. Rewriting (4.3.11.1):

$$\begin{aligned}
g &= A^T g A \\
A^{-1} &= g^{-1} A^T g \\
e^{-L} &= g^{-1} e^{L^T} g = e^{g^{-1} L^T g} \\
-L &= g^{-1} L^T g \\
-gL &= (gL)^T
\end{aligned}$$

The matrix gL is thus antisymmetric and the general form of L is then:

$$L = \begin{pmatrix} 0 & \frac{L_{01}}{c^2} & \frac{L_{02}}{c^2} & \frac{L_{03}}{c^2} \\ L_{01} & 0 & L_{12} & L_{13} \\ L_{02} & -L_{12} & 0 & L_{23} \\ L_{03} & -L_{13} & -L_{23} & 0 \end{pmatrix}$$

One can check, that gL is indeed antisymmetric. However, for a better parametrization, it's better to work with a metric $\operatorname{diag}(-1,1,1,1)$, which can be achieved by putting c into (ct, x, y, z), or equivalently, to work with $x^\mu = (t, x, y, z)$ and multiply this by a matrix $C = \operatorname{diag}(c, 1, 1, 1)$ to get (ct, x, y, z). To get a symmetric $\tilde{L}$, we just have to do $Cx' = \tilde{L}Cx$, so to get an unsymmetric L from the symmetric one, we need to do $C^{-1}\tilde{L}C$, so we get:

$$L = C^{-1} \begin{pmatrix} 0 & \zeta_1 & \zeta_2 & \zeta_3 \\ \zeta_1 & 0 & -\varphi_3 & \varphi_2 \\ \zeta_2 & \varphi_3 & 0 & -\varphi_1 \\ \zeta_3 & -\varphi_2 & \varphi_1 & 0 \end{pmatrix} C = -i\boldsymbol{\varphi} \cdot \mathbf{L} - i\boldsymbol{\zeta} \cdot C^{-1}\mathbf{M}C$$

We have parametrized all the proper Lorentz transformations with just 6 parameters ζ_1, ζ_2, ζ_3, φ_1, φ_2 and φ_3. The matrices $\mathbf{L}$ and $\mathbf{M}$ are defined as:

$$L_1 = -i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

$$L_2 = -i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$L_3 = -i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$M_1 = i \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$M_2 = i \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$M_3 = i \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Straightforward calculation shows:

$$[L_i, L_j] = i\epsilon_{ijk} L_k$$

$$[L_i, M_j] = i\epsilon_{ijk} M_k$$

$$[M_i, M_j] = -i\epsilon_{ijk} L_k$$

The first relation corresponds to the commutation relations for angular momentum, second relation shows that M transforms as a vector under rotations and the final relation shows that boosts do not in general commute.

We get:

$$A = e^{-i\boldsymbol{\varphi}\cdot\mathbf{L} - i\boldsymbol{\zeta}\cdot C^{-1}\mathbf{M}C} = C^{-1}\, e^{-i\boldsymbol{\varphi}\cdot\mathbf{L} - i\boldsymbol{\zeta}\cdot\mathbf{M}}\, C$$

As a special case, the rotation around the z-axis is given by $\boldsymbol{\varphi} = (0, 0, \varphi)$ and $\boldsymbol{\zeta} = 0$:

$$A = e^{-i\varphi L_3} = \mathbb{1} - L_3^2 + iL_3 \sin\varphi + L_3^2 \cos\varphi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi & 0 \\ 0 & -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The boost in the x-direction is $\boldsymbol{\varphi} = 0$ and $\boldsymbol{\zeta} = (\zeta, 0, 0)$, e.g.:

$$A = C^{-1} e^{-i\zeta M_1} C = C^{-1} \left(\mathbb{1} + M_1^2 - iM_1 \sinh\zeta - M_1^2 \cosh\zeta\right) C =$$

$$= C^{-1} \begin{pmatrix} \cosh\zeta & \sinh\zeta & 0 & 0 \\ \sinh\zeta & \cosh\zeta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} C = \begin{pmatrix} \cosh\zeta & \frac{1}{c}\sinh\zeta & 0 & 0 \\ c\sinh\zeta & \cosh\zeta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

from the construction, $-\infty < \zeta < \infty$, so we may do the substitution $\zeta = \operatorname{atanh}\left(\frac{v}{c}\right)$, where $-c < v < c$. The inverse transformation is:

$$\cosh\zeta = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

$$\sinh\zeta = \frac{\frac{v}{c}}{\sqrt{1-\frac{v^2}{c^2}}}$$

and we get the boost given above:

$$A = \begin{pmatrix} \cosh\zeta & \frac{1}{c}\sinh\zeta & 0 & 0 \\ c\sinh\zeta & \cosh\zeta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & \frac{\frac{v}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 \\ \frac{v}{\sqrt{1-\frac{v^2}{c^2}}} & \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Depending on the sign of v, we can also put a minus sign in front of the off-diagonal elements.

Adding two boosts together:

$$A(u)A(v) = \begin{pmatrix} \frac{1}{\sqrt{1-\frac{u^2}{c^2}}} & -\frac{\frac{u}{c^2}}{\sqrt{1-\frac{u^2}{c^2}}} & 0 & 0 \\ -\frac{u}{\sqrt{1-\frac{u^2}{c^2}}} & \frac{1}{\sqrt{1-\frac{u^2}{c^2}}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & -\frac{\frac{v}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 \\ -\frac{v}{\sqrt{1-\frac{v^2}{c^2}}} & \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{1}{\sqrt{1-\frac{w^2}{c^2}}} & -\frac{\frac{w}{c^2}}{\sqrt{1-\frac{w^2}{c^2}}} & 0 & 0 \\ -\frac{w}{\sqrt{1-\frac{w^2}{c^2}}} & \frac{1}{\sqrt{1-\frac{w^2}{c^2}}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

with

$$w = \frac{u+v}{1+\frac{uv}{c^2}}$$

4.3.12 O(4) Group

The group of rotations in 4 dimensions is O(4), e.g. all matrices satisfying:

$$g = \Lambda^T g \Lambda \tag{4.3.12.1}$$

with $g = \mathrm{diag}(c^2, 1, 1, 1)$. Taking the determinant of (4.3.12.1) we get $(\det\Lambda)^2 = 1$ or $\det\Lambda = \pm 1$. Writing the 00 component of (4.3.12.1) we get

$$c^2 = c^2 {A^0}_0^2 + {A^0}_1^2 + {A^0}_2^2 + {A^0}_3^2$$

or

$${A^0}_0^2 = 1 - \frac{1}{c^2}\left({A^0}_1^2 + {A^0}_2^2 + {A^0}_3^2\right)$$

Thus we always have $-1 \le {A^0}_0 \le 1$. That is different to the O(3, 1) group: the O(4) group consists of only 2 continuous parts, that are not connected. (The SO(4) part contains the element $-\mathbb{1}$ though, but one can get to it continuously, so the group is doubly connected.)

Everything proceeds much like for the O(3, 1) group, so gL is antisymmetric, but this time $g = \mathrm{diag}(c^2, 1, 1, 1)$, so we get:

$$L = \begin{pmatrix} 0 & -\frac{L_{01}}{c^2} & -\frac{L_{02}}{c^2} & -\frac{L_{03}}{c^2} \\ L_{01} & 0 & L_{12} & L_{13} \\ L_{02} & -L_{12} & 0 & L_{23} \\ L_{03} & -L_{13} & -L_{23} & 0 \end{pmatrix}$$

and so we also have 6 generators, but this time all of them are rotations:

$$A = C^{-1}\, e^{-i\varphi_a L_a}\, C$$

with $a = 1, 2, 3, 4, 5, 6$. The spatial rotations are the same as for O(3, 1) and the remaining 3 rotations are (t, x), (t, y) and (t, z) plane rotations. So for example the (t, x) rotation is:

$$A = C^{-1} \begin{pmatrix} \cos\varphi_4 & \sin\varphi_4 & 0 & 0 \\ -\sin\varphi_4 & \cos\varphi_4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} C = \begin{pmatrix} \cos\varphi_4 & \frac{1}{c}\sin\varphi_4 & 0 & 0 \\ -c\sin\varphi_4 & \cos\varphi_4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Now we can do this identification:

$$\sin\phi_4 = \frac{\frac{v}{c}}{\sqrt{1 + (\frac{v}{c})^2}}$$

$$\cos\phi_4 = \frac{1}{\sqrt{1 + (\frac{v}{c})^2}}$$

so we get the Galilean transformation in the limit $c \to \infty$:

$$A = \begin{pmatrix} \frac{1}{\sqrt{1+(\frac{v}{c})^2}} & \frac{\frac{v}{c^2}}{\sqrt{1+(\frac{v}{c})^2}} & 0 & 0 \\ -\frac{v}{\sqrt{1+(\frac{v}{c})^2}} & \frac{1}{\sqrt{1+(\frac{v}{c})^2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 0 \\ -v & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Adding two boosts together:

$$A(u)A(v) = \begin{pmatrix} \frac{1}{\sqrt{1+\frac{u^2}{c^2}}} & \frac{\frac{u}{c^2}}{\sqrt{1+\frac{u^2}{c^2}}} & 0 & 0 \\ -\frac{u}{\sqrt{1+\frac{u^2}{c^2}}} & \frac{1}{\sqrt{1+\frac{u^2}{c^2}}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{1+\frac{v^2}{c^2}}} & \frac{\frac{v}{c^2}}{\sqrt{1+\frac{v^2}{c^2}}} & 0 & 0 \\ -\frac{v}{\sqrt{1+\frac{v^2}{c^2}}} & \frac{1}{\sqrt{1+\frac{v^2}{c^2}}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{1}{\sqrt{1+\frac{w^2}{c^2}}} & \frac{\frac{w}{c^2}}{\sqrt{1+\frac{w^2}{c^2}}} & 0 & 0 \\ -\frac{w}{\sqrt{1+\frac{w^2}{c^2}}} & \frac{1}{\sqrt{1+\frac{w^2}{c^2}}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

with

$$w = \frac{u + v}{1 - \frac{uv}{c^2}}$$

However, there is one peculiar thing here that didn't exist in the O(3, 1) case: by adding two velocities less than c, for example $u = v = c/2$, we get:

$$w = \frac{c}{1 - \frac{1}{4}} = \frac{4c}{3} > c$$

(as opposed to $w = \frac{c}{1+\frac{1}{4}} = \frac{4c}{5} < c$ in the O(3, 1) case). So one can get over c easily. By adding $u = v = \frac{4c}{3}$ together:

$$w = \frac{\frac{8c}{3}}{1 - \frac{16}{9}} = -\frac{24c}{7} < 0$$

(as opposed to $w = \frac{\frac{8c}{3}}{1+\frac{16}{9}} = \frac{24c}{25} > 0$ in the O(3, 1) case). So we can also get to negative speeds easily. One also needs to be careful with identifying $\cos\phi_4 = \frac{1}{\sqrt{1+(\frac{v}{c})^2}}$, because for $\varphi_4 > \pi/2$ we should probably set $\cos\varphi_4 = -\frac{1}{\sqrt{1+(\frac{v}{c})^2}}$. All of this follows directly from the structure of SO(4), because one can get from ${\Lambda^0}_0 > 0$ to ${\Lambda^0}_0 < 0$ continuously (this corresponds to increasing φ_4 over $\pi/2$). In fact, by adding two speeds $u = v > c(\sqrt{2}-1)$, one always gets $w > c$. But if $c(\sqrt{2}-1) \doteq 0.414c$ is larger than any speed that we are concerned about, we are fine.

4.3.13 Proper Time

Proper time τ is a time elapsed by (physical) clocks along some (4D) trajectory. Coordinate time t is just some time coordinate assigned to each point in the space and usually one can find some real clocks, that would measure such a time (many times they are in the infinity). To find a formula for a proper time (in terms of the coordinate time), we introduce a local inertial frame at each point of the trajectory – in this frame, the clocks do not move, e.g. x, y, z is constant (zero) and there is no gravity (this follows from the definition of the local inertial frame), so the metric is just a Minkowski metric.

For any metrics, $\mathrm{d}s^2$ is invariant:

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu$$

so coming to the local inertial frame, we have x, y, z constant and we get:

$$\mathrm{d}s^2 = g_{00}\mathrm{d}\tau^2$$

so:

$$\mathrm{d}\tau = \sqrt{\frac{\mathrm{d}s^2}{g_{00}}}$$

since we are still in the local inertial frame (e.g. no gravity), we have $g_{00} = -c^2$ (depending on which metrics we take it could also be $+c^2$), so:

$$\mathrm{d}\tau = \sqrt{-\frac{\mathrm{d}s^2}{c^2}}$$

This formula was derived in the local inertial frame, but the right hand side is the same in any inertial frame, because $\mathrm{d}s^2$ is invariant and c too. So in any frame we have:

$$\mathrm{d}\tau = \sqrt{-\frac{\mathrm{d}s^2}{c^2}} = \sqrt{-\frac{g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu}{c^2}}$$

We'll explain how to calculate the proper time on the 1971 Hafele and Keating experiment. They transported cesium-beam atomic clocks around the Earth on scheduled commercial flights (once flying eastward, once westward) and compared their reading on return to that of a standard clock at rest on the Earth's surface.

We'll calculate it with all the metrics discussed above, to see the difference.

Weak Field Metric

Let's start with the metrics:

$$\mathrm{d}s^2 = -\left(1+\frac{2\phi}{c^2}\right)c^2\mathrm{d}t^2 + \left(1-\frac{2\phi}{c^2}\right)(\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2)$$

Then:

$$\tau_{AB} = \int_A^B \mathrm{d}\tau = \int_A^B \sqrt{-\frac{\mathrm{d}s^2}{c^2}} = \int_A^B \sqrt{\left(1+\frac{2\phi}{c^2}\right)\mathrm{d}t^2 - \frac{1}{c^2}\left(1-\frac{2\phi}{c^2}\right)(\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2)} =$$

$$= \int_A^B \mathrm{d}t \sqrt{\left(1+\frac{2\phi}{c^2}\right) - \frac{1}{c^2}\left(1-\frac{2\phi}{c^2}\right)\left(\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2\right)} =$$

$$= \int_A^B \mathrm{d}t \sqrt{\left(1+\frac{2\phi}{c^2}\right) - \frac{1}{c^2}\left(1-\frac{2\phi}{c^2}\right)|\mathbf{V}|^2}$$

where

$$|\mathbf{V}|^2 = \left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2$$

is the nonrelativistic velocity. Then we expand the square root into power series and only keep terms with low powers of c:

$$\tau_{AB} = \int_A^B \mathrm{d}t \sqrt{\left(1+\frac{2\phi}{c^2}\right) - \frac{1}{c^2}\left(1-\frac{2\phi}{c^2}\right)|\mathbf{V}|^2} = \int_A^B \mathrm{d}t \left(1+\frac{\phi}{c^2} - \frac{1}{2c^2}|\mathbf{V}|^2\right)$$

so

$$\tau_{AB} = \int_A^B \mathrm{d}t \left(1 - \frac{1}{c^2}\left(\frac{1}{2}|\mathbf{V}|^2 - \phi\right)\right)$$

Now let $V_g = V_g(t)$ be the speed of the plane relative to the (rotating) Earth (positive for the eastbound flights, negative for the westbound ones), $V_\oplus = \frac{2\pi R_\oplus}{24}\,\frac{1}{\mathrm{h}}$ the surface speed of the Earth, then the proper time for the clocks on the surface is:

$$\tau_\oplus = \int_A^B \mathrm{d}t \left(1 - \frac{1}{c^2}\left(\frac{1}{2}V_\oplus^2 - \phi_\oplus\right)\right)$$

and for the clocks in the plane

$$\tau = \int_A^B \mathrm{d}t \left(1 - \frac{1}{c^2}\left(\frac{1}{2}(V_g+V_\oplus)^2 - \phi\right)\right)$$

then the difference between the proper times is:

$$\tau - \tau_\oplus = \Delta\tau = \frac{1}{c^2}\int_A^B \mathrm{d}t \left(-\frac{1}{2}(V_g+V_\oplus)^2 + \phi + \frac{1}{2}V_\oplus^2 - \phi_\oplus\right) = \frac{1}{c^2}\int_A^B \mathrm{d}t \left(\phi - \phi_\oplus - \frac{1}{2}V_g(V_g+2V_\oplus)\right)$$

but $\phi - \phi_\oplus = gh$, where $h = h(t)$ is the altitude of the plane, so the final formula is:

$$\Delta\tau = \frac{1}{c^2}\int_A^B \mathrm{d}t \left(gh - \frac{1}{2}V_g(V_g+2V_\oplus)\right)$$

Let's evaluate it for typical altitudes and speeds of commercial aircrafts:

$$R_\oplus = 6378.1\,\mathrm{km} = 6.3781\cdot 10^6\,\mathrm{m}$$

$$V_\oplus = \frac{2\pi R_\oplus}{24}\,\frac{1}{\mathrm{h}} = \frac{2\pi R_\oplus}{24\cdot 3600}\,\frac{1}{\mathrm{s}} = \frac{2\pi\,6.3781\cdot 10^6}{24\cdot 3600}\,\frac{\mathrm{m}}{\mathrm{s}} = 463.83\,\frac{\mathrm{m}}{\mathrm{s}}$$

$$V_g = 870\,\frac{\mathrm{km}}{\mathrm{h}} = 241.67\,\frac{\mathrm{m}}{\mathrm{s}}$$

$$h = 12\,\mathrm{km} = 12000\,\mathrm{m}$$

$$t = \frac{2\pi R_\oplus}{V_g} = \frac{2\pi\,6.3781\cdot 10^6}{241.67}\,\mathrm{s} = 165824.41\,\mathrm{s} \approx 46\,\mathrm{h}$$

$$c = 3 \cdot 10^8 \, \frac{\mathrm{m}}{\mathrm{s}}$$

For eastbound flights we get:

$$\Delta\tau = \frac{t}{c^2}\left(gh - \frac{1}{2}V_g(V_g + 2V_\oplus)\right) = -4.344 \cdot 10^{-8}\,\mathrm{s} = -43.44\,\mathrm{ns}$$

and for westbound flights we get:

$$\Delta\tau = \frac{t}{c^2}\left(gh - \frac{1}{2}V_g(V_g - 2V_\oplus)\right) = 3.6964 \cdot 10^{-7}\,\mathrm{s} = 369.63\,\mathrm{ns}$$

By neglecting gravity, one would get: eastbound flights:

$$\Delta\tau = \frac{t}{c^2}\left(-\frac{1}{2}V_g(V_g + 2V_\oplus)\right) = -260.34\,\mathrm{ns}$$

and for westbound flights:

$$\Delta\tau = \frac{t}{c^2}\left(-\frac{1}{2}V_g(V_g - 2V_\oplus)\right) = 152.73\,\mathrm{ns}$$

By just taking the clocks to the altitude $12\,\mathrm{km}$ and staying there for 46 hours (without moving with respect to the inertial frame, e.g. far galaxies), one gets:

$$\Delta\tau = \frac{ght}{c^2} = 216.90\,\mathrm{ns}$$

Rotating Disk Metric

The rotating disk metrics is (taking weak field gravitation into account):

$$\mathrm{d}s^2 = -\left(1 + \frac{2\phi}{c^2} - \frac{\omega^2}{c^2}(x^2 + y^2)\right)c^2\mathrm{d}t^2 + (\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2) - 2\omega y\,\mathrm{d}x\mathrm{d}t + 2\omega x\,\mathrm{d}y\mathrm{d}t$$

Then:

$$\tau_{AB} = \int_A^B \mathrm{d}\tau = \int_A^B \sqrt{-\frac{\mathrm{d}s^2}{c^2}} =$$

$$= \int_A^B \sqrt{\left(1 + \frac{2\phi}{c^2} - \frac{\omega^2}{c^2}(x^2 + y^2)\right)\mathrm{d}t^2 - \frac{1}{c^2}(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2) + \frac{2\omega y}{c^2}\,\mathrm{d}x\mathrm{d}t - \frac{2\omega x}{c^2}\,\mathrm{d}y\mathrm{d}t} =$$

$$= \int_A^B \mathrm{d}t\sqrt{\left(1 + \frac{2\phi}{c^2} - \frac{\omega^2}{c^2}(x^2 + y^2)\right) - \frac{1}{c^2}|\mathbf{V}|^2 + \frac{2\omega y}{c^2}\,\frac{\mathrm{d}x}{\mathrm{d}t} - \frac{2\omega x}{c^2}\,\frac{\mathrm{d}y}{\mathrm{d}t}}$$

where

$$|\mathbf{V}|^2 = \left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2$$

is the nonrelativistic velocity. Then we expand the square root into power series and only keep terms with low powers of c:

$$\tau_{AB} = \int_A^B \mathrm{d}t\left(1 + \frac{\phi}{c^2} - \frac{1}{2c^2}|\mathbf{V}|^2 + \frac{\omega y}{c^2}\,\frac{\mathrm{d}x}{\mathrm{d}t} - \frac{\omega x}{c^2}\,\frac{\mathrm{d}y}{\mathrm{d}t}\right)$$

so

$$\tau_{AB} = \int_A^B \mathrm{d}t\left(1 - \frac{1}{c^2}\left(\frac{1}{2}|\mathbf{V}|^2 - \phi - \omega y\,\frac{\mathrm{d}x}{\mathrm{d}t} + \omega x\,\frac{\mathrm{d}y}{\mathrm{d}t}\right)\right)$$

Now as before let $V_g = V_g(t)$ be the speed of the plane (relative to the rotating Earth, e.g. relative to our frame), $V_\oplus = \frac{2\pi R_\oplus}{24}\,\frac{1}{\mathrm{h}}$ the surface speed of the Earth, so $\omega R_\oplus = V_\oplus$. For the clocks on the surface, we have:

$$x = R_\oplus$$

$$y = 0$$

$$z = 0$$

so

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\mathrm{d}z}{\mathrm{d}t} = 0$$

$$|\mathbf{V}|^2 = 0$$

then the proper time for the clocks on the surface is:

$$\tau_\oplus = \int_A^B \mathrm{d}t \left(1 - \frac{1}{c^2}\left(-\phi_\oplus\right)\right)$$

and for the clocks in the plane we have:

$$x = (R_\oplus + h)\cos\Omega t$$

$$y = (R_\oplus + h)\sin\Omega t$$

$$z = 0$$

where Ω is defined by $\Omega(R_\oplus + h) = V_g$, so

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -(R_\oplus + h)\Omega\sin\Omega t$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = (R_\oplus + h)\Omega\cos\Omega t$$

$$\frac{\mathrm{d}z}{\mathrm{d}t} = 0$$

$$|\mathbf{V}|^2 = \Omega^2(R_\oplus + h)^2$$

$$\omega y\frac{\mathrm{d}x}{\mathrm{d}t} = -\omega\Omega(R_\oplus + h)^2\sin^2\Omega t$$

$$\omega x\frac{\mathrm{d}y}{\mathrm{d}t} = \omega\Omega(R_\oplus + h)^2\cos^2\Omega t$$

and

$$\tau = \int_A^B \mathrm{d}t \left(1 - \frac{1}{c^2}\left(\frac{1}{2}\Omega^2(R_\oplus + h)^2 - \phi + \omega\Omega(R_\oplus + h)^2\right)\right)$$

then the difference between the proper times is:

$$\tau - \tau_\oplus = \Delta\tau = \frac{1}{c^2}\int_A^B \mathrm{d}t \left(-\frac{1}{2}\Omega^2(R_\oplus + h)^2 - \omega\Omega(R_\oplus + h)^2 + \phi - \phi_\oplus\right) =$$

$$= \frac{1}{c^2}\int_A^B \mathrm{d}t \left(-\frac{1}{2}V_g^2 - V_\oplus V_g\left(1 + \frac{h}{R_\oplus}\right) + \phi - \phi_\oplus\right) =$$

$$= \frac{1}{c^2}\int_A^B \mathrm{d}t \left(\phi - \phi_\oplus - \frac{1}{2}V_g\left(V_g + 2V_\oplus\left(1 + \frac{h}{R_\oplus}\right)\right)\right)$$

but $\phi - \phi_\oplus = gh$, where $h = h(t)$ is the altitude of the plane and we approximate

$$\left(1 + \frac{h}{R_\oplus}\right) \approx 1\,,$$

so the final formula is the same as before:

$$\Delta\tau = \frac{1}{c^2}\int_A^B \mathrm{d}t \left(gh - \frac{1}{2}V_g(V_g + 2V_\oplus)\right)$$

Note: for the values above, the bracket $\left(1 + \frac{h}{R_\oplus}\right)^2 \doteq 1.00377$, so it's effect on the final difference of the proper times is negligible (e.g. less than $1\,\mathrm{ns}$). The difference is caused by a slightly vague definition of the speed of the plane, e.g. the ground speed is a bit different to the speed relative to the rotating Earth (this depends on how much the atmosphere rotates with the Earth).

Concluding Remarks

The coordinate time t in both cases above is totally different. One can find some physical clocks in both cases that measure (e.g. whose proper time is) the particular coordinate time, but the beauty of the differential geometry approach is that we don't have to care about this. t is just a coordinate, that we use to calculate something physical, like a proper time along some trajectory, which is a frame invariant quantity. In both cases above, we got a different formulas for the proper time of the surface clocks (and the clocks in the plane) in terms of the coordinate time (because the coordinate time is different in both cases), however the difference of the proper times is the same in both cases:

$$\Delta\tau = \frac{1}{c^2}\int_A^B \mathrm{d}t \left(gh - \frac{1}{2}V_g(V_g + 2V_\oplus)\right)$$

There is still a slight difference though – the t here used to evaluate the integral is different in both cases. To do it correctly, one should take the total time as measured by any of the clocks and then use the right formula for the proper time of the particular clock to convert to the particular coordinate time. However, the difference is small, of the order of nanoseconds, so it's negligible compared to the total flying time of 46 hours.

4.3.14 FAQ

How does one incorporate the fact, that there are only two possible transformations, into all of this? For more info, see: http://arxiv.org/abs/0710.3398. Answer: in that article there are actually three possible transformations, $K < 0$ corresponds to O(4), $K > 0$ to O(3, 1) and $K = 0$ to either of them in the limit $c \to \infty$.

What is the real difference between the Newtonian physics and special relativity? E.g. how do we derive the Minkowski metrics, how do we know we need to set $c = const$ and how do we incorporate gravity in it? Answer: there are only three possible groups of transformations: O(4), O(3, 1) and a limit of either for $c \to \infty$. All three provide inequivalent predictions for high speeds, so we just choose the right one by experiment. It happens to be the O(3, 1). As to gravity, that can be incorporated in either of them.

4.3.15 Questions Without Answers (Yet)

How can one reformulate the article http://arxiv.org/abs/0710.3398 into the language of the O(4) and O(3, 1) groups above? Basically each assumption and equation must have some counterpart in what we have said above. I'd like to identify those explicitely.

What are all the possible metrics, that generate the Newtonian Christoffel symbols? (Several such are given above, but I want to know all of them) Probable answer: all metrics, whose inverse reduces to $g^{\mu\nu} = \mathrm{diag}(0, 1, 1, 1)$ in the limit $c \to \infty$. I would like to have an explicit proof of this though.

What is the role of the different metrics, that generate the same Christoffel symbols in the limit ($c \to \infty$)? Can one inertial frame be given with one and another frame with a different form of the metrics (e.g. one with $g_{00} = c^2$ and the other one with $g_{00} = -c^2$?) Possible answer: there is no transformation to convert a metrics with signature +4 to signature +2, so one has to choose one and then all other inertial frames have the same one.

What are all the allowed transformations between inertial frames? If we assume that the inertial frames are given with one given metrics (see the previous question), then the answer is: representation of the O(3, 1) group if $g_{00} = -c^2$ or O(4) group if $g_{00} = c^2$. But if one frame is $g_{00} = -c^2$ and we transform to another frame with $g_{00} = c^2$, then it is not clear what happens. Possible answer: one has to choose some signature and stick to it, see also the previous question.

What is the real difference between Newtonian physics and general relativity? Given our formulation of Newtonian physics using the differential geometry, I want to know what the physical differences are between all the three theories are.

CHAPTER

FIVE

CLASSICAL ELECTROMAGNETISM

5.1 Maxwell's Equations

5.1.1 Electromagnetic Field

The electromagnetic field is fully described by a vector field called the 4-potential A^α. It has four components that we can label any way we want, the traditional way is to use:

$$A^\alpha = \left(\frac{\phi}{c}, \mathbf{A}\right)$$

where ϕ is called the electrostatic scalar potential, $\mathbf{A}$ is called the vector potential and c is the speed of light. The Lagrangian density for the free (noninteracting) field is:

$$\mathcal{L} = -\frac{1}{2\mu_0}\partial_\alpha A_\beta \partial^\alpha A^\beta$$

The Lagrangian for a (charged) particle is:

$$L(x^\mu, v^\mu) = -\tfrac{1}{2} m v_\alpha v^\alpha$$

it produces the following charge density:

$$\rho = q\delta(\mathbf{r} - \mathbf{s})$$

The interaction between the charged particle (or in general any charged body) with some charge density and the electromagnetic field is given by the Lagrangian density (this follows from *Local Gauge Invariance*):

$$\mathcal{L} = -j_\alpha A^\alpha$$

where:

$$j^\mu = \rho v^\mu = \gamma\rho(c, \mathbf{v})$$

All together, the Lagrangian of a charged particle and an electromagnetic field is:

$$\begin{aligned} L(x^\mu, v^\mu, A^\mu, \partial_\nu A^\mu) &= -\tfrac{1}{2} m v_\alpha v^\alpha - \int \frac{1}{2\mu_0}\partial_\alpha A_\beta \partial^\alpha A^\beta \mathrm{d}^3x - \int j_\alpha A^\alpha \mathrm{d}^3x = \\ &= -\tfrac{1}{2} m v_\alpha v^\alpha - \int \frac{1}{2\mu_0}\partial_\alpha A_\beta \partial^\alpha A^\beta \mathrm{d}^3x - \int \rho v_\alpha A^\alpha \mathrm{d}^3x = \\ &= -\tfrac{1}{2} m v_\alpha v^\alpha - \int \frac{1}{2\mu_0}\partial_\alpha A_\beta \partial^\alpha A^\beta \mathrm{d}^3x - q v_\alpha A^\alpha \end{aligned} \tag{5.1.1.1}$$

Note that:

$$v_\alpha A^\alpha = -\gamma\phi + \gamma \mathbf{v}\cdot\mathbf{A}$$

There are several approaches how to obtain the above Lagrangian from some other assumptions, but ultimately the exact form of the Lagrangians has to be given by experiment. This Lagrangian is our only assumption and we derive everything else from it.

The Euler-Lagrange equations for the electromagnetic field (in terms of A^μ and $\partial_\nu A^\mu$) are:

$$\begin{aligned}
\partial^\mu \frac{\partial}{\partial(\partial^\mu A^\nu)}\left(-\frac{1}{2\mu_0}\partial_\alpha A_\beta \partial^\alpha A^\beta - j_\alpha A^\alpha\right) &= \frac{\partial}{\partial A^\nu}\left(-\frac{1}{2\mu_0}\partial_\alpha A_\beta \partial^\alpha A^\beta - j_\alpha A^\alpha\right)\\
\partial^\mu \frac{\partial}{\partial(\partial^\mu A^\nu)}\left(\frac{1}{2\mu_0} g_{\delta\alpha} g_{\epsilon\beta}\partial^\delta A^\epsilon \partial^\alpha A^\beta\right) &= \frac{\partial}{\partial A^\nu} j_\alpha A^\alpha\\
\frac{1}{2\mu_0}\partial^\mu g_{\delta\alpha} g_{\epsilon\beta}\left(\delta^\delta_\mu \delta^\epsilon_\nu \partial^\alpha A^\beta + \partial^\delta A^\epsilon \delta^\alpha_\mu \delta^\beta_\nu\right) &= j_\alpha \delta^\alpha_\nu\\
\frac{1}{2\mu_0}\partial^\mu\left(g_{\mu\alpha} g_{\nu\beta}\partial^\alpha A^\beta + g_{\delta\mu} g_{\epsilon\nu}\partial^\delta A^\epsilon\right) &= j_\nu\\
\frac{1}{2\mu_0}\partial^\mu\left(\partial_\mu A_\nu + \partial_\mu A_\nu\right) &= j_\nu\\
\frac{1}{\mu_0}\partial^\mu \partial_\mu A_\nu &= j_\nu\\
\partial^\mu \partial_\mu A_\nu &= \mu_0 j_\nu
\end{aligned} \tag{5.1.1.2}$$

Equations for the particle (in terms of x^μ and v^μ) are:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial L}{\partial v_\mu} &= \frac{\partial L}{\partial x_\mu}\\
\frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial}{\partial v_\mu}\left(\tfrac{1}{2} m g^{\alpha\beta} v_\alpha v_\beta + q v_\alpha A^\alpha\right) &= q v_\alpha \frac{\partial A^\alpha}{\partial x_\mu}\\
\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\tfrac{1}{2} m g^{\alpha\beta}(\delta_{\alpha\mu} v_\beta + v_\alpha \delta_{\beta\mu}) + q\delta_{\alpha\mu} A^\alpha\right) &= q v_\alpha \frac{\partial A^\alpha}{\partial x_\mu}\\
\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\tfrac{1}{2} m(g^{\mu\beta} v_\beta + g^{\alpha\mu} v_\alpha) + q A^\mu\right) &= q v_\alpha \frac{\partial A^\alpha}{\partial x_\mu}\\
\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\tfrac{1}{2} m(v^\mu + v^\mu) + q A^\mu\right) &= q v_\alpha \frac{\partial A^\alpha}{\partial x_\mu}\\
\frac{\mathrm{d}}{\mathrm{d}\tau}\left(m v^\mu + q A^\mu\right) &= q v_\alpha \frac{\partial A^\alpha}{\partial x_\mu}\\
m\frac{\mathrm{d}v^\mu}{\mathrm{d}\tau} &= q\left(-\frac{\mathrm{d}A^\mu}{\mathrm{d}\tau} + v_\alpha \partial^\mu A^\alpha\right)\\
m\frac{\mathrm{d}v^\mu}{\mathrm{d}\tau} &= q\left(-v_\alpha \partial^\alpha A^\mu + v_\alpha \partial^\mu A^\alpha\right)\\
m\frac{\mathrm{d}v^\mu}{\mathrm{d}\tau} &= q(\partial^\mu A^\alpha - \partial^\alpha A^\mu) v_\alpha\\
m\frac{\mathrm{d}v^\mu}{\mathrm{d}\tau} &= q F^{\mu\alpha} v_\alpha
\end{aligned} \tag{5.1.1.3}$$

Where $F^{\mu\nu}$ is called the electromagnetic field strength tensor:

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$$

The only way to measure the electric field is through its interaction with the charge particle. As such, the actual physical field (that can be measured) is $F^{\mu\nu}$, which is invariant under any gauge transformation:

$$A^\alpha \to A^\alpha + \partial^\alpha \psi$$

where ψ is a gauge function:

$$F^{\mu\nu} \to \partial^\mu(A^\nu + \partial^\nu\psi) - \partial^\nu(A^\mu + \partial^\mu\psi) = \partial^\mu A^\nu - \partial^\nu A^\mu + \partial^\mu\partial^\nu\psi - \partial^\nu\partial^\mu\psi = \partial^\mu A^\nu - \partial^\nu A^\mu = F^{\mu\nu}$$

In other words, two different A^μ related by the gauge transformation represent the exact same physical electromagnetic field (as given by the field tensor). As such, we can modify the Lagrangian by applying the gauge transformation to the field A^μ: this changes the equations of motion for the field (thus the numerical values for A^μ will be different), but doesn't change the equation of motion for the particle, so the change will not have any physical effect (cannot be measured).

By choosing ψ as a solution to the equation $\partial_\mu\partial^\mu\psi = -\partial_\mu A^\mu$, we get:

$$\partial_\mu(A^\mu + \partial^\mu\psi) = \partial_\mu A^\mu + \partial_\mu\partial^\mu\psi = \partial_\mu A^\mu - \partial_\mu A^\mu = 0$$

So for any 4-potential we can find ψ such that the transformed 4-potential A^μ obeys the Lorenz gauge condition $\partial_\mu A^\mu = 0$.

In order to obtain a gauge invariant Lagrangian, we need to express it using $F^{\mu\nu}$ using the following identity:

$$\begin{aligned}
\frac{1}{4}F_{\alpha\beta}F^{\alpha\beta} &= \frac{1}{4}(\partial_\alpha A_\beta - \partial_\beta A_\alpha)(\partial^\alpha A^\beta - \partial^\beta A^\alpha) = \\
&= \frac{1}{4}(\partial_\alpha A_\beta\partial^\alpha A^\beta - \partial_\beta A_\alpha\partial^\alpha A^\beta - \partial_\alpha A_\beta\partial^\beta A^\alpha + \partial_\beta A_\alpha\partial^\beta A^\alpha) = \\
&= \frac{1}{2}(\partial_\alpha A_\beta\partial^\alpha A^\beta - \partial_\beta A_\alpha\partial^\alpha A^\beta) = \\
&= \tfrac{1}{2}\partial_\alpha A_\beta\partial^\alpha A^\beta - \tfrac{1}{2}\partial_\beta A_\alpha\partial^\alpha A^\beta = \\
&= \tfrac{1}{2}\partial_\alpha A_\beta\partial^\alpha A^\beta - \tfrac{1}{2}(\partial^\alpha A_\alpha)^2 - \tfrac{1}{2}\partial_\beta(A_\alpha\partial^\alpha A^\beta - A^\beta\partial^\alpha A_\alpha)
\end{aligned}$$

The 4-divergence $\partial_\beta(A_\alpha\partial^\alpha A^\beta - A^\beta\partial^\alpha A_\alpha)$ doesn't change Euler-Lagrange equations, so we can ignore it. We can see, that in the Lorenz gauge $\partial^\alpha A_\alpha = 0$ the term $\frac{1}{4}F_{\alpha\beta}F^{\alpha\beta}$ (which is gauge invariant) simplifies to the term $\partial_\alpha A_\beta\partial^\alpha A^\beta$ in the Lagrangian (5.1.1.1). The gauge invariant Lagrangian is:

$$L(x^\mu, v^\mu, A^\mu, \partial_\nu A^\mu) = -\tfrac{1}{2}mv_\alpha v^\alpha - \int\frac{1}{4\mu_0}F_{\alpha\beta}F^{\alpha\beta}\mathrm{d}^3x - \int j_\alpha A^\alpha\mathrm{d}^3x \tag{5.1.1.4}$$

The E.-L. equation for the particle doesn't change, the equation for the field becomes:

$$\partial^\mu(\partial_\mu A_\nu - \partial_\nu A_\mu) = \mu_0 j_\nu$$

$$\partial^\mu F_{\mu\nu} = \mu_0 j_\nu \tag{5.1.1.5}$$

Which in Lorenz gauge simplifies to equation (5.1.1.2). In order to write equations of motion in terms of $F^{\mu\nu}$ only, we need another equation for it:

$$\begin{aligned}
\epsilon^{\alpha\beta\gamma\delta}\partial_\gamma F_{\alpha\beta} &= \epsilon^{\alpha\beta\gamma\delta}\partial_\gamma(\partial_\alpha A_\beta - \partial_\beta A_\alpha) = \\
&= \epsilon^{\alpha\beta\gamma\delta}\partial_\gamma\partial_\alpha A_\beta - \epsilon^{\alpha\beta\gamma\delta}\partial_\gamma\partial_\beta A_\alpha = 0
\end{aligned} \tag{5.1.1.6}$$

We used the fact, that the partial derivatives are symmetric in the indices $\gamma\alpha$ and $\gamma\beta$ while $\epsilon^{\alpha\beta\gamma\delta}$ is antisymmetric.

5.1.2 Maxwell's Equations

Maxwell's equations are the equations for the electromagnetic field in terms of the physical field strengh tensor, equations (5.1.1.5) and (5.1.1.6):

$$\partial^\mu F_{\mu\nu} = \mu_0 j_\nu$$
$$\epsilon^{\alpha\beta\gamma\delta}\partial_\gamma F_{\alpha\beta} = 0$$

The field strength tensor is antisymmetric, so it has 6 independent components (we use metric tensor with signature -2):

$$F^{0i} = \partial^0 A^i - \partial^i A^0 = \frac{1}{c}\frac{\partial A^i}{\partial t} + \frac{\partial}{\partial x_i}\frac{\phi}{c} = -\frac{1}{c}\left(-\frac{\partial \phi}{\partial x_i} - \frac{\partial A^i}{\partial t}\right)$$
$$F^{ij} = \partial^i A^j - \partial^j A^i = -\frac{\partial A^j}{\partial x_i} + \frac{\partial A^i}{\partial x_j} = -(\delta^i{}_l \delta^j{}_m - \delta^i{}_m \delta^j{}_l)\frac{\partial A^m}{\partial x_l} = -\epsilon^{ij}{}_k \epsilon^k{}_{lm}\frac{\partial A^m}{\partial x_l}$$

There is freedom in how we label the components. The standard way is to express them using physical fields $\mathbf{E}$ and $\mathbf{B}$ that are introduced by:

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}$$
$$\mathbf{B} = \nabla \times \mathbf{A}$$

or in components:

$$E^i = -\frac{\partial \phi}{\partial x_i} - \frac{\partial A^i}{\partial t}$$
$$B^k = \epsilon^k{}_{lm}\nabla^l A^m$$

Comparing to the above, we get:

$$F^{0i} = -\frac{E^i}{c}$$
$$F^{ij} = -\epsilon^{ij}{}_k B^k$$

In particular:

$$F^{12} = -\epsilon^{12}{}_k B^k = -\epsilon^{12}{}_3 B^3 = -B^3$$
$$F^{13} = -\epsilon^{13}{}_k B^k = -\epsilon^{13}{}_2 B^2 = +B^2$$
$$F^{23} = -\epsilon^{23}{}_k B^k = -\epsilon^{23}{}_1 B^1 = -B^1$$

so we get:

$$F^{\mu\nu} = \begin{pmatrix} 0 & -\frac{E^1}{c} & -\frac{E^2}{c} & -\frac{E^3}{c} \\ \frac{E^1}{c} & 0 & -B^3 & B^2 \\ \frac{E^2}{c} & B^3 & 0 & -B^1 \\ \frac{E^3}{c} & -B^2 & B^1 & 0 \end{pmatrix}$$
$$F_{\mu\nu} = g_{\mu\alpha}g_{\nu\beta}F^{\alpha\beta} = \begin{pmatrix} 0 & \frac{E^1}{c} & \frac{E^2}{c} & \frac{E^3}{c} \\ -\frac{E^1}{c} & 0 & -B^3 & B^2 \\ -\frac{E^2}{c} & B^3 & 0 & -B^1 \\ -\frac{E^3}{c} & -B^2 & B^1 & 0 \end{pmatrix}$$

In terms of $\mathbf{E}$ and $\mathbf{B}$ fields, the Maxwell's equations become:

$$\nabla \cdot \mathbf{E} = c^2 \mu_0 \rho$$
$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t}$$
$$\nabla \cdot \mathbf{B} = 0$$
$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

In Lorenz gauge, the equation for the 4-potential is (5.1.1.2):

$$\partial^\mu \partial_\mu A_\nu = \mu_0 j_\nu$$

The solution to this equation is:

$$A^\beta(\mathbf{x},t) = \frac{\mu_0}{4\pi}\int \frac{j^\beta(\mathbf{y}, t-\frac{|\mathbf{x}-\mathbf{y}|}{c})}{|\mathbf{x}-\mathbf{y}|}\mathrm{d}^3y$$

For scalar potential ($\beta = 0$) we get:

$$\frac{\phi(\mathbf{x},t)}{c} = \frac{\mu_0}{4\pi}\int \frac{c\rho(\mathbf{y}, t-\frac{|\mathbf{x}-\mathbf{y}|}{c})}{|\mathbf{x}-\mathbf{y}|}\mathrm{d}^3y$$

$$\phi(\mathbf{x},t) = \frac{\mu_0 c^2}{4\pi}\int \frac{\rho(\mathbf{y}, t-\frac{|\mathbf{x}-\mathbf{y}|}{c})}{|\mathbf{x}-\mathbf{y}|}\mathrm{d}^3y = \frac{1}{4\pi\epsilon_0}\int \frac{\rho(\mathbf{y}, t-\frac{|\mathbf{x}-\mathbf{y}|}{c})}{|\mathbf{x}-\mathbf{y}|}\mathrm{d}^3y \tag{5.1.2.1}$$

And for vector potential ($\beta = i$) we get:

$$\mathbf{A}(\mathbf{x},t) = \frac{\mu_0}{4\pi}\int \frac{\mathbf{j}(\mathbf{y}, t-\frac{|\mathbf{x}-\mathbf{y}|}{c})}{|\mathbf{x}-\mathbf{y}|}\mathrm{d}^3y \tag{5.1.2.2}$$

5.1.3 Lorentz Force

The equation for the charge particle (5.1.1.3) is:

$$m\frac{\mathrm{d}v^\mu}{\mathrm{d}\tau} = qF^{\mu\alpha}v_\alpha$$

In components:

$$m\frac{\mathrm{d}v^0}{\mathrm{d}\tau} = qF^{0\alpha}v_\alpha = -q\frac{E^i}{c}\gamma v_i$$

$$m\frac{\mathrm{d}v^i}{\mathrm{d}\tau} = qF^{i\alpha}v_\alpha = q\left(-\frac{E^i}{c}v_0 - \epsilon^{ij}{}_k B^k v_j\right) = q\left(\frac{E^i}{c}v^0 + \epsilon^i{}_{jk}B^k v^j\right) = q\gamma\left(E^i + (\mathbf{v}\times\mathbf{B})^i\right)$$

Using coordinate time t and coordinates $\mathbf{x}$ instead of the proper time τ and 4-vector x^μ, we need to rewrite the action:

$$S = \int L(x^\mu, v^\mu)\mathrm{d}\tau = \int \frac{1}{\gamma}L(x^\mu, v^\mu)\mathrm{d}t = \int L_{coord}(\mathbf{x},\mathbf{v})\mathrm{d}t$$

where $L_{coord}(\mathbf{x},\mathbf{v})$ is the Lagrangian expressed in coordinates $\mathbf{x}$ and $\mathbf{v}$ (and thus is not Lorentz invariant):

$$\begin{aligned} L_{coord}(\mathbf{x},\mathbf{v}) &= \frac{1}{\gamma}L(x^\mu, v^\mu) = \\ &= -\frac{mc^2}{\gamma} + \frac{e}{\gamma}v_\alpha A^\alpha = \\ &= -\frac{mc^2}{\gamma} + \frac{e}{\gamma}(-c\gamma A^0 + \gamma v_i A^i) = \\ &= -mc^2\sqrt{1-\frac{v^2}{c^2}} - e\phi + e\mathbf{v}\cdot\mathbf{A} \end{aligned}$$

the particle's canonical momentum $\mathbf{P}$ is:

$$P_i = \frac{\partial L(t)}{\partial v_i} = -mc^2\frac{1}{2\sqrt{1-\frac{v^2}{c^2}}}\left(\frac{-2v_i}{c^2}\right) + eA_i = \frac{mv_i}{\sqrt{1-\frac{v^2}{c^2}}} + eA_i$$

$$\mathbf{P} = \frac{m\mathbf{v}}{\sqrt{1-\frac{v^2}{c^2}}} + e\mathbf{A} = \gamma m\mathbf{v} + e\mathbf{A} = \mathbf{p} + e\mathbf{A}$$

where $\mathbf{p} = \mathbf{P} - e\mathbf{A} = \gamma m\mathbf{v}$ is the kinetic momentum. Euler-Lagrange equations are:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L_{coord}}{\partial v_i} = \frac{\partial L_{coord}}{\partial x_i}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}P_i = \frac{\partial L_{coord}}{\partial x_i}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{mv_i}{\sqrt{1-\frac{v^2}{c^2}}} + eA_i\right) = \frac{\partial}{\partial x_i}\left(-mc^2\sqrt{1-\frac{v^2}{c^2}} - e\phi + e\mathbf{v}\cdot\mathbf{A}\right)$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{mv_i}{\sqrt{1-\frac{v^2}{c^2}}}\right) + e\frac{\mathrm{d}A_i}{\mathrm{d}t} = -e\frac{\partial\phi}{\partial x_i} + e\mathbf{v}\cdot\frac{\partial\mathbf{A}}{\partial x_i}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{mv_i}{\sqrt{1-\frac{v^2}{c^2}}}\right) = e\left(-\frac{\partial\phi}{\partial x_i} - \frac{\mathrm{d}A_i}{\mathrm{d}t} + v_j\frac{\partial A_j}{\partial x_i}\right)$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{mv_i}{\sqrt{1-\frac{v^2}{c^2}}}\right) = e\left(-\frac{\partial\phi}{\partial x_i} - \frac{\partial A_i}{\partial t} - v_j\frac{\partial A_i}{\partial x_j} + v_j\frac{\partial A_j}{\partial x_i}\right)$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m\mathbf{v}}{\sqrt{1-\frac{v^2}{c^2}}}\right) = e\left(\mathbf{E} + \mathbf{v}\times\mathbf{B}\right)$$

For continuous case (current), the force due to the magnetic field is:

$$\mathbf{F} = \int \mathbf{j}\times\mathbf{B}\mathrm{d}^3x = I\int \mathrm{d}\mathbf{l}\times\mathbf{B}$$

5.1.4 Hamiltonian

Expressing $\mathbf{v}$ in terms of $\mathbf{P}$ we get:

$$\mathbf{P} = \frac{m\mathbf{v}}{\sqrt{1-\frac{v^2}{c^2}}} + e\mathbf{A}$$
$$\mathbf{P} - e\mathbf{A} = \frac{m\mathbf{v}}{\sqrt{1-\frac{v^2}{c^2}}}$$
$$P_i - eA_i = \frac{mv_i}{\sqrt{1-\frac{v^2}{c^2}}}$$
$$(P_i - eA_i)^2 \left(1 - \frac{v^2}{c^2}\right) = m^2 v_i^2$$
$$(P_i - eA_i)^2 \left(1 - \frac{(v_1^2+v_2^2+v_3^3)^2}{c^2}\right) = m^2 v_i^2$$
$$v_i^2 = \frac{(P_i - eA_i)^2 c^2}{mc^2 + (\mathbf{P} - e\mathbf{A})^2}$$
$$|v_i| = \frac{|P_i - eA_i|}{\sqrt{m + \frac{1}{c^2}(\mathbf{P} - e\mathbf{A})^2}}$$
$$v_i = \frac{P_i - eA_i}{\sqrt{m + \frac{1}{c^2}(\mathbf{P} - e\mathbf{A})^2}}$$
$$\mathbf{v} = \frac{\mathbf{P} - e\mathbf{A}}{\sqrt{m^2 + \frac{1}{c^2}(\mathbf{P} - e\mathbf{A})^2}}$$
$$\mathbf{v} = \frac{c(\mathbf{P} - e\mathbf{A})}{\sqrt{m^2c^2 + (\mathbf{P} - e\mathbf{A})^2}}$$

The system of equations was solved for v_i using the code (in there $v1s = v_1^2$, $vs = v^2$ and $P1 = P_1 - eA_1$):

```
>>> from sympy import var, solve
>>> var("P1 P2 P3 m c v1s v2s v3s")
(P1, P2, P3, m, c, v1s, v2s, v3s)
>>> vs = v1s+v2s+v3s
>>> solve([P1**2*(1-vs/c**2) -v1s*m**2,
...         P2**2*(1-vs/c**2) -v2s*m**2,
...         P3**2*(1-vs/c**2) -v3s*m**2], [v1s, v2s, v3s])
{v1s: P1**2*c**2/(P1**2 + P2**2 + P3**2 + c**2*m**2),
 v2s: P2**2*c**2/(P1**2 + P2**2 + P3**2 + c**2*m**2),
 v3s: P3**2*c**2/(P1**2 + P2**2 + P3**2 + c**2*m**2)}
```

And the absolute value was removed by using the fact, that v_i has the same sign as $p_i = P_i - eA_i$ which follows from the second equation.

The Hamiltonian is:

$$\begin{aligned}
H(\mathbf{x},\mathbf{P},t) = \mathbf{v}\cdot\mathbf{P} - L = \\
= \mathbf{v}\cdot\mathbf{P} + mc^2\sqrt{1-\frac{v^2}{c^2}} + e\phi - e\mathbf{v}\cdot\mathbf{A} = \\
= \mathbf{v}\cdot(\mathbf{P}-e\mathbf{A}) + mc^2\sqrt{1-\frac{v^2}{c^2}} + e\phi = \\
= \frac{c(\mathbf{P}-e\mathbf{A})\cdot(\mathbf{P}-e\mathbf{A})}{\sqrt{m^2c^2+(\mathbf{P}-e\mathbf{A})^2}} + mc^2\sqrt{1-\frac{1}{c^2}\left(\frac{c(\mathbf{P}-e\mathbf{A})}{\sqrt{m^2c^2+(\mathbf{P}-e\mathbf{A})^2}}\right)^2} + e\phi = \\
= \frac{c(\mathbf{P}-e\mathbf{A})^2}{\sqrt{m^2c^2+(\mathbf{P}-e\mathbf{A})^2}} + mc^2\sqrt{1-\frac{(\mathbf{P}-e\mathbf{A})^2}{m^2c^2+(\mathbf{P}-e\mathbf{A})^2}} + e\phi = \\
= \frac{c(\mathbf{P}-e\mathbf{A})^2}{\sqrt{m^2c^2+(\mathbf{P}-e\mathbf{A})^2}} + mc^2\sqrt{\frac{m^2c^2}{m^2c^2+(\mathbf{P}-e\mathbf{A})^2}} + e\phi = \\
= \frac{c\left((\mathbf{P}-e\mathbf{A})^2+m^2c^2\right)}{\sqrt{m^2c^2+(\mathbf{P}-e\mathbf{A})^2}} + e\phi = \\
= c\sqrt{m^2c^2+(\mathbf{P}-e\mathbf{A})^2} + e\phi
\end{aligned}$$

5.1.5 Electromagnetic Stress Tensor

The stress tensor is calculated from the Lagrangian:

$$\mathcal{L} = -\frac{1}{4\mu_0}F_{\alpha\beta}F^{\alpha\beta} = -\frac{1}{2\mu_0}(\partial_\alpha A_\beta \partial^\alpha A^\beta - \partial_\beta A_\alpha \partial^\alpha A^\beta)$$

using the Noether formula:

$$\begin{aligned}
T^\mu{}_\nu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\alpha)}\partial_\nu A_\alpha - \delta^\mu{}_\nu\mathcal{L} = \\
= -\frac{1}{\mu_0}F^{\mu\alpha}\partial_\nu A_\alpha + \frac{1}{4\mu_0}\delta^\mu{}_\nu F_{\alpha\beta}F^{\alpha\beta}
\end{aligned}$$

We raise the ν index:

$$T^{\mu\nu} = g^{\nu\lambda}T^\mu{}_\lambda = -\frac{1}{\mu_0}F^{\mu\alpha}\partial^\nu A_\alpha + \frac{1}{4\mu_0}g^{\mu\nu}F_{\alpha\beta}F^{\alpha\beta}$$

This tensor is not symmetric under the exchange of the $\mu\nu$ indices. To make it symmetric, we add a total derivative term $\partial_\alpha K^{\alpha\mu\nu}$, where $K^{\alpha\mu\nu}$ is antisymmetric in its first two indices. This guarantees that $\partial_\mu\partial_\alpha K^{\alpha\mu\nu} = 0$ so that the new stress energy tensor is still conserved. We choose $K^{\alpha\mu\nu} = \frac{1}{\mu_0}F^{\mu\alpha}A^\nu$ and get:

$$\begin{aligned}
T^{\mu\nu} + \partial_\alpha K^{\alpha\mu\nu} = -\frac{1}{\mu_0}F^{\mu\alpha}\partial^\nu A_\alpha + \frac{1}{4\mu_0}g^{\mu\nu}F_{\alpha\beta}F^{\alpha\beta} + \frac{1}{\mu_0}\partial_\alpha(F^{\mu\alpha}A^\nu) = \\
= \frac{1}{\mu_0}F^{\mu\alpha}(\partial_\alpha A^\nu - \partial^\nu A_\alpha) + \frac{1}{4\mu_0}g^{\mu\nu}F_{\alpha\beta}F^{\alpha\beta} + \frac{1}{\mu_0}(\partial_\alpha F^{\mu\alpha})A^\nu = \\
= \frac{1}{\mu_0}F^{\mu\alpha}F_\alpha{}^\nu + \frac{1}{4\mu_0}g^{\mu\nu}F_{\alpha\beta}F^{\alpha\beta} = \\
= -\frac{1}{\mu_0}\left(F^{\mu\alpha}F^\nu{}_\alpha - \frac{1}{4}g^{\mu\nu}F_{\alpha\beta}F^{\alpha\beta}\right)
\end{aligned}$$

where we used $\partial_\alpha F^{\mu\alpha} = 0$.

Another way to derive the stress energy tensor is from general relativity using the formula:

$$T_{\mu\nu} = -\frac{2}{\sqrt{|\det g|}}\frac{\delta S_{EM}}{\delta g^{\mu\nu}}$$

So we write the action:

$$S_{EM} = -\int \frac{1}{4\mu_0} F_{\alpha\beta}F^{\alpha\beta}\sqrt{|\det g|}\mathrm{d}^4x = -\int \frac{1}{4\mu_0} g^{\alpha\lambda}g^{\beta\rho}F_{\alpha\beta}F_{\lambda\rho}\sqrt{|\det g|}\mathrm{d}^4x$$

And vary with respect to $g^{\mu\nu}$:

$$\begin{aligned}
\delta S_{EM} &= -\delta\int \frac{1}{4\mu_0} g^{\alpha\lambda}g^{\beta\rho}F_{\alpha\beta}F_{\lambda\rho}\sqrt{|\det g|}\mathrm{d}^4x = \\
&= -\frac{1}{4\mu_0}\int \left(\delta(g^{\alpha\lambda}g^{\beta\rho})F_{\alpha\beta}F_{\lambda\rho}\sqrt{|\det g|} + g^{\alpha\mu}g^{\beta\rho}F_{\alpha\beta}F_{\lambda\rho}\left(\delta\sqrt{|\det g|}\right)\right)\mathrm{d}^4x = \\
&= -\frac{1}{4\mu_0}\int \left(2(\delta g^{\alpha\lambda})g^{\beta\rho}F_{\alpha\beta}F_{\lambda\rho}\sqrt{|\det g|} + g^{\alpha\lambda}g^{\beta\rho}F_{\alpha\beta}F_{\lambda\rho}\left(-\tfrac{1}{2}\sqrt{|\det g|}g_{\mu\nu}(\delta g^{\mu\nu})\right)\right)\mathrm{d}^4x = \\
&= -\frac{1}{4\mu_0}\int \left(2(\delta g^{\alpha\lambda})F_{\alpha\beta}F_{\lambda}{}^{\beta} - \tfrac{1}{2}F_{\alpha\beta}F^{\alpha\beta}g_{\mu\nu}(\delta g^{\mu\nu})\right)\sqrt{|\det g|}\mathrm{d}^4x = \\
&= -\frac{1}{2\mu_0}\int \left(F_{\mu\beta}F_{\nu}{}^{\beta} - \frac{1}{4}F_{\alpha\beta}F^{\alpha\beta}g_{\mu\nu}\right)(\delta g^{\mu\nu})\sqrt{|\det g|}\mathrm{d}^4x
\end{aligned}$$

And we get:

$$T_{\mu\nu} = \frac{1}{\mu_0}\left(F_{\mu\beta}F_{\nu}{}^{\beta} - \frac{1}{4}F_{\alpha\beta}F^{\alpha\beta}g_{\mu\nu}\right)$$

5.1.6 Examples

Coulomb Law

Maxwell's equations in Lorenz gauge (5.1.1.2):

$$\partial_\alpha\partial^\alpha A^\beta = \mu_0 j^\beta$$

have the solution for the scalar potential (5.1.2.1):

$$\phi(\mathbf{x},t) = \frac{1}{4\pi\epsilon_0}\int \frac{\rho(\mathbf{y}, t - \frac{|\mathbf{x}-\mathbf{y}|}{c})}{|\mathbf{x}-\mathbf{y}|}\mathrm{d}^3y$$

Assuming $\frac{|\mathbf{x}-\mathbf{y}|}{c} \ll t$:

$$\phi(\mathbf{x},t) = \frac{1}{4\pi\epsilon_0}\int \frac{\rho(\mathbf{y},t)}{|\mathbf{x}-\mathbf{y}|}\mathrm{d}^3y$$

Assuming the vector potential $\mathbf{A}(\mathbf{x},t) = \mathbf{A}(\mathbf{x})$ is time independent, we get for the electric field:

$$\begin{aligned}
\mathbf{E}(\mathbf{x},t) = -\nabla\phi(\mathbf{x},t) - \frac{\partial\mathbf{A}(\mathbf{x},t)}{\partial t} &= -\nabla\phi(\mathbf{x},t) = -\nabla\frac{1}{4\pi\epsilon_0}\int\frac{\rho(\mathbf{y},t)}{|\mathbf{x}-\mathbf{y}|}\mathrm{d}^3y = \\
&= -\frac{1}{4\pi\epsilon_0}\int\rho(\mathbf{y},t)\nabla\frac{1}{|\mathbf{x}-\mathbf{y}|}\mathrm{d}^3y = \\
&= \frac{1}{4\pi\epsilon_0}\int\rho(\mathbf{y},t)\frac{\mathbf{x}-\mathbf{y}}{|\mathbf{x}-\mathbf{y}|^3}\mathrm{d}^3y
\end{aligned}$$

If the charge distribution can be approximated by an infinitely-narrow wire with linear charge density $\lambda(y,t) = \frac{\mathrm{d}Q(t)}{\mathrm{d}y}$, we get:

$$\rho(\mathbf{y},t)\mathrm{d}^3y = \lambda(\mathbf{y},t)\mathrm{d}l$$

and:

$$\mathbf{E}(\mathbf{x},t) = \frac{1}{4\pi\epsilon_0}\int \lambda(\mathbf{l},t)\frac{\mathbf{x}-\mathbf{l}}{|\mathbf{x}-\mathbf{l}|^3}\mathrm{d}l$$

Example: Straight Wire

Let's assume infinite straight wire with constant linear charge density λ:

$$\begin{aligned}
\mathbf{l} &= (0,0,l)\\
\mathrm{d}\mathbf{l} &= (0,0,1)\mathrm{d}l\\
\mathbf{x} &= (x,y,z)\\
\mathbf{x}-\mathbf{l} &= (x,y,z-l)\\
\mathbf{E}(\mathbf{x}) &= \frac{\lambda}{4\pi\epsilon_0}\int_{-\infty}^{\infty}\frac{\mathbf{x}-\mathbf{l}}{|\mathbf{x}-\mathbf{l}|^3}\mathrm{d}l =\\
&= \frac{\lambda}{4\pi\epsilon_0}\int_{-\infty}^{\infty}\frac{(x,y,z-l)\mathrm{d}l}{(x^2+y^2+(z-l)^2)^{\frac{3}{2}}} =\\
&= \frac{\lambda}{4\pi\epsilon_0}\int_{-\infty}^{\infty}\frac{(x,y,0)\mathrm{d}l}{(x^2+y^2+(z-l)^2)^{\frac{3}{2}}} =\\
&= (x,y,0)\frac{\lambda}{4\pi\epsilon_0}\int_{-\infty}^{\infty}\frac{\mathrm{d}u}{(x^2+y^2+u^2)^{\frac{3}{2}}} =\\
&= (x,y,0)\frac{\lambda}{4\pi\epsilon_0}\frac{2}{x^2+y^2} =\\
&= (x,y,0)\frac{\lambda}{2\pi\epsilon_0}\frac{1}{x^2+y^2}
\end{aligned}$$

For $y=0$:

$$\mathbf{E}(x,0,z) = (x,0,0)\frac{\lambda}{2\pi\epsilon_0}\frac{1}{x^2} = (1,0,0)\frac{\lambda}{2\pi\epsilon_0 x}$$

We can also calculate the scalar potential as follows:

$$\begin{aligned}
\phi(\mathbf{x}) &= \frac{\lambda}{4\pi\epsilon_0}\int_{-\infty}^{\infty}\frac{\mathrm{d}l}{|\mathbf{x}-\mathbf{l}|} =\\
&= \frac{\lambda}{4\pi\epsilon_0}\int_{-\infty}^{\infty}\frac{\mathrm{d}l}{(x^2+y^2+(z-l)^2)^{\frac{1}{2}}} =\\
&= \frac{\lambda}{4\pi\epsilon_0}\int_{-\infty}^{\infty}\frac{\mathrm{d}u}{\sqrt{x^2+y^2+u^2}} =\\
&= \infty
\end{aligned}$$

Note that in the radial direction (let's set for example $y=0$) the result is scale (translation) invariant, i.e. $\phi(kx) = \phi(x)$.

In order to calculate with $\phi(\mathbf{x})$, we need to regularize it first. Cutoff regularization is:

$$\phi(\mathbf{x}) = \frac{\lambda}{4\pi\epsilon_0}\int_{-L}^{L}\frac{\mathrm{d}u}{\sqrt{x^2+y^2+u^2}} = \\ = \frac{\lambda}{4\pi\epsilon_0}\log\frac{\sqrt{x^2+y^2+L^2}+L}{\sqrt{x^2+y^2+L^2}-L}$$

where L is the regulator and also an auxiliary scale. In this regularization, we lost the translational symmetry. The physical quantities don't depend on L in the limit $L \to \infty$:

$$E_x = -\frac{\partial}{\partial x}\phi(x) = \frac{\lambda}{2\pi\epsilon_0 x}\frac{L}{\sqrt{L^2+x^2}} \to \frac{\lambda}{2\pi\epsilon_0 x}$$

and

$$\Delta\phi = \phi(x_2)-\phi(x_1) = \frac{\lambda}{4\pi\epsilon_0}\log\frac{\sqrt{x_2^2+L^2}+L}{\sqrt{x_2^2+L^2}-L}\frac{\sqrt{x_1^2+L^2}-L}{\sqrt{x_1^2+L^2}+L} \to \frac{\lambda}{4\pi\epsilon_0}\log\frac{x_1^2}{x_2^2}$$

Dimensional regularization expresses the integral in the dimension $n = 1-2\epsilon$ as follows:

$$\phi(\mathbf{x}) = \frac{\lambda}{4\pi\epsilon_0}\int \mathrm{d}\Omega_n \int_0^\infty \frac{u^{n-1}}{\Lambda^{n-1}}\frac{\mathrm{d}u}{\sqrt{x^2+y^2+u^2}} = \\ = \frac{\lambda}{4\pi\epsilon_0}\frac{\Gamma\left(\frac{1-n}{2}\right)}{\left(\frac{\sqrt{x^2+y^2}}{\Lambda}\sqrt{\pi}\right)^{1-n}} = \\ = \frac{\lambda}{4\pi\epsilon_0}\frac{\Gamma(\epsilon)}{\left(\frac{\sqrt{x^2+y^2}}{\Lambda}\right)^{2\epsilon}\pi^\epsilon} = \\ = \frac{\lambda}{4\pi\epsilon_0}\left[\frac{1}{\epsilon}-\gamma-\log\pi+\log\frac{\Lambda^2}{x^2+y^2}+O(\epsilon)\right]$$

Here ϵ is the regulator and Λ is the auxiliary scale. This regularization preserves the translational symmetry. Now we can renormalize the integral. The minimal subtraction (MS) renormalization is:

$$\phi_{\mathrm{MS}}(\mathbf{x}) = \frac{\lambda}{4\pi\epsilon_0}\left[-\gamma-\log\pi+\log\frac{\Lambda^2}{x^2+y^2}\right]$$

Another option is the modified minimal subtraction ($\overline{\mathrm{MS}}$) renormalization is:

$$\phi_{\overline{\mathrm{MS}}}(\mathbf{x}) = \frac{\lambda}{4\pi\epsilon_0}\log\frac{\Lambda^2}{x^2+y^2}$$

Once we choose a renormalization scheme, we can calculate the electric field as follows:

$$E_x = -\frac{\partial}{\partial x}\phi_{\overline{\mathrm{MS}}}(x) = -\frac{\partial}{\partial x}\frac{\lambda}{4\pi\epsilon_0}\log\frac{\Lambda^2}{x^2} = \\ = -\frac{\lambda}{4\pi\epsilon_0}\frac{x^2}{\Lambda^2}\Lambda^2\left(-\frac{2}{x^3}\right) = \\ = \frac{\lambda}{2\pi\epsilon_0}\frac{1}{x}$$

and the potential difference as:

$$\Delta\phi = \phi_{\overline{\mathrm{MS}}}(x_2)-\phi_{\overline{\mathrm{MS}}}(x_1) = \frac{\lambda}{4\pi\epsilon_0}\log\frac{\Lambda^2}{x_2^2}\frac{x_1^2}{\Lambda^2} = \frac{\lambda}{4\pi\epsilon_0}\log\frac{x_1^2}{x_2^2}$$

In agreement with the previous result. The final results don't depend on the auxiliary scale Λ and we are not doing any limits.

Biot-Savart Law

Maxwell's equations in Lorenz gauge (5.1.1.2):

$$\partial_\alpha \partial^\alpha A^\beta = \mu_0 j^\beta$$

have the solution for the vector potential (5.1.2.2):

$$\mathbf{A}(\mathbf{x}, t) = \frac{\mu_0}{4\pi} \int \frac{\mathbf{j}(\mathbf{y}, t - \frac{|\mathbf{x}-\mathbf{y}|}{c})}{|\mathbf{x}-\mathbf{y}|} \mathrm{d}^3 y$$

Assuming $\frac{|\mathbf{x}-\mathbf{y}|}{c} \ll t$:

$$\mathbf{A}(\mathbf{x}, t) = \frac{\mu_0}{4\pi} \int \frac{\mathbf{j}(\mathbf{y}, t)}{|\mathbf{x}-\mathbf{y}|} \mathrm{d}^3 y$$

The magnetic field is then:

$$\begin{aligned}
\mathbf{B}(\mathbf{x}, t) = \nabla \times \mathbf{A}(\mathbf{x}, t) = \nabla \times \frac{\mu_0}{4\pi} \int \frac{\mathbf{j}(\mathbf{y}, t)}{|\mathbf{x}-\mathbf{y}|} \mathrm{d}^3 y = \\
= \frac{\mu_0}{4\pi} \int \left(\nabla \frac{1}{|\mathbf{x}-\mathbf{y}|} \right) \times \mathbf{j}(\mathbf{y}, t) \mathrm{d}^3 y = \\
= \frac{\mu_0}{4\pi} \int \left(-\frac{\mathbf{x}-\mathbf{y}}{|\mathbf{x}-\mathbf{y}|^3} \right) \times \mathbf{j}(\mathbf{y}, t) \mathrm{d}^3 y = \\
= \frac{\mu_0}{4\pi} \int \mathbf{j}(\mathbf{y}, t) \times \frac{\mathbf{x}-\mathbf{y}}{|\mathbf{x}-\mathbf{y}|^3} \mathrm{d}^3 y
\end{aligned}$$

If the current can be approximated by an infinitely-narrow wire, we get:

$$\mathbf{j}(\mathbf{y}, t) \mathrm{d}^3 y = I(t) \mathrm{d}\mathbf{l}$$

and:

$$\mathbf{B}(\mathbf{x}, t) = \frac{\mu_0}{4\pi} \int I(t) \mathrm{d}\mathbf{l} \times \frac{\mathbf{x}-\mathbf{l}}{|\mathbf{x}-\mathbf{l}|^3}$$

Example: Straight Wire

Let's assume infinite straight wire carrying constant current I:

$$\begin{aligned}
\mathbf{l} = (0, 0, l) \\
\mathrm{d}\mathbf{l} = (0, 0, 1) \mathrm{d}l \\
\mathbf{x} = (x, y, z) \\
\mathbf{x} - \mathbf{l} = (x, y, z - l) \\
\mathbf{B}(\mathbf{x}) = \frac{\mu_0 I}{4\pi} \int \mathrm{d}\mathbf{l} \times \frac{\mathbf{x}-\mathbf{l}}{|\mathbf{x}-\mathbf{l}|^3} = \\
= \frac{\mu_0 I}{4\pi} \int_{-\infty}^{\infty} (0, 0, 1) \times \frac{(x, y, z-l) \mathrm{d}l}{(x^2 + y^2 + (z-l)^2)^{\frac{3}{2}}} = \\
= (y, -x, 0) \frac{\mu_0 I}{4\pi} \int_{-\infty}^{\infty} \frac{\mathrm{d}l}{(x^2 + y^2 + (z-l)^2)^{\frac{3}{2}}} = \\
= (y, -x, 0) \frac{\mu_0 I}{4\pi} \frac{2}{x^2 + y^2} = \\
= (y, -x, 0) \frac{\mu_0 I}{2\pi} \frac{1}{x^2 + y^2}
\end{aligned}$$

Where we used the value of the folowing integral:

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}l}{(x^2+y^2+(z-l)^2)^{\frac{3}{2}}} = \int_{-\infty}^{\infty} \frac{\mathrm{d}u}{(x^2+y^2+u^2)^{\frac{3}{2}}} =$$
$$= \left[\frac{u}{(x^2+y^2)\sqrt{x^2+y^2+u^2}}\right]_{-\infty}^{\infty} = \left[\frac{\operatorname{sign} u}{(x^2+y^2)\sqrt{\left(\frac{x}{u}\right)^2+\left(\frac{y}{u}\right)^2+1}}\right]_{-\infty}^{\infty} =$$
$$= \frac{1}{x^2+y^2} - \left(-\frac{1}{x^2+y^2}\right) = \frac{2}{x^2+y^2}$$

For $y = 0$:

$$\mathbf{B}(x,0,z) = (0,-x,0)\frac{\mu_0 I}{2\pi}\frac{1}{x^2} = (0,-1,0)\frac{\mu_0 I}{2\pi x}$$

Example: Circular Loop

Let's assume a circular loop:

$$\mathbf{l} = (r\cos\phi, r\sin\phi, 0)$$
$$\frac{\mathrm{d}\mathbf{l}}{\mathrm{d}\phi} = (-r\sin\phi, r\cos\phi, 0)$$
$$\mathbf{x} = (x,y,z)$$
$$\mathbf{x}-\mathbf{l} = (x-r\cos\phi, y-r\sin\phi, z)$$
$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0 I}{4\pi}\int \mathrm{d}\mathbf{l}\times\frac{\mathbf{x}-\mathbf{l}}{|\mathbf{x}-\mathbf{l}|^3} =$$
$$= \frac{\mu_0 I}{4\pi}\int_0^{2\pi} (-r\sin\phi, r\cos\phi, 0)\times\frac{(x-r\cos\phi, y-r\sin\phi, z)\mathrm{d}\phi}{((x-r\cos\phi)^2+(y-r\sin\phi)^2+z^2)^{\frac{3}{2}}} =$$
$$= \frac{\mu_0 I}{4\pi}\int_0^{2\pi} \frac{(-z\cos\phi, -z\sin\phi, (x-r\cos\phi)\cos\phi+(y-r\sin\phi)\sin\phi)r\mathrm{d}\phi}{((x-r\cos\phi)^2+(y-r\sin\phi)^2+z^2)^{\frac{3}{2}}} =$$
$$= \frac{\mu_0 I}{4\pi}\int_0^{2\pi} \frac{(-z\cos\phi, -z\sin\phi, x\cos\phi+y\sin\phi-r)r\mathrm{d}\phi}{(x^2+y^2+z^2+r^2-2xr\cos\phi-2yr\sin\phi)^{\frac{3}{2}}}$$

Due to the symmetry of the problem, we can set $y = 0$:

$$\mathbf{B}(x,0,z) = \frac{\mu_0 I}{4\pi}\int_0^{2\pi} \frac{(-z\cos\phi, -z\sin\phi, x\cos\phi-r)r\mathrm{d}\phi}{(x^2+z^2+r^2-2xr\cos\phi)^{\frac{3}{2}}} =$$
$$= \frac{\mu_0 I}{4\pi}\int_0^{2\pi} \frac{(-z\cos\phi, 0, x\cos\phi-r)r\mathrm{d}\phi}{(x^2+z^2+r^2-2xr\cos\phi)^{\frac{3}{2}}}$$

In the last equation we used the fact, that $\sin\phi$ is odd and $\cos\phi$ is even on the interval $(0, 2\pi)$. For $x = y = 0$ we get:

$$\mathbf{B}(0,0,z) = \frac{\mu_0 I}{4\pi}\int_0^{2\pi} \frac{(-z\cos\phi, 0, -r)r\mathrm{d}\phi}{(r^2+z^2)^{\frac{3}{2}}} =$$
$$= (0,0,-1)\frac{\mu_0 I}{4\pi}\int_0^{2\pi} \frac{r^2\mathrm{d}\phi}{(r^2+z^2)^{\frac{3}{2}}} =$$
$$= (0,0,-1)\frac{\mu_0 I}{2}\frac{r^2}{(r^2+z^2)^{\frac{3}{2}}}$$

Helmholtz Coil

Helmholtz coil is a set of two circular loops of radius r, that are d apart, where $d = r$. Let's calculate the magnetic field on the axis. Magnetic field of the first coil is (see the previous example):

$$\mathbf{B}_1(0,0,z) = (0,0,-1)\frac{\mu_0 I}{2}\frac{r^2}{(r^2+z^2)^{\frac{3}{2}}}$$

Second coil is positioned d above the first one:

$$\mathbf{B}_2(0,0,z) = (0,0,-1)\frac{\mu_0 I}{2}\frac{r^2}{(r^2+(z-d)^2)^{\frac{3}{2}}}$$

The total magnetic field is:

$$\begin{aligned}\mathbf{B}(0,0,z) &= \mathbf{B}_1(0,0,z) + \mathbf{B}_2(0,0,z) = \\ &= (0,0,-1)\frac{\mu_0 I}{2}\frac{r^2}{(r^2+z^2)^{\frac{3}{2}}} + (0,0,-1)\frac{\mu_0 I}{2}\frac{r^2}{(r^2+(z-d)^2)^{\frac{3}{2}}} = \\ &= (0,0,-1)\frac{\mu_0 I r^2}{2}\left(\frac{1}{(r^2+z^2)^{\frac{3}{2}}} + \frac{1}{(r^2+(z-d)^2)^{\frac{3}{2}}}\right)\end{aligned}$$

The field in the middle:

$$\begin{aligned}\mathbf{B}(0,0,\frac{d}{2}) &= (0,0,-1)\frac{\mu_0 I r^2}{2}\left(\frac{1}{(r^2+\left(\frac{d}{2}\right)^2)^{\frac{3}{2}}} + \frac{1}{(r^2+\left(\frac{d}{2}\right)^2)^{\frac{3}{2}}}\right) = \\ &= (0,0,-1)\frac{\mu_0 I r^2}{(r^2+\left(\frac{d}{2}\right)^2)^{\frac{3}{2}}}\end{aligned}$$

For $r = d$ we get:

$$\begin{aligned}\mathbf{B}(0,0,\frac{d}{2}) &= (0,0,-1)\frac{\mu_0 I r^2}{(r^2+\left(\frac{r}{2}\right)^2)^{\frac{3}{2}}} = \\ &= (0,0,-1)\frac{\mu_0 I}{r(1+\left(\frac{1}{2}\right)^2)^{\frac{3}{2}}} = \\ &= (0,0,-1)\frac{\mu_0 I 4^{\frac{3}{2}}}{r5^{\frac{3}{2}}} = \\ &= (0,0,-1)\frac{8}{5\sqrt{5}}\frac{\mu_0 I}{r} = \\ &= (0,0,-1)B\end{aligned}$$

where the magnitude of $\mathbf{B}$ is:

$$B = \frac{8}{5\sqrt{5}}\frac{\mu_0 I}{r}$$

For $r = 0.15\,\mathrm{m}$ and $N = 130$ turns we get the magnitude of the field as (we use SI units, so I is in A and B in tesla):

$$B = \frac{8}{5\sqrt{5}}\frac{\mu_0 N I}{r} = \frac{8}{5\sqrt{5}}\frac{4\pi 10^{-7}\cdot 130 I}{0.15} = 7.79\cdot 10^{-4} I$$

Code:

```
>>> from math import pi, sqrt
>>> "%e" % (8*4*pi*1e-7*130 / (5*sqrt(5)*0.15))
'7.792861e-04'
```

Equation of motion for an electron in this field is:

$$m\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}t^2} = e\left(\mathbf{v}\times\mathbf{B}\right)$$
$$m\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}t^2} = eB\left(v_y, -v_x, 0\right)$$

The general solution is:

$$\mathbf{x} = \frac{vm}{eB}\left(x + \cos\frac{eB}{m}(t-t_0), y - \sin\frac{eB}{m}(t-t_0), z\right)$$

So the electron is moving in a circle with a center (x, y, z), t_0 depends on the initial direction of the velocity and v is the magnitude of the initial velocity. There can also be a possible movement in the z direction, but for the following initial conditions there is none:

$$\mathbf{x}_0 = (0, 0, 0)$$
$$\mathbf{v}_0 = (0, -v, 0)$$

Then we get:

$$\mathbf{x} = \frac{vm}{eB}\left(-1 + \cos\frac{eB}{m}t, -\sin\frac{eB}{m}t, 0\right)$$
$$\mathbf{v} = v\left(-\sin\frac{eB}{m}t, -\cos\frac{eB}{m}t, 0\right)$$

So the radius of the circle is $R = \frac{vm}{eB}$. Let the electrons by accelerated by the electric potential V:

$$\tfrac{1}{2}mv^2 = eV$$

So the initial velocity is:

$$v = \sqrt{\frac{2eV}{m}}$$

and we get for the radius:

$$R = \frac{vm}{eB} = \frac{m}{eB}\sqrt{\frac{2eV}{m}} = \frac{1}{B}\sqrt{\frac{2mV}{e}}$$

from which the electron charge versus mass ratio is:

$$\frac{e}{m} = \frac{2V}{R^2B^2} = \frac{2V}{R^2\left(\frac{8}{5\sqrt{5}}\frac{\mu_0 NI}{r}\right)^2} = \\ = \frac{125Vr^2}{32\mu_0^2R^2N^2I^2}$$

For $r = 0.15\,\mathrm{m}$, $N = 130$, $V = 300\,\mathrm{V}$, $R = 0.05\,\mathrm{m}$, $I = 1.48\,\mathrm{A}$ we get:

$$\frac{e}{m} = 1.80\cdot 10^{11}\,\mathrm{C}\cdot\mathrm{kg}^{-1}$$

Code:

```
>>> from math import pi
>>> r = 0.15
>>> N = 130
>>> V = 300
>>> R = 0.05
>>> I = 1.48
>>> mu0 = 4*pi*1e-7
>>> "%e" % (125 * V * r**2 / (32 * mu0**2 * R**2 * N**2 * I**2))
'1.804238e+11'
```

Reference value is:

$$\frac{e}{m} = 1.7588 \cdot 10^{11}\,\mathrm{C} \cdot \mathrm{kg}^{-1}$$

Code:

```
>>> e = 1.6021766e-19
>>> c = 299792458
>>> eV = e
>>> KeV = 1e3 * eV
>>> m = 510.998910 * KeV / c**2
>>> m
9.109382795192204e-31
>>> "%e" % (e / m)
'1.758820e+11'
```

or even simpler (we do not actually need the value of the electron charge e):

```
>>> c = 299792458
>>> KeV = 1e3
>>> m = 510.998910 * KeV / c**2
>>> "%e" % (1/m)
'1.758820e+11'
```

We can use the experimental value to calculate the electron rest mass energy:

$$mc^2 = \frac{c^2}{1.804238 \cdot 10^{11}}\,\mathrm{eV} = 498.1356\,\mathrm{KeV}$$

Ampère's Force Law

The force on a wire 1 due to a magnetic field of a wire 2 is:

$$\mathbf{F} = I_1 \int \mathrm{d}\mathbf{l}_1 \times \mathbf{B}(\mathbf{l}_1)$$

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} \int I_2(t)\mathrm{d}\mathbf{l}_2 \times \frac{\mathbf{x} - \mathbf{l}_2}{|\mathbf{x} - \mathbf{l}_2|^3}$$

Where $\mathbf{B}(\mathbf{x})$ is the magnetic field produced by the wire 2. Combining these two equations we get:

$$\begin{aligned}\mathbf{F} = I_1 \int \mathrm{d}\mathbf{l}_1 \times \mathbf{B}(\mathbf{l}_1) = \\ = I_1 \int \mathrm{d}\mathbf{l}_1 \times \left(\frac{\mu_0}{4\pi} \int I_2(t) \mathrm{d}\mathbf{l}_2 \times \frac{\mathbf{l}_1 - \mathbf{l}_2}{|\mathbf{l}_1 - \mathbf{l}_2|^3} \right) = \\ = \frac{\mu_o I_1 I_2}{4\pi} \int \int \frac{\mathrm{d}\mathbf{l}_1 \times (\mathrm{d}\mathbf{l}_2 \times (\mathbf{l}_1 - \mathbf{l}_2))}{|\mathbf{l}_1 - \mathbf{l}_2|^3} = \\ = \frac{\mu_o I_1 I_2}{4\pi} \int \int \frac{\mathrm{d}\mathbf{l}_2 (\mathrm{d}\mathbf{l}_1 \cdot (\mathbf{l}_1 - \mathbf{l}_2)) - (\mathbf{l}_1 - \mathbf{l}_2)(\mathrm{d}\mathbf{l}_2 \cdot \mathrm{d}\mathbf{l}_1)}{|\mathbf{l}_1 - \mathbf{l}_2|^3}\end{aligned}$$

Parallel Straight Wires

We calculate the force between two parallel straight infinite wires:

$$\begin{aligned}\mathbf{l}_1 = (\frac{d}{2}, 0, l_1) \\ \mathrm{d}\mathbf{l}_1 = (0, 0, \mathrm{d}l_1) \\ \mathbf{l}_2 = (-\frac{d}{2}, 0, l_2) \\ \mathrm{d}\mathbf{l}_2 = (0, 0, \mathrm{d}l_2) \\ \mathbf{l}_1 - \mathbf{l}_2 = (d, 0, l_1 - l_2) \\ \mathrm{d}\mathbf{l}_2 (\mathrm{d}\mathbf{l}_1 \cdot (\mathbf{l}_1 - \mathbf{l}_2)) - (\mathbf{l}_1 - \mathbf{l}_2)(\mathrm{d}\mathbf{l}_2 \cdot \mathrm{d}\mathbf{l}_1) = (0, 0, \mathrm{d}l_2)(l_1 - l_2)\mathrm{d}l_1 - (d, 0, l_1 - l_2)\mathrm{d}l_2 \mathrm{d}l_1 = (-d, 0, 0)\mathrm{d}l_1 \mathrm{d}l_2 \\ \mathbf{F} = \frac{\mu_o I_1 I_2}{4\pi} \int \int \frac{\mathrm{d}\mathbf{l}_2 (\mathrm{d}\mathbf{l}_1 \cdot (\mathbf{l}_1 - \mathbf{l}_2)) - (\mathbf{l}_1 - \mathbf{l}_2)(\mathrm{d}\mathbf{l}_2 \cdot \mathrm{d}\mathbf{l}_1)}{|\mathbf{l}_1 - \mathbf{l}_2|^3} = \\ = \frac{\mu_o I_1 I_2}{4\pi} \int \int \frac{(-d, 0, 0)\mathrm{d}l_1 \mathrm{d}l_2}{(d^2 + (l_1 - l_2)^2)^{\frac{3}{2}}} = \\ = (-1, 0, 0) \frac{\mu_o I_1 I_2}{4\pi} \int \mathrm{d}l_1 \int_{-\infty}^{\infty} \mathrm{d}l_2 \frac{d}{(d^2 + (l_1 - l_2)^2)^{\frac{3}{2}}} = \\ = (-1, 0, 0) \frac{\mu_o I_1 I_2}{4\pi} \int \mathrm{d}l_1 \frac{2}{d} = \\ = (-1, 0, 0) \frac{\mu_o I_1 I_2}{2\pi d} \int \mathrm{d}l_1\end{aligned}$$

Where we used the value of the folowing integral:

$$\begin{aligned}\int_{-\infty}^{\infty} \mathrm{d}l_2 \frac{d}{(d^2 + (l_1 - l_2)^2)^{\frac{3}{2}}} = \int_{-\infty}^{\infty} \mathrm{d}x \frac{d}{(d^2 + x^2)^{\frac{3}{2}}} = \\ = \left[\frac{x}{d\sqrt{d^2 + x^2}} \right]_{-\infty}^{\infty} = \left[\frac{\operatorname{sign} x}{d\sqrt{\left(\frac{d}{x}\right)^2 + 1}} \right]_{-\infty}^{\infty} = \\ = \frac{1}{d} - \left(-\frac{1}{d} \right) = \frac{2}{d}\end{aligned}$$

As such, the direction of the force on the first wire (at coordinates $(\frac{d}{2}, 0, 0)$ going in the z direction) will be to the left and the force per unit length is:

$$F_m = \frac{\mu_o I_1 I_2}{2\pi d}$$

Because the second wire is at the coordinates $(-\frac{d}{2}, 0, 0)$ and the force on the first wire is in the direction $(-1, 0, 0)$, the force between the wires is attractive, as long as I_1 and I_2 have the same sign (either both currents go up, or both down) and repulsive if I_1 and I_2 have opposite signs.

Let $d = 1\,\mathrm{m}$, $I_1 = I_2 = 1\,\mathrm{A}$, then the force is attractive and (we also use $\mu_0 = 4\pi \cdot 10^{-7}$):

$$F_m = \frac{4\pi \cdot 10^{-7}}{2\pi}\,\mathrm{N}\cdot\mathrm{m}^{-1} = 2 \cdot 10^{-7}\,\mathrm{N}\cdot\mathrm{m}^{-1}$$

Perpendicular Straight Wires

We calculate the force between two perpendicular straight infinite wires:

$$\begin{aligned}
\mathbf{l}_1 &= (\frac{d}{2}, 0, l_1) \\
\mathrm{d}\mathbf{l}_1 &= (0, 0, \mathrm{d}l_1) \\
\mathbf{l}_2 &= (-\frac{d}{2}, l_2, 0) \\
\mathrm{d}\mathbf{l}_2 &= (0, \mathrm{d}l_2, 0) \\
\mathbf{l}_1 - \mathbf{l}_2 &= (d, -l_2, l_1) \\
\mathrm{d}\mathbf{l}_2(\mathrm{d}\mathbf{l}_1 \cdot (\mathbf{l}_1 - \mathbf{l}_2)) - (\mathbf{l}_1 - \mathbf{l}_2)(\mathrm{d}\mathbf{l}_2 \cdot \mathrm{d}\mathbf{l}_1) &= (0, \mathrm{d}l_2, 0) l_1 \mathrm{d}l_1 = (0, l_1, 0)\mathrm{d}l_1\mathrm{d}l_2 \\
\mathbf{F} = \frac{\mu_o I_1 I_2}{4\pi} \int\int \frac{\mathrm{d}\mathbf{l}_2(\mathrm{d}\mathbf{l}_1 \cdot (\mathbf{l}_1 - \mathbf{l}_2)) - (\mathbf{l}_1 - \mathbf{l}_2)(\mathrm{d}\mathbf{l}_2 \cdot \mathrm{d}\mathbf{l}_1)}{|\mathbf{l}_1 - \mathbf{l}_2|^3} &= \\
= \frac{\mu_o I_1 I_2}{4\pi} \int\int \frac{(0, l_1, 0)\mathrm{d}l_1\mathrm{d}l_2}{(d^2 + l_1^2 + l_2^2)^{\frac{3}{2}}} &= \\
= (0,1,0)\frac{\mu_o I_1 I_2}{4\pi} \int_{-\infty}^{\infty} \mathrm{d}l_1 \int_{-\infty}^{\infty} \mathrm{d}l_2 \frac{l_1}{(d^2 + l_1^2 + l_2^2)^{\frac{3}{2}}} &= \\
= (-1,0,0)\frac{\mu_o I_1 I_2}{4\pi} \int_{-\infty}^{\infty} \mathrm{d}l_1 \frac{2l_1}{d^2 + l_1^2} &= \\
= 0 &
\end{aligned}$$

The integral is an odd functin of l_1, so it is zero. We used the value of the folowing integral (but in fact it is already seen before this integral is needed that the double integral must be zero):

$$\begin{aligned}
\int_{-\infty}^{\infty} \mathrm{d}l_2 \frac{l_1}{(d^2 + l_1^2 + l_2^2)^{\frac{3}{2}}} & \\
= \left[\frac{l_1 l_2}{(d^2 + l_1^2)\sqrt{d^2 + l_1^2 + l_2^2}}\right]_{-\infty}^{\infty} = \left[\frac{l_1 \operatorname{sign} l_2}{(d^2 + l_1^2)\sqrt{\left(\frac{d}{l_2}\right)^2 + \left(\frac{l_1}{l_2}\right)^2 + 1}}\right]_{-\infty}^{\infty} = & \\
= \frac{l_1}{d^2 + l_1^2} - \left(-\frac{l_1}{d^2 + l_1^2}\right) = \frac{2l_1}{d^2 + l_1^2} &
\end{aligned}$$

As such, there will be no net force.

Infinitely Long Wire and a Square Loop

We calculate the net force on a square loop with current I_1 of side a, whose center is d far from an infinitely long wire with current I_2:

The wire has coordinates $(0, 0, z)$ and the magnetic field from it is (see the example above):

$$\mathbf{B}(x, 0, z) = (0, -1, 0)\frac{\mu_0 I}{2\pi x}$$

The four sides of the loop are ($0 \leq l_1 \leq a$):

$$\begin{aligned}
\mathbf{l}_1 &= (d - \frac{a}{2} + l_1, 0, \frac{a}{2}) \\
\mathbf{l}_1 &= (d + \frac{a}{2}, 0, \frac{a}{2} - l_1) \\
\mathbf{l}_1 &= (d + \frac{a}{2} - l_1, 0, -\frac{a}{2}) \\
\mathbf{l}_1 &= (d - \frac{a}{2}, 0, -\frac{a}{2} + l_1)
\end{aligned}$$

and the differentials are:

$$\begin{aligned}
\mathrm{d}\mathbf{l}_1 &= (1, 0, 0)\mathrm{d}l_1 \\
\mathrm{d}\mathbf{l}_1 &= (0, 0, -1)\mathrm{d}l_1 \\
\mathrm{d}\mathbf{l}_1 &= (-1, 0, 0)\mathrm{d}l_1 \\
\mathrm{d}\mathbf{l}_1 &= (0, 0, 1)\mathrm{d}l_1
\end{aligned}$$

The net force on the loop is:

$$\begin{aligned}
\mathbf{F} = I_1 \int \mathrm{d}\mathbf{l}_1 \times \mathbf{B} = I_1 \int \mathrm{d}\mathbf{l}_1 \times (0, -1, 0) \frac{\mu_0 I_2}{2\pi (\mathbf{l}_1)_x} = \\
= \frac{\mu_0 I_1 I_2}{2\pi} \left(\int_0^a \frac{(0, 0, 1)\mathrm{d}l_1}{d - \frac{a}{2} + l_1} + \int_0^a \frac{(1, 0, 0)\mathrm{d}l_1}{d + \frac{a}{2}} + \int_0^a \frac{(0, 0, -1)\mathrm{d}l_1}{d + \frac{a}{2} - l_1} + \int_0^a \frac{(-1, 0, 0)\mathrm{d}l_1}{d - \frac{a}{2}} \right) = \\
= \frac{\mu_0 I_1 I_2}{2\pi} \left((0, 0, 1) \left[\log \left| d - \frac{a}{2} + l_1 \right| - \log \left| d + \frac{a}{2} - l_1 \right| \right]_0^a + (1, 0, 0) \left(\frac{a}{d + \frac{a}{2}} - \frac{a}{d - \frac{a}{2}} \right) \right) = \\
= \frac{\mu_0 I_1 I_2}{2\pi} \left((0, 0, 1) \cdot 0 + (1, 0, 0) \frac{a^2}{d^2 - \left(\frac{a}{2}\right)^2} \right) = \\
= (1, 0, 0) \frac{\mu_0 I_1 I_2}{2\pi} \frac{a^2}{d^2 - \left(\frac{a}{2}\right)^2}
\end{aligned}$$

Magnetic Dipole

$$\begin{aligned}
\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi} \frac{\mathbf{m} \times \mathbf{r}}{r^3} \\
\mathbf{B}(\mathbf{r}) = \nabla \times \mathbf{A} = \frac{\mu_0}{4\pi} \nabla \times \left(\frac{\mathbf{m} \times \mathbf{r}}{r^3} \right) = \\
= \frac{\mu_0}{4\pi} \left(\mathbf{m} \nabla \cdot \left(\frac{\mathbf{r}}{r^3} \right) - \mathbf{m} \cdot \nabla \left(\frac{\mathbf{r}}{r^3} \right) \right) = \\
= \frac{\mu_0}{4\pi} \left(\mathbf{m} \left(\left(\nabla \frac{1}{r^3} \right) \cdot \mathbf{r} + \frac{1}{r^3} \nabla \cdot \mathbf{r} \right) - \mathbf{m} \cdot \left(\left(\nabla \frac{1}{r^3} \right) \mathbf{r} + \frac{1}{r^3} \nabla \mathbf{r} \right) \right) = \\
= \frac{\mu_0}{4\pi} \left(\mathbf{m} \left(\left(-\frac{3\mathbf{r}}{r^5} \right) \cdot \mathbf{r} + \frac{1}{r^3} 3 \right) - \mathbf{m} \cdot \left(\left(-\frac{3\mathbf{r}}{r^5} \right) \mathbf{r} + \frac{1}{r^3} \mathbb{1} \right) \right) = \\
= \frac{\mu_0}{4\pi} \left(\mathbf{m} \left(-\frac{3}{r^3} + \frac{3}{r^3} \right) + \mathbf{m} \cdot \left(\frac{3\mathbf{r}\mathbf{r}}{r^5} - \frac{\mathbb{1}}{r^3} \right) \right) = \\
= \frac{\mu_0}{4\pi} \left(\frac{3\mathbf{r}(\mathbf{m} \cdot \mathbf{r})}{r^5} - \frac{\mathbf{m}}{r^3} \right)
\end{aligned}$$

Bar Magnet

A good model of a bar magnet of the length L and width W is a combination of two magnetic monopoles (that sit inside the magnet, so one cannot actually see them, just their behavior outside the magnet):

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0 Q_m}{4\pi}\left(\frac{\mathbf{x}-\mathbf{p}_1}{|\mathbf{x}-\mathbf{p}_1|^3} - \frac{\mathbf{x}-\mathbf{p}_2}{|\mathbf{x}-\mathbf{p}_2|^3}\right)$$

where:

$$\begin{aligned}\mathbf{p}_1 &= (0,0,d)\\ \mathbf{p}_2 &= (0,0,-d)\\ d &= \frac{L-W}{2}\end{aligned}$$

The magnetic moment vector is:

$$\mathbf{m} = Q_m(\mathbf{p}_1 - \mathbf{p}_2)$$

and its magnitude then is:

$$m = 2Q_m d$$

The permeability is:

$$\mu_0 = 4\pi\cdot 10^{-7}\,\mathrm{H}\cdot\mathrm{m}^{-1} = 4\pi\cdot 10^{-7}\,\mathrm{V}\cdot\mathrm{s}\cdot\mathrm{A}^{-1}\cdot\mathrm{m}^{-1}$$

For a typical bar magnet, we have for example:

$$\begin{aligned}L &= 5\,\mathrm{cm}\\ W &= 1\,\mathrm{cm}\\ Q_m &= 3.3\,\mathrm{A}\cdot\mathrm{m}\\ d &= \frac{L-W}{2} = 0.02\,\mathrm{m}\\ m &= 2Q_m d = 2\times 3.3\times 0.02\,\mathrm{A}\cdot\mathrm{m}^2 = 0.132\,\mathrm{A}\cdot\mathrm{m}^2\end{aligned}$$

The unit of $\mathbf{B}$ is Tesla: $1\mathrm{T} = \mathrm{V}\cdot\mathrm{s}\cdot\mathrm{m}^{-2}$.

Bar Magnet in a Coil

We throw a magnet through a coil and calculate the voltage on the coil. We use two model of the bar magnet: a magnetic dipole and two monopoles $2d$ apart.

Geometry:

$$\begin{aligned}\mathbf{v} &= (0,0,v)\\ \mathbf{l} &= (a\cos\phi, a\sin\phi, z)\\ \frac{\mathrm{d}\mathbf{l}}{\mathrm{d}\phi} &= (-a\sin\phi, a\cos\phi, 0)\end{aligned}$$

Field of the dipole:

$$\begin{aligned}\mathbf{E} &= 0\\ \mathbf{B}(\mathbf{r}) &= \frac{\mu_0}{4\pi}\left(\frac{3\mathbf{r}(\mathbf{m}\cdot\mathbf{r})}{r^5} - \frac{\mathbf{m}}{r^3}\right)\\ \mathbf{m} &= (0,0,m)\end{aligned}$$

we will need:

$$\mathbf{v} \times \mathbf{B}(\mathbf{l}) = \frac{\mu_0}{4\pi}\mathbf{v} \times \left(\frac{3\mathbf{l}(\mathbf{m}\cdot\mathbf{l})}{l^5} - \frac{\mathbf{m}}{l^3}\right) =$$
$$= \frac{\mu_0}{4\pi}\left(\frac{3(\mathbf{v}\times\mathbf{l})(\mathbf{m}\cdot\mathbf{l})}{l^5} - \frac{\mathbf{v}\times\mathbf{m}}{l^3}\right) =$$
$$= \frac{\mu_0}{4\pi}\frac{3(\mathbf{v}\times\mathbf{l})(\mathbf{m}\cdot\mathbf{l})}{l^5} =$$
$$= \frac{\mu_0}{4\pi}\frac{3(va\sin\theta, -va\cos\theta, 0)mz}{(a^2+z^2)^{\frac{5}{2}}} =$$
$$= \frac{3\mu_0 m}{4\pi}\frac{avz}{(a^2+z^2)^{\frac{5}{2}}}(\sin\theta, -\cos\theta, 0)$$

and

$$\mathbf{v}\times\mathbf{B}\cdot\frac{\mathrm{d}\mathbf{l}}{\mathrm{d}\phi} =$$
$$= \frac{3\mu_0 m}{4\pi}\frac{avz}{(a^2+z^2)^{\frac{5}{2}}}(\sin\theta, -\cos\theta, 0)\cdot(-a\sin\phi, a\cos\phi, 0) =$$
$$= -\frac{3\mu_0 m}{4\pi}\frac{a^2 vz}{(a^2+z^2)^{\frac{5}{2}}}$$

Field of two monopoles:

$$\mathbf{E} = 0$$
$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0 Q_m}{4\pi}\left(\frac{\mathbf{x}-\mathbf{p}_1}{|\mathbf{x}-\mathbf{p}_1|^3} - \frac{\mathbf{x}-\mathbf{p}_2}{|\mathbf{x}-\mathbf{p}_2|^3}\right)$$
$$\mathbf{p}_1 = (0,0,d)$$
$$\mathbf{p}_2 = (0,0,-d)$$
$$d = \frac{L-W}{2}$$

we will need:

$$\mathbf{v}\times\mathbf{B}(\mathbf{l}) = \frac{\mu_0 Q_m}{4\pi}\left(\frac{\mathbf{v}\times(\mathbf{l}-\mathbf{p}_1)}{|\mathbf{l}-\mathbf{p}_1|^3} - \frac{\mathbf{v}\times(\mathbf{l}-\mathbf{p}_2)}{|\mathbf{l}-\mathbf{p}_2|^3}\right) =$$
$$= \frac{\mu_0 Q_m}{4\pi}\left(\frac{(0,0,v)\times(a\cos\phi, a\sin\phi, z-d)}{(a^2+(z-d)^2)^{\frac{3}{2}}} - \frac{(0,0,v)\times(a\cos\phi, a\sin\phi, z+d)}{(a^2+(z+d)^2)^{\frac{3}{2}}}\right) =$$
$$= \frac{\mu_0 Q_m a v}{4\pi}\left(\frac{1}{(a^2+(z-d)^2)^{\frac{3}{2}}} - \frac{1}{(a^2+(z+d)^2)^{\frac{3}{2}}}\right)(\sin\phi, -\cos\phi, 0)$$

and

$$\mathbf{v}\times\mathbf{B}\cdot\frac{\mathrm{d}\mathbf{l}}{\mathrm{d}\phi} =$$
$$= \frac{\mu_0 Q_m a v}{4\pi}\left(\frac{1}{(a^2+(z-d)^2)^{\frac{3}{2}}} - \frac{1}{(a^2+(z+d)^2)^{\frac{3}{2}}}\right)(\sin\phi, -\cos\phi, 0)\cdot(-a\sin\phi, a\cos\phi, 0) =$$
$$= -\frac{\mu_0 Q_m a^2 v}{4\pi}\left(\frac{1}{(a^2+(z-d)^2)^{\frac{3}{2}}} - \frac{1}{(a^2+(z+d)^2)^{\frac{3}{2}}}\right)$$

Now we can calculate the voltage:

$$V = \oint (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \mathrm{d}\mathbf{l} =$$
$$= \oint \mathbf{v} \times \mathbf{B} \cdot \mathrm{d}\mathbf{l} =$$
$$= \int_0^{2\pi} \mathbf{v} \times \mathbf{B} \cdot \frac{\mathrm{d}\mathbf{l}}{\mathrm{d}\phi} \mathrm{d}\phi$$

for the dipole we get

$$V = \cdots = -\int_0^{2\pi} \frac{3\mu_0 m}{4\pi} \frac{a^2 v z}{(a^2+z^2)^{\frac{5}{2}}} \mathrm{d}\phi =$$
$$= -\frac{3\mu_0 m}{2} \frac{a^2 v z}{(a^2+z^2)^{\frac{5}{2}}}$$

For two monopoles we get

$$V = \cdots = -\int_0^{2\pi} \frac{\mu_0 Q_m a^2 v}{4\pi} \left(\frac{1}{(a^2+(z-d)^2)^{\frac{3}{2}}} - \frac{1}{(a^2+(z+d)^2)^{\frac{3}{2}}} \right) \mathrm{d}\phi =$$
$$= -\frac{\mu_0 Q_m a^2 v}{2} \left(\frac{1}{(a^2+(z-d)^2)^{\frac{3}{2}}} - \frac{1}{(a^2+(z+d)^2)^{\frac{3}{2}}} \right)$$

For the dipole, the function

$$\frac{z}{(a^2+z^2)^{\frac{5}{2}}}$$

has a maximum and minimum for:

$$z = \pm \frac{a}{2}$$

with the max value:

$$\frac{z}{(a^2+z^2)^{\frac{5}{2}}} = \frac{\frac{a}{2}}{(a^2+\left(\frac{a}{2}\right)^2)^{\frac{5}{2}}} = \frac{16\sqrt{5}}{125a^4}$$

Code:

```
>>> from sympy import var, solve, S, refine, Q
>>> var("a z")
(a, z)
>>> f = z / (a**2+z**2)**(S(5)/2)
>>> solve(f.diff(z), z)
[-a/2, a/2]
>>> f.subs(z, a/2)
16*sqrt(5)*a/(125*(a**2)**(5/2))
>>> refine(f.subs(z, a/2), Q.positive(a))
16*sqrt(5)/(125*a**4)
```

So the maximum voltage is:

$$V = \frac{\mu_0}{2} \frac{3va^2mz}{(a^2+z^2)^{\frac{5}{2}}} = \frac{\mu_0}{2} 3mva^2 \frac{16\sqrt{5}}{125a^4} =$$
$$= \frac{24\sqrt{5}}{125} \frac{\mu_0 mv}{a^2}$$

If we drop the magnet from height h above the coil into it, then its speed will be $v_0 = \sqrt{2hg}$ in the middle of the coil, when $t = 0$. Then:

$$z = v_0 t + \tfrac{1}{2} g t^2$$
$$v = v_0 + gt$$

And we get for the voltage dependence for dipole:

$$V = -\frac{\mu_0}{2} \frac{3va^2mz}{(a^2+z^2)^{\frac{5}{2}}} = -\frac{\mu_0}{2} \frac{3(v_0+gt)a^2m(v_0t+\frac{1}{2}gt^2)}{(a^2+(v_0t+\frac{1}{2}gt^2)^2)^{\frac{5}{2}}}$$

The time difference between the maximum and minimum is the time difference between $z = -\frac{a}{2}$ and $z = +\frac{a}{2}$, so:

$$\Delta t = \sqrt{\frac{2h+a}{g}} - \sqrt{\frac{2h-a}{g}}$$

The total flux doesn't depend on the particular dependence of $z(t)$ and $v(t)$:

$$\begin{aligned}
\Phi &= \int_0^\infty V(t)\mathrm{d}t = \\
&= -\frac{3\mu_0 m}{2} \int_0^\infty \frac{v(t)a^2 z(t)}{(a^2+z(t)^2)^{\frac{5}{2}}}\mathrm{d}t = \\
&= -\frac{3\mu_0 m}{2} \int_0^\infty \frac{\frac{\mathrm{d}z}{\mathrm{d}t}a^2 z(t)}{(a^2+z(t)^2)^{\frac{5}{2}}}\mathrm{d}t = \\
&= -\frac{3\mu_0 m}{2} \int_0^\infty \frac{a^2 z}{(a^2+z^2)^{\frac{5}{2}}}\mathrm{d}z = \\
&= -\frac{3\mu_0 m}{4} \int_{a^2}^\infty \frac{a^2}{u^{\frac{5}{2}}}\mathrm{d}u = \\
&= -\frac{3\mu_0 m a^2}{4} \left(-\frac{2}{3}\right) \left[\frac{1}{u^{\frac{3}{2}}}\right]_{a^2}^\infty = \\
&= -\frac{3\mu_0 m a^2}{4} \left(-\frac{2}{3}\right) \left[\frac{-1}{a^3}\right] = \\
&= -\frac{\mu_0 m}{2a}
\end{aligned}$$

For the voltage dependence of two monopoles, we get:

$$\begin{aligned}
V &= -\frac{\mu_0 Q_m a^2 v}{2} \left(\frac{1}{(a^2+(z-d)^2)^{\frac{3}{2}}} - \frac{1}{(a^2+(z+d)^2)^{\frac{3}{2}}}\right) = \\
&= -\frac{\mu_0 Q_m a^2 (v_0+gt)}{2} \left(\frac{1}{(a^2+(v_0t+\frac{1}{2}gt^2-d)^2)^{\frac{3}{2}}} - \frac{1}{(a^2+(v_0t+\frac{1}{2}gt^2+d)^2)^{\frac{3}{2}}}\right)
\end{aligned}$$

The total flux doesn't depend on the particular dependence of $z(t)$ and $v(t)$:

$$\begin{aligned}
\Phi = \int_0^\infty V(t)\mathrm{d}t = \\
= -\int_0^\infty \frac{\mu_0 Q_m a^2 v(t)}{2}\left(\frac{1}{(a^2+(z(t)-d)^2)^{\frac{3}{2}}} - \frac{1}{(a^2+(z(t)+d)^2)^{\frac{3}{2}}}\right)\mathrm{d}t = \\
= -\int_0^\infty \frac{\mu_0 Q_m a^2 \frac{\mathrm{d}z}{\mathrm{d}t}}{2}\left(\frac{1}{(a^2+(z(t)-d)^2)^{\frac{3}{2}}} - \frac{1}{(a^2+(z(t)+d)^2)^{\frac{3}{2}}}\right)\mathrm{d}t = \\
= -\int_0^\infty \frac{\mu_0 Q_m a^2}{2}\left(\frac{1}{(a^2+(z-d)^2)^{\frac{3}{2}}} - \frac{1}{(a^2+(z+d)^2)^{\frac{3}{2}}}\right)\mathrm{d}z = \\
= -\frac{\mu_0 Q_m a^2}{2}\left(\int_0^\infty \frac{1}{(a^2+(z-d)^2)^{\frac{3}{2}}}\mathrm{d}z - \int_0^\infty \frac{1}{(a^2+(z+d)^2)^{\frac{3}{2}}}\mathrm{d}z\right) = \\
= -\frac{\mu_0 Q_m a^2}{2}\left(\frac{1}{a^2}\left(1+\frac{d}{\sqrt{a^2+d^2}}\right) - \frac{1}{a^2}\left(1-\frac{d}{\sqrt{a^2+d^2}}\right)\right) = \\
= -\frac{\mu_0 Q_m d}{\sqrt{a^2+d^2}}
\end{aligned}$$

Note that in the limit $d \to 0$, we get the magnetic moment $m = 2dQ_m$ and the last formula for two monopoles flux becomes the dipole flux.

As a particular example, consider a coil with $N = 500$ loops, $a = 1.4\,\mathrm{cm}$, $d = 1.8\,\mathrm{cm}$, $Q_m = 43\,\mathrm{A}\cdot\mathrm{m}$. Then the total flux from the second peak is:

$$\Phi = -\frac{N\mu_0 Q_m d}{\sqrt{a^2+d^2}} = -0.021\,\mathrm{V}\cdot\mathrm{s}$$

Code:

```
>>> from math import pi, sqrt
>>> mu0 = 4*pi*1e-7
>>> cm = 0.01
>>> Q_m = 43.
>>> d = 1.8*cm
>>> a = 1.4*cm
>>> N = 500
>>> -N*mu0*Q_m*d/sqrt(a**2+d**2)
-0.02132647889395681
```

For a single loop with $a = 1.25\,\mathrm{cm}$ we get:

$$\Phi = -\frac{\mu_0 Q_m d}{\sqrt{a^2+d^2}} = -4.44\times 10^{-5}\,\mathrm{V}\cdot\mathrm{s}$$

and for a single loop with $a = 1.8\,\mathrm{cm}$ we get:

$$\Phi = -\frac{\mu_0 Q_m d}{\sqrt{a^2+d^2}} = -3.82\times 10^{-5}\,\mathrm{V}\cdot\mathrm{s}$$

Code:

```
>>> a = 1.25*cm
>>> -mu0*Q_m*d/sqrt(a**2+d**2)
-4.438304942066266e-05
>>> a = 1.8*cm
>>> -mu0*Q_m*d/sqrt(a**2+d**2)
-3.820879326816195e-05
```

RC Circuit

Let's consider resistor (with voltage $V = RI$) and capacitor (with voltage $V = \frac{Q}{C}$ and current $I(t) = Q'(t)$) in a series. Voltage on the battery is V, then the equation for the circuit is:

$$RI(t) + \frac{Q(t)}{C} = V$$

with initial condition $Q(0) = 0$. We differentiate it:

$$RI'(t) + \frac{I(t)}{C} = 0$$

and the initial condition follows from the first equation $I(0) = \frac{V}{R}$. The solution is:

$$I(t) = \frac{V}{R} e^{-\frac{t}{RC}}$$

Now we calculate the charge (using the initial condition for the charge above for the lower bound of the integral):

$$\begin{aligned} Q(t) = \int_0^t I(t')\mathrm{d}t' = \frac{V}{R}\int_0^t e^{-\frac{t'}{RC}}\,\mathrm{d}t' = \frac{V}{R}\left[-RCe^{-\frac{t'}{RC}}\right]_0^t = \\ = \frac{V}{R}\left[-RCe^{-\frac{t}{RC}} + RC\right] = VC\left(1 - e^{-\frac{t}{RC}}\right) \end{aligned}$$

The voltage on the resistor is:

$$RI(t) = R\frac{V}{R}e^{-\frac{t}{RC}} = Ve^{-\frac{t}{RC}}$$

The voltage on the capacitor is:

$$\frac{Q(t)}{C} = \frac{VC\left(1 - e^{-\frac{t}{RC}}\right)}{C} = V\left(1 - e^{-\frac{t}{RC}}\right)$$

Half life of the capacitor is defined as the time τ so that the charge is half of the total charge, and we get:

$$\begin{aligned} Q(\tau) &= \tfrac{1}{2}Q(\infty) \\ VC\left(1 - e^{-\frac{\tau}{RC}}\right) &= \tfrac{1}{2}VC \\ 1 - e^{-\frac{\tau}{RC}} &= \tfrac{1}{2} \\ \tfrac{1}{2} &= e^{-\frac{\tau}{RC}} \\ \log\tfrac{1}{2} &= -\frac{\tau}{RC} \\ \tau = -RC\log\tfrac{1}{2} &= RC\log 2 \end{aligned}$$

5.2 Semiconductor Device Physics

In general, the task is to find the five quantities:

$$n(\mathbf{x}, t), p(\mathbf{x}, t), \mathbf{J}_n(\mathbf{x}, t), \mathbf{J}_h(\mathbf{x}, t), \mathbf{E}(\mathbf{x}, t)$$

where n (p) is the electron (hole) concentration, $\mathbf{J}_n$ ($\mathbf{J}_p$) is the electron (hole) current density, $\mathbf{E}$ is the electric field.

And we have five equations that relate them. We start with the continuity equation:

$$\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = 0$$

where the current density $\mathbf{J}$ is composed of electron and hole current densities:

$$\mathbf{J} = \mathbf{J}_n + \mathbf{J}_p$$

and the charge density ρ is composed of mobile (electrons and holes) and fixed charges (ionized donors and acceptors):

$$\rho = q(p - n + C)$$

where n and p is the electron and hole concetration, C is the net doping concetration ($C = p_D - n_A$ where p_D is the concentration of ionized donors, charged positive, and n_A is the concentration of ionized acceptors, charged negative) and q is the electron charge (positive). We get:

$$\nabla \cdot \mathbf{J}_n + \nabla \cdot \mathbf{J}_p + q\left(\frac{\partial p}{\partial t} - \frac{\partial n}{\partial t} + \frac{\partial C}{\partial t}\right) = 0$$

Assuming the fixed charges C are time invariant, we get:

$$\nabla \cdot \mathbf{J}_n - q\frac{\partial n}{\partial t} = -\left(\nabla \cdot \mathbf{J}_p + q\frac{\partial p}{\partial t}\right) \equiv qR$$

where R is the net recombination rate for electrons and holes (a positive value means recombination, a negative value generation of carriers). We get the carrier continuity equations:

$$\begin{aligned} \frac{\partial n}{\partial t} &= -R + \frac{1}{q}\nabla \cdot \mathbf{J}_n \\ \frac{\partial p}{\partial t} &= -R - \frac{1}{q}\nabla \cdot \mathbf{J}_p \end{aligned} \tag{5.2.1}$$

Then we need material relations that express how the current $\mathbf{J}$ is generated using $\mathbf{E}$ and n and p. A drift-diffusion model is to assume a drift current ($q\mu_n n\mathbf{E}$) and a diffusion ($qD_n\nabla n$), which gives:

$$\begin{aligned} \mathbf{J}_n &= q\mu_n n\mathbf{E} + qD_n\nabla n \\ \mathbf{J}_p &= q\mu_p p\mathbf{E} - qD_p\nabla p \end{aligned} \tag{5.2.2}$$

where μ_n, μ_p, D_n, D_p are the carrier mobilities and diffusivities.

Final equation is the Gauss's law:

$$\begin{aligned} \nabla \cdot (\varepsilon\mathbf{E}) &= \rho \\ \nabla \cdot (\varepsilon\mathbf{E}) &= q(p - n + C) \end{aligned} \tag{5.2.3}$$

5.2.1 Equations

Combining (5.2.2) and (5.2.1) we get the following three equations for three unknowns n, p and $\mathbf{E}$:

$$\begin{aligned} \frac{\partial n}{\partial t} &= -R + \nabla \cdot (\mu_n n\mathbf{E}) + \nabla \cdot (D_n\nabla n) \\ \frac{\partial p}{\partial t} &= -R - \nabla \cdot (\mu_p p\mathbf{E}) + \nabla \cdot (D_p\nabla p) \\ \nabla \cdot (\varepsilon\mathbf{E}) &= q(p - n + C) \end{aligned}$$

And it is usually assumed that the magnetic field is time independent, so $\mathbf{E} = -\nabla\phi$ and we get:

$$\begin{aligned} \frac{\partial n}{\partial t} &= -R - \nabla \cdot (\mu_n n\nabla\phi) + \nabla \cdot (D_n\nabla n) \\ \frac{\partial p}{\partial t} &= -R + \nabla \cdot (\mu_p p\nabla\phi) + \nabla \cdot (D_p\nabla p) \\ \nabla \cdot (\varepsilon\nabla\phi) &= -q(p - n + C) \end{aligned} \tag{5.2.1.1}$$

These are three nonlinear (due to the terms $\mu_n n\nabla\phi$ and $\mu_p p\nabla\phi$) equations for three unknown functions n, p and ϕ.

Example 1

We can substract the first two equations and we get:

$$\frac{\partial q(p-n)}{\partial t} = -q\nabla\cdot((\mu_p p + \mu_n n)\mathbf{E}) + q\nabla\cdot(D_p\nabla p - D_n\nabla n)$$
$$\nabla\cdot(\varepsilon\mathbf{E}) = q(p-n+C)$$

and using $\rho = q(p-n+C)$ and $\sigma = q(\mu_p p + \mu_n n)$, we get:

$$\frac{\partial\rho}{\partial t} - q\frac{\partial C}{\partial t} = -\nabla\cdot(\sigma\mathbf{E}) + q\nabla\cdot(D_p\nabla p - D_n\nabla n)$$
$$\nabla\cdot(\varepsilon\mathbf{E}) = \rho$$

So far we didn't make any assumptions. Most of the times the net doping concetration C is time independent, which gives:

$$\frac{\partial\rho}{\partial t} = -\nabla\cdot(\sigma\mathbf{E}) + q\nabla\cdot(D_p\nabla p - D_n\nabla n)$$
$$\nabla\cdot(\varepsilon\mathbf{E}) = \rho$$

Assuming further $D_p\nabla p - D_n\nabla n = 0$, we just get the equation of continuity and the Gauss law:

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\sigma\mathbf{E}) = 0$$
$$\nabla\cdot(\varepsilon\mathbf{E}) = \rho$$

Finally, assuming also that that ρ doesn't depend on time, we get:

$$\nabla\cdot(\sigma\mathbf{E}) = 0$$
$$\nabla\cdot(\varepsilon\mathbf{E}) = \rho$$

Example 2

As a simple model, assume D_n, D_p, μ_n, μ_p and ε are position independent and $C = 0$, $R = 0$:

$$\frac{\partial n}{\partial t} = +\mu_n n\nabla\cdot\mathbf{E} + \mu_n\mathbf{E}\cdot\nabla n + D_n\nabla^2 n$$
$$\frac{\partial p}{\partial t} = -\mu_p p\nabla\cdot\mathbf{E} - \mu_p\mathbf{E}\cdot\nabla p + D_p\nabla^2 p$$
$$\varepsilon\nabla\cdot\mathbf{E} = q(p-n)$$

Using $\mathbf{E} = -\nabla\phi$ we get:

$$\frac{\partial n}{\partial t} = -\mu_n n\nabla^2\phi - \mu_n\nabla\phi\cdot\nabla n + D_n\nabla^2 n$$
$$\frac{\partial p}{\partial t} = +\mu_p p\nabla^2\phi + \mu_p\nabla\phi\cdot\nabla p + D_p\nabla^2 p$$
$$\varepsilon\nabla^2\phi = -q(p-n)$$

5.2.2 Example 3

Let's calculate the 1D pn-junction. We take the equations (5.2.1.1) and write them in 1D for the stationary state ($\frac{\partial n}{\partial t} = \frac{\partial p}{\partial t} = 0$):

$$0 = -R - (\mu_n n\phi')' + (D_n n')'$$
$$0 = -R + (\mu_p p\phi')' + (D_p p')'$$
$$(\varepsilon\phi')' = -q(p-n+C)$$

We expand the derivatives and assume that μ and D is constant:

$$\begin{aligned} 0 &= -R - \mu_n n'\phi' - \mu_n n\phi'' + D_n n'' \\ 0 &= -R + \mu_p p'\phi' + \mu_p p\phi'' + D_p p'' \\ \varepsilon\phi'' &= -q(p - n + C) \end{aligned}$$

and we put the second derivatives on the left hand side:

$$\begin{aligned} n'' &= \frac{1}{D_n}(R + \mu_n n'\phi' + \mu_n n\phi'') \\ p'' &= \frac{1}{D_p}(R - \mu_p p'\phi' - \mu_p p\phi'') \\ \phi'' &= -\frac{q}{\varepsilon}(p - n + C) \end{aligned} \tag{5.2.2.1}$$

now we introduce the variables y_i:

$$\begin{aligned} y_0 &= n \\ y_1 = y_0' &= n' \\ y_2 &= p \\ y_3 = y_2' &= p' \\ y_4 &= \phi \\ y_5 = y_4' &= \phi' \end{aligned}$$

and rewrite (5.2.2.1):

$$\begin{aligned} y_1' &= \frac{1}{D_n}(R + \mu_n y_1 y_5 + \mu_n y_0 y_5') \\ y_3' &= \frac{1}{D_p}(R - \mu_p y_3 y_5 - \mu_p y_2 y_5') \\ y_5' &= -\frac{q}{\varepsilon}(y_2 - y_0 + C) \end{aligned}$$

So we are solving the following six nonlinear first order ODE:

$$\begin{aligned} y_5' &= -\frac{q}{\varepsilon}(y_2 - y_0 + C) \\ y_0' &= y_1 \\ y_1' &= \frac{1}{D_n}(R + \mu_n y_1 y_5 + \mu_n y_0 y_5') \\ y_2' &= y_3 \\ y_3' &= \frac{1}{D_p}(R - \mu_p y_3 y_5 - \mu_p y_2 y_5') \\ y_4' &= y_5 \end{aligned} \tag{5.2.2.2}$$

CHAPTER

SIX

THERMODYNAMICS AND STATISTICAL PHYSICS

6.1 Thermodynamics

6.1.1 Thermodynamic Potentials

We start by writing the internal energy

$$U = U(S, V, N) = \int (T\mathrm{d}S - p\mathrm{d}V + \mu\mathrm{d}N)$$

as a function of entropy S, volume V and a number of particles N. Now we want to express it as a function of temperature T, pressure p or a chemical potential μ, without losing any information, i.e. we still want to just differentiate to obtain other quantities. In order to do that, we have to use the Legendre transform. Including U, there are 8 possible combinations of Legendre transforms that one can do (three of them are applying it to just one variable, three of them to two variables, one to all three variables):

$$\begin{aligned}
U = U(S, V, N) &= \int (T\mathrm{d}S - p\mathrm{d}V + \mu\mathrm{d}N) \\
F = F(T, V, N) = U - TS &= \int (-S\mathrm{d}T - p\mathrm{d}V + \mu\mathrm{d}N) \\
H = H(S, p, N) = U + pV &= \int (T\mathrm{d}S + V\mathrm{d}p + \mu\mathrm{d}N) \\
X_1 = X_1(S, V, \mu) = U - \mu N &= \int (T\mathrm{d}S - p\mathrm{d}V - N\mathrm{d}\mu) \\
G = G(T, p, N) = U - TS + pV &= \int (-S\mathrm{d}T + V\mathrm{d}p + \mu\mathrm{d}N) \\
\Omega = \Omega(T, V, \mu) = U - TS - \mu N &= \int (-S\mathrm{d}T - p\mathrm{d}V - N\mathrm{d}\mu) \\
X_2 = X_2(S, p, \mu) = U + pV - \mu N &= \int (T\mathrm{d}S + V\mathrm{d}p - N\mathrm{d}\mu) \\
X_3 = X_3(T, p, \mu) = U - TS + pV - \mu N &= \int (-S\mathrm{d}T + V\mathrm{d}p - N\mathrm{d}\mu)
\end{aligned}$$

Of these, U is the internal energy, F is the Helmholtz free energy, H is the enthalpy, G is the Gibbs free energy, Ω is the grand potential (sometimes also called a Landau potential). The unnamed potentials are simply labeled X_1, X_2 and X_3. The X_3 is sometimes called a null function.

From the differentials, we can then read off the derivatives (and what other variables are constant), here are all the

combinations:

$$
\begin{aligned}
T &= \left(\frac{\partial U}{\partial S}\right)_{V,N} = \left(\frac{\partial H}{\partial S}\right)_{p,N} = \left(\frac{\partial X_1}{\partial S}\right)_{V,\mu} = \left(\frac{\partial X_2}{\partial S}\right)_{p,\mu} \\
S &= -\left(\frac{\partial F}{\partial T}\right)_{V,N} = -\left(\frac{\partial G}{\partial T}\right)_{p,N} = -\left(\frac{\partial \Omega}{\partial T}\right)_{V,\mu} = -\left(\frac{\partial X_3}{\partial T}\right)_{p,\mu} \\
p &= -\left(\frac{\partial U}{\partial V}\right)_{S,N} = -\left(\frac{\partial F}{\partial V}\right)_{T,N} = -\left(\frac{\partial X_1}{\partial V}\right)_{S,\mu} = -\left(\frac{\partial \Omega}{\partial V}\right)_{T,\mu} \\
V &= \left(\frac{\partial H}{\partial p}\right)_{S,N} = \left(\frac{\partial G}{\partial p}\right)_{T,N} = \left(\frac{\partial X_2}{\partial p}\right)_{S,\mu} = \left(\frac{\partial X_3}{\partial p}\right)_{T,\mu} \\
\mu &= \left(\frac{\partial U}{\partial N}\right)_{S,V} = \left(\frac{\partial F}{\partial N}\right)_{T,V} = \left(\frac{\partial H}{\partial N}\right)_{S,p} = \left(\frac{\partial G}{\partial N}\right)_{T,p} \\
N &= -\left(\frac{\partial X_1}{\partial \mu}\right)_{S,V} = -\left(\frac{\partial \Omega}{\partial \mu}\right)_{T,V} = -\left(\frac{\partial X_2}{\partial \mu}\right)_{S,p} = -\left(\frac{\partial X_3}{\partial \mu}\right)_{T,p}
\end{aligned}
$$

A large system is defined as: if the number of particles N is made λ times as large, U, V, and S all become λ times larger. In other words, the internal energy of a large system is a homogeneous function of S, V, and N of order one:

$$U(\lambda S, \lambda V, \lambda N) = \lambda U(S, V, N)$$

Now we can apply the Euler's theorem (see *Homogeneous Functions (Euler's Theorem)*):

$$U(S, V, N) = S\left(\frac{\partial U}{\partial S}\right)_{V,N} + V\left(\frac{\partial U}{\partial V}\right)_{S,N} + N\left(\frac{\partial U}{\partial N}\right)_{S,V} = TS - pV + \mu N$$

And from the definitions of all the potentials we can calculate their forms for large systems:

$$
\begin{aligned}
U(S, V, N) &= TS - pV + \mu N \\
F(T, V, N) &= U - TS = -pV + \mu N \\
H(S, p, N) &= U + pV = TS + \mu N \\
X_1(S, V, \mu) &= U - \mu N = TS - pV \\
G(T, p, N) &= U - TS + pV = \mu N \\
\Omega(T, V, \mu) &= U - TS - \mu N = -pV \\
X_2(S, p, \mu) &= U + pV - \mu N = TS \\
X_3(T, p, \mu) &= U - TS + pV - \mu N = 0
\end{aligned}
$$

6.1.2 Examples

Ideal Gas

The internal energy as a function of S, V and N is equal to:

$$U(S, V, N) = c_V N k_{\mathrm{B}} \left(\frac{N\Phi}{V} e^{\frac{S}{N k_{\mathrm{B}}}}\right)^{\frac{1}{c_V}} \tag{6.1.2.1}$$

where c_V is the heat capacity at a constant volume ($\frac{3}{2}$ for monoatomic gases, $\frac{5}{2}$ for diatomic gases), k_{B} is the Boltzman constant and Φ is a constant that may vary for different gases, but it is independent of the thermodynamic state of the gas.

At this level, the above expression is simply given. We would have to use statistical physics in order to calculate any of the thermodynamic potentials.

Now we calculate the free energy $F(T,V,N)$. First we must calculate the temperature T:

$$\begin{aligned} T = \left(\frac{\partial U}{\partial S}\right)_{V,N} = \\ = \frac{\partial}{\partial S}\left(c_V N k_\mathrm{B}\left(\frac{N\Phi}{V}e^{\frac{S}{Nk_\mathrm{B}}}\right)^{\frac{1}{c_V}}\right) = \\ = \left(\frac{N\Phi}{V}e^{\frac{S}{Nk_\mathrm{B}}}\right)^{\frac{1}{c_V}}. \end{aligned} \tag{6.1.2.2}$$

In order to calculate the the free energy, we must use (6.1.2.2) to eliminate S:

$$S = Nk_\mathrm{B}\log\left(\frac{VT^{c_V}}{N\Phi}\right) \tag{6.1.2.3}$$

and then express F as a function of T, V and N only:

$$\begin{aligned} F(T,V,N) = U - TS = \\ = c_V N k_\mathrm{B}\left(\frac{N\Phi}{V}e^{\frac{S}{Nk_\mathrm{B}}}\right)^{\frac{1}{c_V}} - TS = \\ = c_V N k_\mathrm{B} T - TNk_\mathrm{B}\log\left(\frac{VT^{c_V}}{N\Phi}\right) = \\ = Nk_\mathrm{B}T\left(c_V - \log\left(\frac{VT^{c_V}}{N\Phi}\right)\right). \end{aligned} \tag{6.1.2.4}$$

This calculation shows that one can also express the internal energy as a function of T, V and N as $U = U(T,V,N) = c_V N k_\mathrm{B} T$. This is a valid expression, but unlike $U = U(S,V,N)$, this is not a thermodynamic potential, because we lost some information. In particular, if we use $U = U(T,V,N)$ to find $U = U(S,V,N)$:

$$\begin{aligned} U = U(T,V,N) = c_V N k_\mathrm{B} T = c_V N k_\mathrm{B}\left(\frac{\partial U}{\partial S}\right)_{V,N} \\ \mathrm{d}S = c_V N k_\mathrm{B}\frac{\mathrm{d}U}{U} \qquad (V \text{ and } N \text{ constant}) \\ S = c_V N k_\mathrm{B}\log U + C \qquad (V \text{ and } N \text{ constant}) \\ U(S,V,N) = f(V,N)\left(e^{\frac{S}{Nk_\mathrm{B}}}\right)^{\frac{1}{c_V}}, \end{aligned}$$

we can see, that we recovered the correct formula for $U(S,V,N)$ except an arbitrary function $f(V,N)$ of V and N. Compared to (6.1.2.1) we can see that it must be $f(V,N) = c_V N k_\mathrm{B}\left(\frac{N\Phi}{V}\right)^{\frac{1}{c_V}}$, but this information got lost. For this reason, only $U = U(S,V,N)$ as well as $F = F(T,V,N)$, that we just calculated, are thermodynamic potentials and both contain equivalent information. But $U = U(T,V,N)$ is not and it does not contain full information.

To convert $F(T,V,N)$ back to $U(S,V,N)$, we first calculate the entropy S:

$$\begin{aligned} S = -\left(\frac{\partial F}{\partial T}\right)_{V,N} = \\ = -\frac{\partial}{\partial T}\left(Nk_\mathrm{B}T\left(c_V - \log\left(\frac{VT^{c_V}}{N\Phi}\right)\right)\right) = \\ = -Nk_\mathrm{B}c_V + Nk_\mathrm{B}\log\left(\frac{VT^{c_V}}{N\Phi}\right) + Nk_\mathrm{B}T\frac{N\Phi}{VT^{c_V}}\frac{Vc_V T^{c_V-1}}{N\Phi} = \\ = Nk_\mathrm{B}\log\left(\frac{VT^{c_V}}{N\Phi}\right), \end{aligned}$$

which is the same equation as (6.1.2.3). From this, we express T, we get (6.1.2.2). Finally, we can calculate the internal energy and substitute T for S using (6.1.2.2):

$$\begin{aligned} U(S,V,N) = F + TS = \\ = Nk_{\mathrm{B}}T\left(c_V - \log\left(\frac{VT^{c_V}}{N\Phi}\right)\right) + TS = \\ = Nk_{\mathrm{B}}Tc_V - Nk_{\mathrm{B}}T\log\left(\frac{VT^{c_V}}{N\Phi}\right) + TS = \\ = Nk_{\mathrm{B}}Tc_V - TS + TS = \\ = c_V Nk_{\mathrm{B}}T = \\ = c_V Nk_{\mathrm{B}}\left(\frac{N\Phi}{V}e^{\frac{S}{Nk_{\mathrm{B}}}}\right)^{\frac{1}{c_V}}. \end{aligned}$$

This is the same equation as (6.1.2.1). This shows that all thermodynamic potentials contain the same information and can be converted to one another using the Legendre transformation.

Note: in equations like $F(T,V,N) = U - TS$, we can use any expressions for U and S (e.g. we can use $U = U(S,V,N)$ or $U = (T,V,N)$, etc.) in the intermediate steps, but at the end, we must express the final formula using T, V and N only.

To calculate the Gibbs energy, we need to calculate pressure first. We can use any of the potentials U, F, X_1 or Ω to do so. Since the equation of state is typicaly expressed as $p = p(T,V,N)$, then the free energy $F(T,V,N)$ is the natural choice:

$$\begin{aligned} p = -\left(\frac{\partial F}{\partial V}\right)_{T,N} = \\ = -\frac{\partial}{\partial V}\left(Nk_{\mathrm{B}}T\left(c_V - \log\left(\frac{VT^{c_V}}{N\Phi}\right)\right)\right) = \\ = Nk_{\mathrm{B}}T\frac{\partial}{\partial V}\log\left(\frac{VT^{c_V}}{N\Phi}\right) = \\ = Nk_{\mathrm{B}}T\frac{1}{V}, \end{aligned}$$

and we get the ideal gas law $pV = Nk_{\mathrm{B}}T$. The Gibbs energy is equal to:

$$\begin{aligned} G(T,p,N) = U - TS + pV = F + pV = \\ = Nk_{\mathrm{B}}T\left(c_V - \log\left(\frac{VT^{c_V}}{N\Phi}\right)\right) + Nk_{\mathrm{B}}T = \\ = Nk_{\mathrm{B}}T\left((c_V+1) - \log\left(\frac{k_{\mathrm{B}}T^{c_V+1}}{p\Phi}\right)\right). \end{aligned} \tag{6.1.2.5}$$

For the enthalpy, we first need:

$$\begin{aligned} p = -\left(\frac{\partial U}{\partial V}\right)_{S,N} = \\ = -c_V Nk_{\mathrm{B}}\frac{1}{c_V}\left(\frac{N\Phi}{V}e^{\frac{S}{Nk_{\mathrm{B}}}}\right)^{\frac{1}{c_V}-1}\frac{N\Phi}{V}e^{\frac{S}{Nk_{\mathrm{B}}}}\left(-\frac{1}{V}\right) = \\ = \frac{1}{V}Nk_{\mathrm{B}}\left(\frac{N\Phi}{V}e^{\frac{S}{Nk_{\mathrm{B}}}}\right)^{\frac{1}{c_V}}, \end{aligned}$$

we need to use this to express the volume V:

$$V^{\frac{c_V+1}{c_V}} = \frac{Nk_\mathrm{B}}{p}\left(N\Phi e^{\frac{S}{Nk_\mathrm{B}}}\right)^{\frac{1}{c_V}} = \left(\frac{N^{c_V+1}k_\mathrm{B}^{c_V}}{p^{c_V}}\Phi e^{\frac{S}{Nk_\mathrm{B}}}\right)^{\frac{1}{c_V}}$$
$$V = \left(\frac{N^{c_V+1}k_\mathrm{B}^{c_V}}{p^{c_V}}\Phi e^{\frac{S}{Nk_\mathrm{B}}}\right)^{\frac{1}{c_V+1}} = \frac{Nk_\mathrm{B}}{p}\left(\frac{p\Phi}{k_\mathrm{B}}e^{\frac{S}{Nk_\mathrm{B}}}\right)^{\frac{1}{c_V+1}}$$

now we can calculate $H(S,p,N)$:

$$\begin{aligned} H(S,p,N) = U + pV = \\ = c_V Nk_\mathrm{B}\left(\frac{N\Phi}{V}e^{\frac{S}{Nk_\mathrm{B}}}\right)^{\frac{1}{c_V}} + pV = \\ = (c_V+1)pV = \\ = (c_V+1)Nk_\mathrm{B}\left(\frac{p\Phi}{k_\mathrm{B}}e^{\frac{S}{Nk_\mathrm{B}}}\right)^{\frac{1}{c_V+1}}. \end{aligned} \tag{6.1.2.6}$$

The enthalpy in terms of temperature $H = H(T,p,N)$ can be calculated as:

$$H(T,p,N) = (c_V+1)pV = (c_V+1)Nk_\mathrm{B}T\,.$$

The specific heat capacity at a constant volume can be calculated as:

$$\begin{aligned} c_V \equiv \frac{1}{Nk_\mathrm{B}}C_V = \frac{1}{Nk_\mathrm{B}}\left(\frac{\partial U}{\partial T}\right)_{V,N} = \\ = \frac{1}{Nk_\mathrm{B}}\frac{\partial}{\partial T}\left(c_V Nk_\mathrm{B}T\right) = c_V \end{aligned}$$

This provides proof that the c_V in (6.1.2.1) is indeed the specific heat capacity at a constant volume.

The specific heat capacity at a constant pressure can be calculated as:

$$\begin{aligned} c_p \equiv \frac{1}{Nk_\mathrm{B}}C_p = \frac{1}{Nk_\mathrm{B}}\left(\frac{\partial H}{\partial T}\right)_{p,N} = \\ = \frac{1}{Nk_\mathrm{B}}\frac{\partial H(T,p,N)}{\partial T} = \\ = \frac{1}{Nk_\mathrm{B}}\frac{\partial}{\partial T}\left((c_V+1)Nk_\mathrm{B}T\right) = c_V + 1\,. \end{aligned}$$

Using this relation $c_p = c_V + 1$ we can then express (6.1.2.5):

$$\begin{aligned} G(T,p,N) = Nk_\mathrm{B}T\left((c_V+1) - \log\left(\frac{k_\mathrm{B}T^{c_V+1}}{p\Phi}\right)\right) = \\ = Nk_\mathrm{B}T\left(c_p - \log\left(\frac{k_\mathrm{B}T^{c_p}}{p\Phi}\right)\right), \end{aligned}$$

and (6.1.2.6) as:

$$\begin{aligned} H(S,p,N) = (c_V+1)Nk_\mathrm{B}\left(\frac{p\Phi}{k_\mathrm{B}}e^{\frac{S}{Nk_\mathrm{B}}}\right)^{\frac{1}{c_V+1}} = \\ = c_p Nk_\mathrm{B}\left(\frac{p\Phi}{k_\mathrm{B}}e^{\frac{S}{Nk_\mathrm{B}}}\right)^{\frac{1}{c_p}}. \end{aligned}$$

In order to calculate the grand potential, we first need to find the chemical potential:

$$\begin{aligned}\mu = \left(\frac{\partial F}{\partial N}\right)_{T,V} = \\ = \frac{\partial}{\partial N}\left(Nk_{\mathrm{B}}T\left(c_V - \log\left(\frac{VT^{c_V}}{N\Phi}\right)\right)\right) = \\ = k_{\mathrm{B}}T\left((c_V+1) - \log\left(\frac{VT^{c_V}}{N\Phi}\right)\right),\end{aligned}$$

and express N using μ:

$$N = \frac{VT^{c_V}}{\Phi e^{c_V+1-\frac{\mu}{k_{\mathrm{B}}T}}}$$

Now we can calculate $\Omega(T,V,\mu)$:

$$\begin{aligned}\Omega(T,V,\mu) = U - TS - \mu N = F - \mu N = \\ = Nk_{\mathrm{B}}T\left(c_V - \log\left(\frac{VT^{c_V}}{N\Phi}\right)\right) - \mu N = \\ = Nk_{\mathrm{B}}T\left(\frac{\mu}{k_{\mathrm{B}}T} - 1\right) - \mu N = \\ = -Nk_{\mathrm{B}}T = \\ = -\frac{k_{\mathrm{B}}VT^{c_V+1}}{\Phi e^{c_V+1-\frac{\mu}{k_{\mathrm{B}}T}}} = \\ = -\frac{k_{\mathrm{B}}VT^{c_p}}{\Phi e^{c_p-\frac{\mu}{k_{\mathrm{B}}T}}}.\end{aligned} \tag{6.1.2.7}$$

6.2 Statistical Physics

6.2.1 Microcanonical Ensemble

The entropy is equal to:

$$S = k_{\mathrm{B}}\log W = \frac{1}{\beta T}\log W$$

where W is the micronanonical partition function, the number of microstates within the range of energy.

6.2.2 Canonical Ensemble

The partition function is:

$$Z_{can} = \sum_n e^{-\beta E_n}$$

The Helmholtz free energy is equal to:

$$F = -\frac{1}{\beta}\log Z_{can} = -\frac{1}{\beta}\log\sum_n e^{-\beta E_n}$$

6.2.3 Grand Canonical Ensemble

The partition function for fermions is:

$$\begin{aligned}
Z_{gr} = \sum_n e^{-\beta(E_n - \mu N_n)} = \\
= \left(\prod_\alpha \sum_{n_\alpha=0}^{1} \right) e^{-\beta(E_n - \mu N_n)} = \\
= \left(\prod_\alpha \sum_{n_\alpha=0}^{1} \right) e^{-\beta\left(\sum_\alpha n_\alpha \epsilon_\alpha - \mu \sum_\alpha n_\alpha\right)} = \\
= \left(\prod_\alpha \sum_{n_\alpha=0}^{1} \right) \prod_\alpha e^{-\beta n_\alpha(\epsilon_\alpha - \mu)} = \\
= \prod_\alpha \left(\sum_{n_\alpha=0}^{1} e^{-\beta n_\alpha(\epsilon_\alpha - \mu)} \right) = \\
= \prod_\alpha \left(1 + e^{-\beta(\epsilon_\alpha - \mu)}\right)
\end{aligned}$$

Similarly, for bosons we would get:

$$\begin{aligned}
Z_{gr} = \sum_n e^{-\beta(E_n - \mu N_n)} = \\
= \left(\prod_\alpha \sum_{n_\alpha=0}^{\infty} \right) e^{-\beta(E_n - \mu N_n)} = \\
= \left(\prod_\alpha \sum_{n_\alpha=0}^{\infty} \right) e^{-\beta\left(\sum_\alpha n_\alpha \epsilon_\alpha - \mu \sum_\alpha n_\alpha\right)} = \\
= \left(\prod_\alpha \sum_{n_\alpha=0}^{\infty} \right) \prod_\alpha e^{-\beta n_\alpha(\epsilon_\alpha - \mu)} = \\
= \prod_\alpha \left(\sum_{n_\alpha=0}^{\infty} e^{-\beta n_\alpha(\epsilon_\alpha - \mu)} \right) = \\
= \prod_\alpha \left(1 - e^{-\beta(\epsilon_\alpha - \mu)}\right)^{-1}
\end{aligned}$$

The grand potential for fermions is then equal to:

$$\begin{aligned}
\Omega = -\frac{1}{\beta} \log Z_{gr} = \\
= -\frac{1}{\beta} \log \left(\prod_\alpha \left(1 + e^{-\beta(\epsilon_\alpha - \mu)}\right) \right) = \\
= -\frac{1}{\beta} \sum_\alpha \log \left(1 + e^{-\beta(\epsilon_\alpha - \mu)}\right)
\end{aligned}$$

Similarly, the grand potential for bosons is equal to:

$$\begin{aligned}\Omega = -\frac{1}{\beta}\log Z_{gr} = \\ = -\frac{1}{\beta}\log\left(\prod_\alpha \left(1-e^{-\beta(\epsilon_\alpha-\mu)}\right)^{-1}\right) = \\ = -\frac{1}{\beta}\sum_\alpha \log\left(1-e^{-\beta(\epsilon_\alpha-\mu)}\right)^{-1} = \\ = \frac{1}{\beta}\sum_\alpha \log\left(1-e^{-\beta(\epsilon_\alpha-\mu)}\right)\end{aligned}$$

6.2.4 Examples

Ideal Gas

Ideal gas is simply a system of classical particles, where for a given microstate specified by a set of coordinates $\mathbf{x}_i$ and momenta $\mathbf{p}_i$, the total energy of the microstate is given by the following Hamiltonian:

$$H(\mathbf{x}_i, \mathbf{p}_i) = \sum_{i=1}^{N} \frac{p_i^2}{2m},$$

that is, the particles are non-interacting, each has a mass m and a momentum $\mathbf{p}_i$. The canonical partition function is then equal to:

$$\begin{aligned}Z_{can}(T,V,N) = \sum_n e^{-\beta E_n} = \\ = \int \frac{\mathrm{d}^{3N}x\,\mathrm{d}^{3N}p}{N!(2\pi\hbar)^{3N}} e^{-\beta H(\mathbf{x}_i,\mathbf{p}_i)} = \\ = \int \frac{\mathrm{d}^{3N}x\,\mathrm{d}^{3N}p}{N!(2\pi\hbar)^{3N}} e^{-\beta\sum_{i=1}^{N}\frac{p_i^2}{2m}} = \\ = \frac{1}{N!}\left(\int \frac{\mathrm{d}^3x\mathrm{d}^3p}{(2\pi\hbar)^3} e^{-\beta\frac{p^2}{2m}}\right)^N = \\ = \frac{1}{N!}\left(V\int_0^\infty \frac{4\pi p^2\mathrm{d}p}{(2\pi\hbar)^3} e^{-\beta\frac{p^2}{2m}}\right)^N = \\ = \frac{1}{N!}\left(V\frac{4\pi}{(2\pi\hbar)^3}\frac{\sqrt{\pi}(2m)^{\frac{3}{2}}}{4\beta^{\frac{3}{2}}}\right)^N = \\ = \frac{1}{N!}\left(\left(\frac{m}{2\pi\hbar^2\beta}\right)^{\frac{3}{2}}V\right)^N,\end{aligned}$$

where we used the following integral:

$$\int_0^\infty p^2 e^{-\alpha p^2}\mathrm{d}p = \frac{\sqrt{\pi}}{4\alpha^{\frac{3}{2}}}.$$

The Helmholtz free energy is then equal to:

$$\begin{aligned}
F(T,V,N) = -\frac{1}{\beta}\log Z_{can}(T,V,N) = \\
= -\frac{1}{\beta}\log\left(\frac{1}{N!}\left(\left(\frac{m}{2\pi\hbar^2\beta}\right)^{\frac{3}{2}}V\right)^N\right) = \\
= \frac{N}{\beta}\left(\frac{\log N!}{N} - \log\left(\left(\frac{m}{2\pi\hbar^2\beta}\right)^{\frac{3}{2}}V\right)\right) = \\
= \frac{N}{\beta}\left(\log N - 1 + O\left(\frac{\log N}{N}\right) - \log\left(\left(\frac{m}{2\pi\hbar^2\beta}\right)^{\frac{3}{2}}V\right)\right) = \\
= \frac{N}{\beta}\left(-\log\left(\left(\frac{m}{2\pi\hbar^2\beta}\right)^{\frac{3}{2}}\frac{Ve}{N}\right) + O\left(\frac{\log N}{N}\right)\right) = \\
= \frac{N}{\beta}\left(\frac{3}{2} - \log\left(\left(\frac{m}{2\pi\hbar^2\beta}\right)^{\frac{3}{2}}\frac{Ve^{\frac{5}{2}}}{N}\right) + O\left(\frac{\log N}{N}\right)\right) = \\
= Nk_{\mathrm{B}}T\left(\frac{3}{2} - \log\left(\frac{VT^{\frac{3}{2}}}{N\left(\frac{2\pi\hbar^2}{mk_{\mathrm{B}}e^{\frac{5}{3}}}\right)^{\frac{3}{2}}}\right) + O\left(\frac{\log N}{N}\right)\right),
\end{aligned}$$

where we used the Stirling's approximation for $N!$. For large N this is equal to the Helmholtz free energy of the ideal gas (see *Ideal Gas*):

$$F(T,V,N) = Nk_{\mathrm{B}}T\left(c_V - \log\left(\frac{VT^{c_V}}{N\Phi}\right)\right),$$

with $c_V = \frac{3}{2}$ and $\Phi = \left(\frac{2\pi\hbar^2}{mk_{\mathrm{B}}e^{\frac{5}{3}}}\right)^{\frac{3}{2}}$. See that section where all other thermodynamic properties are derived from it.

We can also start from the grand canonical partition function:

$$\begin{aligned}
Z_{gr}(T,V,\mu) = \sum_{N=0}^{\infty} e^{\beta\mu N} Z_{can}(T,V,N) = \\
= \sum_{N=0}^{\infty} e^{\beta\mu N}\frac{1}{N!}\left(\left(\frac{m}{2\pi\hbar^2\beta}\right)^{\frac{3}{2}}V\right)^N = \\
= \sum_{N=0}^{\infty}\frac{1}{N!}\left(e^{\beta\mu}\left(\frac{m}{2\pi\hbar^2\beta}\right)^{\frac{3}{2}}V\right)^N = \\
= e^{e^{\beta\mu}\left(\frac{m}{2\pi\hbar^2\beta}\right)^{\frac{3}{2}}V}
\end{aligned}$$

And the grand potential is:

$$\begin{aligned}\Omega(T,V,\mu) = -\frac{1}{\beta}\log Z_{gr}(T,V,\mu) = \\ = -\frac{1}{\beta}e^{\beta\mu}\left(\frac{m}{2\pi\hbar^2\beta}\right)^{\frac{3}{2}}V = \\ = -k_{\mathrm{B}}Te^{\frac{\mu}{k_{\mathrm{B}}T}}\left(\frac{mk_{\mathrm{B}}T}{2\pi\hbar^2}\right)^{\frac{3}{2}}V = \\ = -\frac{k_{\mathrm{B}}VT^{\frac{5}{2}}}{\left(\frac{2\pi\hbar^2}{mk_{\mathrm{B}}e^{\frac{5}{3}}}\right)^{\frac{3}{2}}e^{\frac{5}{2}-\frac{\mu}{k_{\mathrm{B}}T}}}.\end{aligned}$$

This is equal to the grand potential of an ideal gas:

$$\Omega(T,V,\mu) = -\frac{k_{\mathrm{B}}VT^{c_p}}{\Phi e^{c_p-\frac{\mu}{k_{\mathrm{B}}T}}},$$

with $c_p = \frac{5}{2}$ and $\Phi = \left(\frac{2\pi\hbar^2}{mk_{\mathrm{B}}e^{\frac{5}{3}}}\right)^{\frac{3}{2}}$. The thermodynamics section then shows that the corresponding Helmholtz free energy is the same as we obtained above from the canonical ensemble. Note that we also obtained the same Φ as before.

CHAPTER

SEVEN

FLUID DYNAMICS

7.1 Fluid Dynamics

7.1.1 Stress-Energy Tensor

In general, the stress energy tensor is the flux of momentum p^μ over the surface x^ν. It is a machine that contains a knowledge of the energy density, momentum density and stress as measured by any observer of the event.

Imagine a (small) box in the spacetime. Then the observer with a 4-velocity u^μ measures the density of 4-momentum $\frac{\mathrm{d}p^\alpha}{\mathrm{d}V}$ in his frame as:

$$\frac{\mathrm{d}p^\alpha}{\mathrm{d}V} = -T^\alpha{}_\beta u^\beta$$

and the energy density that he measures is:

$$\rho = \frac{E}{V} = -\frac{u^\alpha p_\alpha}{V} = -u^\alpha \frac{\mathrm{d}p_\alpha}{\mathrm{d}V} = u^\alpha T_{\alpha\beta} u^\beta$$

One can also obtain the stress energy tensor from the Lagrangian $\mathcal{L} = \mathcal{L}(\eta_\rho, \partial_\nu \eta_\rho, x^\nu)$ by combining the Euler-Lagrange equations

$$\frac{\partial \mathcal{L}}{\partial \eta_\rho} - \partial_\nu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\nu \eta_\rho)} \right) = 0$$

with the total derivative $\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}x^\mu}$:

$$\begin{aligned} \frac{\mathrm{d}\mathcal{L}}{\mathrm{d}x^\mu} &= \frac{\partial \mathcal{L}}{\partial \eta_\rho} \partial_\mu \eta_\rho + \frac{\partial \mathcal{L}}{\partial(\partial_\nu \eta_\rho)} \partial_\mu \partial_\nu \eta_\rho + \partial_\mu \mathcal{L} = \\ &= \partial_\nu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\nu \eta_\rho)} \right) \partial_\mu \eta_\rho + \frac{\partial \mathcal{L}}{\partial(\partial_\nu \eta_\rho)} \partial_\nu \partial_\mu \eta_\rho + \partial_\mu \mathcal{L} = \\ &= \partial_\nu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\nu \eta_\rho)} \partial_\mu \eta_\rho \right) + \partial_\mu \mathcal{L} \end{aligned}$$

or

$$\partial_\nu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\nu \eta_\rho)} \partial_\mu \eta_\rho - \mathcal{L} \delta_\mu{}^\nu \right) + \partial_\mu \mathcal{L} = 0$$

This can be written as:

$$\partial_\nu T_\mu{}^\nu + f_\mu = 0$$

where

$$T_\mu{}^\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\nu \eta_\rho)} \partial_\mu \eta_\rho - \mathcal{L}\delta_\mu{}^\nu$$
$$f_\mu = \partial_\mu \mathcal{L}$$

The Navier-Stokes equations can be derived from the conservation law:

$$\partial_\nu T^{\mu\nu} + f^\mu = 0$$

To obtain some Lagrangian (and action) for the perfect fluid, so that we can derive the stress energy tensor $T^{\mu\nu}$ from that, is not trivial, see for example arXiv:gr-qc/9304026. One has to take into account the equation of state and incorporate the particle number conservation $\nabla_\mu(nu^\mu) = 0$ and no entropy exchange $\nabla_\mu(nsu^\mu) = 0$ constraints.

The equation of continuity follows from the conservation of the baryon number — the volume V that contains certain number of baryons can change, but the total number of baryons nV must remain constant:

$$\frac{\mathrm{d}(nV)}{\mathrm{d}\tau} = 0$$
$$\frac{\mathrm{d}n}{\mathrm{d}\tau} V + n\frac{\mathrm{d}V}{\mathrm{d}\tau} = 0$$
$$u^\alpha(\partial_\alpha n)V + n(\partial_\alpha u^\alpha)V = 0$$
$$\partial_\alpha(nu^\alpha) = 0$$

Perfect Fluids

Perfect fluids have no heat conduction ($T^{i0} = T^{0i} = 0$) and no viscosity ($T^{ij} = p\mathbb{1}$), so in the comoving frame:

$$T^{\alpha\beta} = \mathrm{diag}(\rho c^2, p, p, p) = \left(\rho + \frac{p}{c^2}\right) u^\alpha u^\beta + p g^{\alpha\beta}$$

where in the comoving frame we have $g^{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$, $u^0 = c$ and $u^i = 0$, but $\partial_\alpha U^i \neq 0$. p is the pressure with units $[p] = \mathrm{N\,m^{-2}} = \mathrm{kg\,m^{-1}\,s^{-2}}$ (then $[\frac{p}{c^2}] = \mathrm{kg\,m^{-3}}$), ρ is the rest mass density with units $[\rho] = \mathrm{kg\,m^{-3}}$, and ρc^2 is the energy density with units $[\rho c^2] = \mathrm{kg\,m^{-1}\,s^{-2}}$.

The last equation is a tensor equation so it holds in any frame. Let's write the components explicitly:

$$T^{00} = \left(\rho + \frac{p}{c^2}\right) u^0 u^0 - p = \left(\rho + \frac{p}{c^2}\right) c^2\gamma^2 - p = \left(\rho c^2 + p\left(1 - \frac{1}{\gamma^2}\right)\right)\gamma^2 = \left(\rho c^2 + p\frac{v^2}{c^2}\right)\gamma^2$$
$$T^{0i} = T^{i0} = \left(\rho + \frac{p}{c^2}\right) u^0 u^i = \left(\rho + \frac{p}{c^2}\right) cv^i\gamma^2 = \frac{1}{c}\left(\rho c^2 + p\right) v^i\gamma^2$$
$$T^{ij} = \left(\rho + \frac{p}{c^2}\right) u^i u^j + p\delta^{ij} = \left(\rho + \frac{p}{c^2}\right) v^i v^j \gamma^2 + p\delta^{ij}$$

We now use the conservation of the stress energy tensor and the conservation of the number of particles:

$$\partial_\nu T^{\mu\nu} = 0 \tag{7.1.1.1}$$

$$\partial_\mu(nu^\mu) = 0 \tag{7.1.1.2}$$

The equation (7.1.1.2) gives:

$$\partial_t(n\gamma) + \partial_i(nv^i\gamma) = 0$$

$$\partial_t(nm\gamma) + \partial_i(nmv^i\gamma) = 0 \tag{7.1.1.3}$$

$$\partial_t(nmc^2\gamma) + \partial_i(nmc^2v^i\gamma) = 0 \tag{7.1.1.4}$$

The equation (7.1.1.1) gives for $\mu = 0$:

$$\begin{aligned}
\partial_\nu T^{0\nu} &= 0 \\
\partial_0 T^{00} + \partial_i T^{0i} &= 0 \\
\partial_t \left(\frac{1}{c} \left(\rho c^2 + p \frac{v^2}{c^2} \right) \gamma^2 \right) + \partial_i \left(\frac{1}{c} \left(\rho c^2 + p \right) v^i \gamma^2 \right) &= 0 \\
\partial_t \left(\left(\rho c^2 + p \frac{v^2}{c^2} \right) \gamma^2 \right) + \partial_i \left(\left(\rho c^2 + p \right) v^i \gamma^2 \right) &= 0 \qquad (7.1.1.5)
\end{aligned}$$

We now substract the equation (7.1.1.4) from (7.1.1.5):

$$\partial_t \left(\left(\rho c^2 \gamma - nmc^2 + p \frac{v^2}{c^2} \gamma \right) \gamma \right) + \partial_i \left(\left(\rho c^2 \gamma - nmc^2 + p\gamma \right) v^i \gamma \right) = 0$$

We define the nonrelativistic energy as:

$$E = \rho c^2 \gamma - nmc^2 = \tfrac{1}{2} \rho v^2 + (\rho - nm) c^2 + O\left(\frac{v^4}{c^2} \right)$$

so it contains the kinetic plus internal energies. We substitute back into (7.1.1.5):

$$\partial_t \left(\left(E + p \frac{v^2}{c^2} \gamma \right) \gamma \right) + \partial_i \left(\left(E + p\gamma \right) v^i \gamma \right) = 0 \qquad (7.1.1.6)$$

This is the relativistic equation for the energy. Substituting $nm = \rho\gamma - \frac{E}{c^2}$ into (7.1.1.3):

$$\partial_t \left(\rho\gamma^2 - \frac{E\gamma}{c^2} \right) + \partial_i \left(\left(\rho\gamma^2 - \frac{E\gamma}{c^2} \right) v^i \right) = 0 \qquad (7.1.1.7)$$

The equation (7.1.1.1) for $\mu = i$ gives:

$$\begin{aligned}
\partial_\nu T^{i\nu} &= 0 \\
\partial_0 T^{i0} + \partial_j T^{ij} &= 0 \\
\partial_t \left(\frac{1}{c^2} \left(\rho c^2 + p \right) v^i \gamma^2 \right) + \partial_j \left(\left(\rho + \frac{p}{c^2} \right) v^i v^j \gamma^2 + p\delta^{ij} \right) &= 0 \\
\partial_t \left(\left(\rho + \frac{p}{c^2} \right) v^i \gamma^2 \right) + \partial_j \left(\left(\rho + \frac{p}{c^2} \right) v^i v^j \gamma^2 + p\delta^{ij} \right) &= 0 \qquad (7.1.1.8)
\end{aligned}$$

This is the momentum equation. The equations (7.1.1.7), (7.1.1.8) and (7.1.1.6) are the correct relativistic equations for the perfect fluid (no approximations were done). We can take either (7.1.1.7) or (7.1.1.5) as the equation of continuity (both give the same nonrelativistic equation of continuity). Their Newtonian limit is obtained by $c \to \infty$ (which implies $\gamma \to 1$):

$$\begin{aligned}
\partial_t \rho + \partial_i (\rho v^i) &= 0 \\
\partial_t \left(\rho v^i \right) + \partial_j \left(\rho v^i v^j + p\delta^{ij} \right) &= 0 \\
\partial_t E + \partial_j \left(v^j \left(E + p \right) \right) &= 0
\end{aligned}$$

those are the Euler equations, also sometimes written as:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \qquad (7.1.1.9)$$

$$\frac{\partial (\rho \mathbf{v})}{\partial t} + \nabla \cdot (\rho \mathbf{v}\mathbf{v}^T) + \nabla p = 0 \qquad (7.1.1.10)$$

$$\frac{\partial E}{\partial t} + \nabla \cdot \left(\mathbf{v} \left(E + p \right) \right) = 0 \qquad (7.1.1.11)$$

The momentum equation can be further simplified by expanding the parentheses and using the continuity equation:

$$\begin{aligned}\frac{\partial(\rho\mathbf{v})}{\partial t} + \nabla\cdot(\rho\mathbf{v}\mathbf{v}^T) + \nabla p = 0\\ \underbrace{\left(\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v})\right)}_{0}\mathbf{v} + \rho\left(\frac{\partial\mathbf{v}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{v}\right) + \nabla p = 0\\ \rho\left(\frac{\partial\mathbf{v}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{v}\right) + \nabla p = 0\end{aligned} \tag{7.1.1.12}$$

Where we used:

$$\left[\nabla\cdot(\rho\mathbf{v}\mathbf{v}^T)\right]^i = \partial_j(\rho v^i v^j) = v^i\partial_j(\rho v^j) + \rho v^j\partial_j v^i = \left[\mathbf{v}\nabla\cdot(\rho\mathbf{v}) + \rho\mathbf{v}\cdot\nabla\mathbf{v}\right]^i$$

Alternative Derivation

We can also take the non-relativistic limit in the stress energy tensor:

$$\begin{aligned}T^{00} &\to \rho c^2\\ T^{0i} = T^{i0} &\to \frac{1}{c}\rho c^2 v^i\\ T^{ij} &\to \rho v^i v^j + p\delta^{ij}\end{aligned}$$

and plug it into the equation (7.1.1.1). For $\mu = 0$ we get the equation of continuity:

$$\begin{aligned}\partial_\nu T^{0\nu} &= 0\\ \partial_0 T^{00} + \partial_i T^{0i} &= 0\\ \partial_t\left(\frac{1}{c}\rho c^2\right) + \partial_i\left(\frac{1}{c}\rho c^2 v^i\right) &= 0\\ \partial_t\rho + \partial_i\left(\rho v^i\right) &= 0\end{aligned}$$

and for $\mu = i$ we get the momentum equation:

$$\begin{aligned}\partial_\nu T^{i\nu} &= 0\\ \partial_0 T^{i0} + \partial_j T^{ij} &= 0\\ \partial_t\left(\frac{1}{c^2}\rho c^2 v^i\right) + \partial_j\left(\rho v^i v^j + p\delta^{ij}\right) &= 0\\ \partial_t\left(\rho v^i\right) + \partial_j\left(\rho v^i v^j + p\delta^{ij}\right) &= 0\end{aligned}$$

However, in order to derive the equation for energy E, one needs to take into account the full relativistic stress energy tensor, see the previous section for details.

Energy Equation

The energy equation can also be derived from thermodynamic and the other two Euler equations. We have the following two Euler equations:

$$\begin{aligned}\partial_t\rho + \partial_i(\rho u^i) &= 0\\ \rho\partial_t u^i + \rho u^j\partial_j u^i + \delta^{ij}\partial_j p &= 0\end{aligned}$$

We'll need the following formulas:

$$\begin{aligned}
\partial_t(u_i u^i) = (\partial_t u_i)u^i + u_i\partial_t u^i = (\partial_t u_i)\delta^{ij}u_j + u_i\partial_t u^i = \\
= (\partial_t u_i\delta^{ij})u_j + u_i\partial_t u^i = (\partial_t u^j)u_j + u_i\partial_t u^i = 2u_i\partial_t u^i \\
\partial_j(u_i u^i) = 2u_i\partial_j u^i \\
\partial_t\rho = -\partial_i(\rho u^i) \\
\partial_t u^i = -u^j\partial_j u^i - \frac{\delta^{ij}}{\rho}\partial_j p \\
-u^j\partial_j p + \partial_t(\rho U) = \\
= -\frac{\mathrm{d}p}{\mathrm{d}t} + \partial_t p + \partial_t(\rho U) = \\
= -\frac{\mathrm{d}p}{\mathrm{d}t} + \partial_t(\rho U + p) = \\
= -\frac{\mathrm{d}p}{\mathrm{d}t} + \frac{\mathrm{d}}{\mathrm{d}t}(\rho U + p) - u^j\partial_j(\rho U + p) = \\
= -\frac{\mathrm{d}p}{\mathrm{d}t} + \frac{\mathrm{d}\rho}{\mathrm{d}t}\left(U + \frac{p}{\rho}\right) + \rho\frac{\mathrm{d}}{\mathrm{d}t}\left(U + \frac{p}{\rho}\right) - u^j\partial_j(\rho U + p) = \\
= -\frac{\mathrm{d}p}{\mathrm{d}t} + \frac{\mathrm{d}\rho}{\mathrm{d}t}\left(U + \frac{p}{\rho}\right) + \rho\frac{\mathrm{d}}{\mathrm{d}t}\left(U + \frac{p}{\rho}\right) + (\rho U + p)\partial_j u^j - \partial_j(\rho U u^j + p u^j) = \\
= \left[\rho\frac{\mathrm{d}}{\mathrm{d}t}\left(U + \frac{p}{\rho}\right) - \frac{\mathrm{d}p}{\mathrm{d}t}\right] + \left(U + \frac{p}{\rho}\right)\left[\frac{\mathrm{d}\rho}{\mathrm{d}t} + \rho\partial_j u^j\right] - \partial_j(\rho U u^j + p u^j) = \\
= -\partial_j(\rho U u^j + p u^j) \\
0 = \mathrm{d}Q = T\mathrm{d}S = \mathrm{d}U + p\mathrm{d}V = \mathrm{d}(U + pV) - V\mathrm{d}p = \mathrm{d}\left(U + \frac{p}{\rho}\right) - \frac{1}{\rho}\mathrm{d}p = \mathrm{d}H - \frac{1}{\rho}\mathrm{d}p
\end{aligned}$$

where $V = \frac{1}{\rho}$ is the specific volume and $H = U + \frac{p}{\rho}$ is entalphy (heat content).

Then:

$$\begin{aligned}
\partial_t E = \\
= \partial_t(\tfrac{1}{2}\rho u_i u^i + \rho U) = \\
= \tfrac{1}{2}u_i u^i\partial_t\rho + \tfrac{1}{2}\rho\partial_t(u_i u^i) + \partial_t(\rho U) = \\
= -\tfrac{1}{2}u_i u^i\partial_j(\rho u^j) + \rho u_i\partial_t u^i + \partial_t(\rho U) = \\
= -\tfrac{1}{2}u_i u^i\partial_j(\rho u^j) - \rho u_i u^j\partial_j u^i - u_i\delta^{ij}\partial_j p + \partial_t(\rho U) = \\
= -\tfrac{1}{2}u_i u^i\partial_j(\rho u^j) - \tfrac{1}{2}\rho u^j\partial_j(u_i u^i) - u_i\delta^{ij}\partial_j p + \partial_t(\rho U) = \\
= -\tfrac{1}{2}\partial_j(\rho u_i u^i u^j) - u^j\partial_j p + \partial_t(\rho U) = \\
= -\tfrac{1}{2}\partial_j(\rho u_i u^i u^j) - \partial_j(\rho U u^j + p u^j) = \\
= -\partial_j\left(u^j\left(\tfrac{1}{2}\rho u_i u^i + \rho U + p\right)\right) = \\
= -\partial_j\left(u^j\left(E + p\right)\right)
\end{aligned}$$

so:

$$\begin{aligned}
\partial_t E + \partial_j\left(u^j\left(E + p\right)\right) = 0 \\
\frac{\partial E}{\partial t} + \nabla\cdot\left(\mathbf{u}\left(E + p\right)\right) = 0
\end{aligned}$$

7.1.2 Navier-Stokes Equations

We start with the following nonrelativistic components of the stress energy tensor:

$$\begin{aligned} T^{00} &\to \rho c^2 \\ T^{0i} = T^{i0} &\to \frac{1}{c}\rho c^2 v^i \\ T^{ij} &\to \rho v^i v^j - \sigma^{ij} \end{aligned}$$

where $\sigma^{ij} = -p\delta^{ij} + \mathbb{T}$ (more below) and plug it into the equation (7.1.1.1). For $\mu = 0$ we get the equation of continuity as for perfect fluids:

$$\begin{aligned} \partial_\nu T^{0\nu} &= 0 \\ \partial_0 T^{00} + \partial_i T^{0i} &= 0 \\ \partial_t \left(\frac{1}{c}\rho c^2\right) + \partial_i \left(\frac{1}{c}\rho c^2 v^i\right) &= 0 \\ \partial_t \rho + \partial_i \left(\rho v^i\right) &= 0 \end{aligned}$$

and for $\mu = i$ we get the momentum equation:

$$\begin{aligned} \partial_\nu T^{i\nu} &= f^i \\ \partial_0 T^{i0} + \partial_j T^{ij} &= f^i \\ \partial_t \left(\frac{1}{c^2}\rho c^2 v^i\right) + \partial_j \left(\rho v^i v^j - \sigma^{ij}\right) &= f^i \\ \partial_t \left(\rho v^i\right) + \partial_j \left(\rho v^i v^j - \sigma^{ij}\right) &= f^i \end{aligned}$$

By using the continuity equation in the momentum equation (as in perfect fluids), we get:

$$\rho \left(\partial_t v^i + v^i \partial_j v^j\right) - \partial_j \sigma^{ij} = f^i$$

This is sometimes called the Cauchy momentum equation:

$$\rho \left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v}\right) = \nabla \cdot \sigma + \mathbf{f}$$

where the stress tensor σ can be written as:

$$\sigma = -p\mathbb{1} + \mathbb{T}$$

and we get the Navier-Stokes equations:

$$\rho \left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v}\right) = -\nabla p + \nabla \cdot \mathbb{T} + \mathbf{f}$$

Those are the most general equations. If we assume some more things about the fluid, they can be further simplified.

For Newtonian fluids, we want $\mathbb{T}$ to be isotropic, linear in strain rates and it's divergence zero for fluid at rest. It follows that the only way to write the tensor under these conditions is:

$$T_{ij} = 2\mu\epsilon_{ij} + \delta_{ij}\lambda \nabla \cdot \mathbf{v}$$

where the strain rate is:

$$\epsilon_{ij} = \frac{1}{2}\left(\partial_j v_i + \partial_i v_j\right)$$

The trace of $\mathbb{T}$ is:

$$\operatorname{Tr}\mathbb{T} = T_{ii} = 2\mu\epsilon_{ii} + \delta_{ii}\lambda\nabla\cdot\mathbf{v} = (2\mu + 3\lambda)\nabla\cdot\mathbf{v}$$

Note that $\mathbb{T}$ has zero trace, which is automatically satisfied for incompressible flow ($\nabla\cdot\mathbf{v} = 0$), but for compressible flow this imposes:

$$\lambda = -\frac{2}{3}\mu$$

The divergence of the tensor is:

$$\partial_j T_{ij} = 2\mu\partial_j\epsilon_{ij} + \partial_j\delta_{ij}\lambda\nabla\cdot\mathbf{v} = \mu\partial_j\partial_j v_i + \mu\partial_i\nabla\cdot\mathbf{v} + \lambda\partial_i\nabla\cdot\mathbf{v} = \mu\partial_j\partial_j v_i + (\mu+\lambda)\partial_i\nabla\cdot\mathbf{v}$$

or in vector form (these are usually called the compressible Navier-Stokes equations):

$$\nabla\cdot\mathbb{T} = \mu\nabla^2\mathbf{v} + (\mu+\lambda)\nabla\nabla\cdot\mathbf{v}$$

For incompressible fluid we have $\nabla\cdot\mathbf{v} = \mathbf{0}$, so we get the incompressible Navier-Stokes equations:

$$\nabla\cdot\mathbb{T} = \mu\nabla^2\mathbf{v}$$

and for a perfect fluid we have no viscosity, e.g. $\mu = 0$, then we get the Euler equations (for perfect fluid):

$$\nabla\cdot\mathbb{T} = 0$$

7.1.3 Incompressible Equations

Incompressible flow means that the material derivative of density is zero:

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \frac{\partial\rho}{\partial t} + \mathbf{v}\cdot\nabla\rho = 0\,. \tag{7.1.3.1}$$

Putting this into the equation of continuity (7.1.1.9) one obtains $\rho\nabla\cdot\mathbf{v} = 0$ or equivalently:

$$\nabla\cdot\mathbf{v} = 0\,. \tag{7.1.3.2}$$

But also (7.1.3.2) implies (7.1.3.1), so these two equations are equivalent: the divergence of the velocity field is zero if and only if the material derivative of the density is zero.

Using the condition $\nabla\cdot\mathbf{v} = 0$ in (7.1.1.9) and (7.1.1.12) we obtain:

$$\begin{aligned}
\nabla\cdot\mathbf{v} &= 0\,,\\
\frac{\partial\rho}{\partial t} + \mathbf{v}\cdot\nabla\rho &= 0\,,\\
\rho\left(\frac{\partial\mathbf{v}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{v}\right) + \nabla p &= \mu\nabla^2\mathbf{v}\,.
\end{aligned}$$

In addition to incompressibility, we can also assume a constant density $\rho(x,y,z) = \rho_0$, then we obtain the incompressible Navier-Stokes equations:

$$\nabla\cdot\mathbf{v} = 0\,, \tag{7.1.3.3}$$

$$\rho_0\left(\frac{\partial\mathbf{v}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{v}\right) + \nabla p = \mu\nabla^2\mathbf{v}\,. \tag{7.1.3.4}$$

For $\mu = 0$ they become the incompressible Euler equations. At the given time step with known $\mathbf{v}$ and p, the equation (7.1.3.4) is solved for $\mathbf{v}$ at the new time step. Then we solve for new p as follows. Apply divergence to (7.1.3.4):

$$\rho_0 \nabla \cdot \left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} \right) + \nabla \cdot \nabla p = \mu \nabla \cdot \nabla^2 \mathbf{v} \,,$$
$$\rho_0 \left(\frac{\partial (\nabla \cdot \mathbf{v})}{\partial t} + \nabla \cdot (\mathbf{v} \cdot \nabla \mathbf{v}) \right) + \nabla^2 p = \mu \nabla \cdot \nabla^2 \mathbf{v} \,,$$

now we use the following identities:

$$\nabla \cdot (\mathbf{v} \cdot \nabla \mathbf{v}) = \partial_i (v^j \partial_j v^i) = (\partial_i v^j)(\partial_j v^i) + v^j \partial_j \partial_i v^i = \operatorname{Tr}(\nabla \mathbf{v})^2 + \mathbf{v} \cdot \nabla (\nabla \cdot \mathbf{v}) \,,$$
$$\nabla \cdot \nabla^2 \mathbf{v} = \partial_i \partial^j \partial_j v^i = \partial^j \partial_j \partial_i v^i = \nabla^2 (\nabla \cdot \mathbf{v}) \,,$$

to get:

$$\rho_0 \left(\frac{\partial (\nabla \cdot \mathbf{v})}{\partial t} + \operatorname{Tr}(\nabla \mathbf{v})^2 + \mathbf{v} \cdot \nabla (\nabla \cdot \mathbf{v}) \right) + \nabla^2 p = \mu \nabla^2 (\nabla \cdot \mathbf{v}) \,.$$

Finally we use the equation (7.1.3.3) to simplify:

$$-\nabla^2 p = \rho_0 \operatorname{Tr}(\nabla \mathbf{v})^2 \,, \tag{7.1.3.5}$$

which is a Poisson equation for p. Note again that $\operatorname{Tr}(\nabla \mathbf{v})^2 = (\partial_i v^j)(\partial_j v^i)$. The equation (7.1.3.5) is then used to solve for p at the new time step.

Divergence Free Velocity

Typically by propagating (7.1.3.4), we obtain a velocity $\mathbf{v}^*$ that is not divergence free. To make it so, we want to find such a divergence free $\mathbf{v}$ that is closest to $\mathbf{v}^*$ in the L^2 norm $\|\mathbf{v} - \mathbf{v}^*\| = \sqrt{\int (\mathbf{v} - \mathbf{v}^*)^2 \,\mathrm{d}^3 x}$, in other words we want to find the L^2 projection onto the divergence free subspace, so we have to minimize the following functional:

$$R[\mathbf{v}, \lambda] = \tfrac{1}{2} \|\mathbf{v} - \mathbf{v}^*\|^2 - \int \lambda \nabla \cdot \mathbf{v} \,\mathrm{d}^3 x = \int \tfrac{1}{2} (\mathbf{v} - \mathbf{v}^*)^2 - \lambda \nabla \cdot \mathbf{v} \,\mathrm{d}^3 x \,,$$

where we used a Langrange multiplier $\lambda = \lambda(\mathbf{x})$ in the second term to impose the zero divergence on $\mathbf{v} = \mathbf{v}(\mathbf{x})$ for all points $\mathbf{x}$ (that is why λ is a function of $\mathbf{x}$ and not a constant) and in the first term we ensure that $\mathbf{v}$ is as close as possible to the original field $\mathbf{v}^*$ in the L^2 sense. Let's calculate the variation:

$$\delta R[\mathbf{v}, \lambda] = \int (\mathbf{v} - \mathbf{v}^*) \cdot \delta \mathbf{v} - \lambda \nabla \cdot \delta \mathbf{v} - (\nabla \cdot \mathbf{v}) \delta \lambda \,\mathrm{d}^3 x =$$
$$= \int (\mathbf{v} - \mathbf{v}^*) \cdot \delta \mathbf{v} + (\nabla \lambda) \cdot \delta \mathbf{v} - (\nabla \cdot \mathbf{v}) \delta \lambda \,\mathrm{d}^3 x + \int \lambda \delta \mathbf{v} \cdot \mathbf{n} \,\mathrm{d}S =$$
$$= \int (\mathbf{v} - \mathbf{v}^* + \nabla \lambda) \cdot \delta \mathbf{v} - (\nabla \cdot \mathbf{v}) \delta \lambda \,\mathrm{d}^3 x + \int \lambda \delta \mathbf{v} \cdot \mathbf{n} \,\mathrm{d}S \,.$$

From the condition $\delta R[\mathbf{v}, \lambda] = 0$ and assuming the surface integral vanishes (i.e. either $\lambda = 0$ or $\delta \mathbf{v} \cdot \mathbf{n} = 0$ everywhere on the boundary) we obtain the two Euler-Lagrange equations:

$$\frac{\delta R[\mathbf{v}, \lambda]}{\delta \mathbf{v}} = \mathbf{v} - \mathbf{v}^* + \nabla \lambda = 0 \,, \tag{7.1.3.6}$$

$$\frac{\delta R[\mathbf{v}, \lambda]}{\delta \lambda} = -\nabla \cdot \mathbf{v} = 0 \,. \tag{7.1.3.7}$$

Applying divergence to (7.1.3.6) and using (7.1.3.7) we obtain:

$$\nabla^2 \lambda = \nabla \cdot \mathbf{v}^* \,. \tag{7.1.3.8}$$

After solving this Poisson equation for λ we can calculate the divergence free $\mathbf{v}$ from (7.1.3.6):

$$\mathbf{v} = \mathbf{v}^* - \nabla \lambda \,. \tag{7.1.3.9}$$

Time Discretization

The incompressible Euler equations are:

$$\nabla \cdot \mathbf{v} = 0\,,$$
$$\frac{\partial \mathbf{v}}{\partial t} = -\mathbf{v}\cdot\nabla\mathbf{v} - \frac{1}{\rho_0}\nabla p\,. \tag{7.1.3.10}$$

We use first order time discretization:

$$\nabla \cdot \mathbf{v}^n = 0\,, \tag{7.1.3.11}$$
$$\nabla \cdot \mathbf{v}^{n+1} = 0\,, \tag{7.1.3.12}$$
$$\frac{\mathbf{v}^{n+1}-\mathbf{v}^n}{\Delta t} = -\mathbf{v}^n\cdot\nabla\mathbf{v}^n - \frac{1}{\rho_0}\nabla p^{n+1}\,. \tag{7.1.3.13}$$

The velocity at time steps n and $n+1$ must be divergence free, per (7.1.3.11) and (7.1.3.12). The simplest discretization of (7.1.3.10) is to use an explicit scheme, so we evaluate the term $\mathbf{v}\cdot\nabla\mathbf{v}$ at the time step n. Regarding the pressure term ∇p, if we evaluated it at the time step n, then from (7.1.3.13) we could calculate $\mathbf{v}^{n+1}$ that would not be divergence free, per (7.1.3.12). So we are led to evaluate the pressure term at the time step $n+1$, then all the equations (7.1.3.11), (7.1.3.12) and (7.1.3.13) can be satisfied.

To solve this system of equations, we use an operator splitting on (7.1.3.13), the most natural is probably the following:

$$\frac{\mathbf{v}^{*}-\mathbf{v}^n}{\Delta t} = -\mathbf{v}^n\cdot\nabla\mathbf{v}^n - \frac{1}{\rho_0}\nabla p^{n}\,, \tag{7.1.3.14}$$
$$\frac{\mathbf{v}^{n+1}-\mathbf{v}^*}{\Delta t} = -\frac{1}{\rho_0}\nabla (p^{n+1}-p^n)\,. \tag{7.1.3.15}$$

The first equation (7.1.3.14) is just like (7.1.3.13), except that the pressure term is evaluated at the time step n, which forces us to change $\mathbf{v}^{n+1}$ into $\mathbf{v}^*$, which is not divergence free. The second equation (7.1.3.15) is then uniquely given by the condition that the sum of (7.1.3.14) and (7.1.3.15) is equal to (7.1.3.13).

The equation (7.1.3.15) is equivalent to (7.1.3.9), with $\lambda = \frac{\Delta t}{\rho_0}(p^{n+1}-p^n)$, so this is an L^2 projection of $\mathbf{v}^*$ onto the divergence free subspace to obtain $\mathbf{v}^{n+1}$, also sometimes called a pressure projection. We use the same method as was used to obtain the Poisson equation (7.1.3.8) for λ, i.e. take a divergence and rearrange:

$$\nabla^2\lambda = \nabla^2\left(\frac{\Delta t}{\rho_0}(p^{n+1}-p^n)\right) = \nabla\cdot\mathbf{v}^*\,. \tag{7.1.3.16}$$

One solves (7.1.3.14) for $\mathbf{v}^*$, then the Poisson equation (7.1.3.16) for λ (i.e. the pressure update $p^{n+1}-p^n$), and then one computes $\mathbf{v}^{n+1}$ using (7.1.3.15) (or equivalently (7.1.3.9)).

These equations are derived from Euler equations (7.1.3.11) and (7.1.3.12) using a time discretization and an operator splitting technique. The theory of the L^2 projection onto the divergence free subspace is not needed to derive these equations, but it helps with understanding of what is going on.

Note 1: the operator splitting of (7.1.3.13) into (7.1.3.14) and (7.1.3.15) is not unique. Another option is:

$$\frac{\mathbf{v}^{*}-\mathbf{v}^n}{\Delta t} = -\mathbf{v}^n\cdot\nabla\mathbf{v}^n\,, \tag{7.1.3.17}$$
$$\frac{\mathbf{v}^{n+1}-\mathbf{v}^*}{\Delta t} = -\frac{1}{\rho_0}\nabla p^{n+1}\,. \tag{7.1.3.18}$$

The sum of (7.1.3.17) and (7.1.3.18) is still (7.1.3.13) and the equation (7.1.3.18) is still equivalent to (7.1.3.9), only this time with $\lambda = \frac{\Delta t}{\rho_0}p^{n+1}$. The Poisson equation then becomes:

$$\nabla^2\lambda = \nabla^2\left(\frac{\Delta t}{\rho_0}p^{n+1}\right) = \nabla\cdot\mathbf{v}^*\,. \tag{7.1.3.19}$$

The only difference to the previous scheme is that now the L^2 norm of $\|\mathbf{v}^{n+1} - \mathbf{v}^*\| = \|\nabla\lambda\|$ is larger, because λ now depends on the full pressure instead of the pressure difference, so $\mathbf{v}^*$ is not as close to $\mathbf{v}^{n+1}$ as in the previous scheme.

Note 2: By applying divergence to (7.1.3.17) we obtain:

$$\frac{\nabla\cdot\mathbf{v}^*}{\Delta t} = -\nabla\cdot(\mathbf{v}^n\cdot\nabla\mathbf{v}^n) = -\operatorname{Tr}(\nabla\mathbf{v}^n)^2\,,$$

and substituting into (7.1.3.19) we obtain:

$$\nabla^2\left(\frac{\Delta t}{\rho_0}p^{n+1}\right) = \nabla\cdot\mathbf{v}^* = -\Delta t\operatorname{Tr}(\nabla\mathbf{v}^n)^2\,,$$

or

$$-\nabla^2 p^{n+1} = \rho_0\operatorname{Tr}(\nabla\mathbf{v}^n)^2\,, \tag{7.1.3.20}$$

which is the discrete analog of the equation (7.1.3.5). The same result is obtained by applying a divergence to (7.1.3.14) and substituting into (7.1.3.16):

$$\nabla^2\left(\frac{\Delta t}{\rho_0}(p^{n+1}-p^n)\right) = \nabla\cdot\mathbf{v}^* = -\Delta t\operatorname{Tr}(\nabla\mathbf{v}^n)^2 - \frac{\Delta t}{\rho_0}\nabla^2 p^n$$

Which simplifies to (7.1.3.20).

7.1.4 Bernoulli's Principle

Bernoulli's principle works for a perfect fluid, so we take the Euler equations:

$$\rho\left(\frac{\partial\mathbf{v}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{v}\right) = -\nabla p + \mathbf{f}$$

and put it into a vertical gravitational field $\mathbf{f} = (0, 0, -\rho g) = -\rho g\nabla z$, so:

$$\rho\left(\frac{\partial\mathbf{v}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{v}\right) = -\nabla p - \rho g\nabla z\,,$$

we divide by ρ:

$$\frac{\partial\mathbf{v}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{v} = -\nabla\left(\frac{p}{\rho} + gz\right)$$

and use the identity $\mathbf{v}\cdot\nabla\mathbf{v} = \frac{1}{2}\nabla v^2 + (\nabla\times\mathbf{v})\times\mathbf{v}$:

$$\frac{\partial\mathbf{v}}{\partial t} + \frac{1}{2}\nabla v^2 + (\nabla\times\mathbf{v})\times\mathbf{v} + \nabla\left(\frac{p}{\rho} + gz\right) = 0\,,$$

so:

$$\frac{\partial\mathbf{v}}{\partial t} + (\nabla\times\mathbf{v})\times\mathbf{v} + \nabla\left(\frac{v^2}{2} + gz + \frac{p}{\rho}\right) = 0\,.$$

If the fluid is moving, we integrate this along a streamline from the point A to B:

$$\int\frac{\partial\mathbf{v}}{\partial t}\cdot\mathrm{d}\mathbf{l} + \left[\frac{v^2}{2} + gz + \frac{p}{\rho}\right]_A^B = 0\,.$$

So far we didn't do any approximation (besides having a perfect fluid in a vertical gravitation field). Now we assume a steady flow, so $\frac{\partial\mathbf{v}}{\partial t} = 0$ and since points A and B are arbitrary, we get:

$$\frac{v^2}{2} + gz + \frac{p}{\rho} = \text{const.}$$

along the streamline. This is called the Bernoulli's principle. If the fluid is not moving, we set $\mathbf{v} = 0$ in the equations above and immediately get:

$$gz + \frac{p}{\rho} = \text{const.}$$

The last equation then holds everywhere in the (nonmoving) fluid (as opposed to the previous equation that only holds along the streamline).

Hydrostatic Pressure

Let p_1 be the pressure on the water surface and p_2 the pressure h meters below the surface. From the Bernoulli's principle:

$$\frac{p_1}{\rho} = g \cdot (-h) + \frac{p_2}{\rho}$$

so

$$p_1 + h\rho g = p_2$$

and we can see, that the pressure h meters below the surface is $h\rho g$ plus the (atmospheric) pressure p_1 on the surface.

Torricelli's Law

We want to find the speed v of the water flowing out of the tank (of the height h) through a small hole at the bottom. The (atmospheric) pressure at the water surface and also near the small hole is p_1. From the Bernoulli's principle:

$$\frac{p_1}{\rho} = \frac{v^2}{2} + g \cdot (-h) + \frac{p_1}{\rho}$$

so:

$$v = \sqrt{2gh}$$

This is called the Torricelli's law.

Venturi Effect

A pipe with a cross section A_1, pressure p_1 and the speed of a perfect liquid v_1 changes it's cross section to A_2, so the pressure changes to p_2 and the speed to v_2. Given $\Delta p = p_1 - p_2$, A_1 and A_2, calculate v_1 and v_2.

We use the continuity equation:

$$A_1 v_1 = A_2 v_2$$

and the Bernoulli's principle:

$$\frac{v_1^2}{2} + \frac{p_1}{\rho} = \frac{v_2^2}{2} + \frac{p_2}{\rho}$$

so we have two equations for two unknowns v_1 and v_2, after solving it we get:

$$v_1 = A_2 \sqrt{\frac{2\Delta p}{\rho(A_1^2 - A_2^2)}}$$

$$v_2 = A_1 \sqrt{\frac{2\Delta p}{\rho(A_1^2 - A_2^2)}}$$

Hagen-Poiseuille Law

We assume incompressible (but viscuous) Newtonean fluid (in no external force field):

$$\rho\left(\frac{\partial \mathbf{v}}{\partial t}+\mathbf{v}\cdot\nabla\mathbf{v}\right)=-\nabla p+\mu\nabla^2\mathbf{v}$$

flowing in the vertical pipe of radius R and we further assume steady flow $\frac{\partial \mathbf{v}}{\partial t}=0$, axis symmetry $v_r=v_\theta=\partial_\theta(\cdots)=0$ and a fully developed flow $\partial_z v_z=0$. We write the Navier-Stokes equations above in the cylindrical coordinates and using the stated assumptions, the only nonzero equations are:

$$0=-\partial_r p$$

$$0=-\partial_z p+\mu\frac{1}{r}\partial_r(r\partial_r v_z)$$

from the first one we can see the $p=p(z)$ is a function of z only and we can solve the second one for $v_z=v_z(r)$:

$$v_z(r)=\frac{1}{4\mu}(\partial_z p)r^2+C_1\log r+C_2$$

We want $v_z(r=0)$ to be finite, so $C_1=0$, next we assume the no slip boundary conditions $v_z(r=R)=0$, so $C_2=-\frac{1}{4\mu}(\partial_z p)R^2$ and we get the parabolic velocity profile:

$$v_z(r)=\frac{1}{4\mu}(-\partial_z p)(R^2-r^2)$$

Assuming that the pressure decreases linearly across the length of the pipe, we have $-\partial_z p=\frac{\Delta P}{L}$ and we get:

$$v_z(r)=\frac{\Delta P}{4\mu L}(R^2-r^2)$$

We can now calculate the volumetric flow rate:

$$Q=\frac{\mathrm{d}V}{\mathrm{d}t}=\frac{\mathrm{d}}{\mathrm{d}t}\int z\,\mathrm{d}S=\int\frac{\mathrm{d}z}{\mathrm{d}t}\mathrm{d}S=\int v_z\,\mathrm{d}S=\int_0^{2\pi}\int_0^R v_z\,r\,\mathrm{d}r\,\mathrm{d}\phi=$$

$$=\frac{\Delta P\pi}{2\mu L}\int_0^R(R^2-r^2)r\,\mathrm{d}r=\frac{\Delta P\pi R^4}{8\mu L}$$

so we can see that it depends on the 4th power of R. This is called the Hagen-Poiseuille law.

7.2 MHD Equations

7.2.1 Introduction

The magnetohydrodynamics (MHD) equations are:

$$\frac{\partial\rho}{\partial t}+\nabla\cdot(\rho\mathbf{v})=0 \tag{7.2.1.1}$$

$$\rho\left(\frac{\partial \mathbf{v}}{\partial t}+(\mathbf{v}\cdot\nabla)\mathbf{v}\right)=-\nabla p+\frac{1}{\mu}(\nabla\times\mathbf{B})\times\mathbf{B}+\rho\mathbf{g} \tag{7.2.1.2}$$

$$\frac{\partial\mathbf{B}}{\partial t}=\nabla\times(\mathbf{v}\times\mathbf{B})+\eta\nabla^2\mathbf{B} \tag{7.2.1.3}$$

$$\nabla\cdot\mathbf{B}=0 \tag{7.2.1.4}$$

assuming η is constant. See the next section for a derivation. We can now apply the following identities (we use the fact that $\nabla \cdot \mathbf{B} = 0$):

$$[(\nabla \times \mathbf{B}) \times \mathbf{B}]_i = \varepsilon_{ijk}(\nabla \times \mathbf{B})_j B_k = \varepsilon_{ijk}\varepsilon_{jlm}(\partial_l B_m)B_k = (\delta_{kl}\delta_{im} - \delta_{km}\delta_{il})(\partial_l B_m)B_k =$$
$$= (\partial_k B_i)B_k - (\partial_i B_k)B_k = \left[(\mathbf{B}\cdot\nabla)\mathbf{B} - \frac{1}{2}\nabla|\mathbf{B}|^2\right]_i$$
$$(\nabla \times \mathbf{B}) \times \mathbf{B} = (\mathbf{B}\cdot\nabla)\mathbf{B} - \frac{1}{2}\nabla|\mathbf{B}|^2 = (\mathbf{B}\cdot\nabla)\mathbf{B} +$$
$$\nabla \times (\mathbf{v} \times \mathbf{B}) = (\mathbf{B}\cdot\nabla)\mathbf{v} - \mathbf{B}(\nabla\cdot\mathbf{v}) + \mathbf{v}(\nabla\cdot\mathbf{B}) -$$
$$\nabla \cdot (\rho\mathbf{v}\mathbf{v}^T) = (\nabla\cdot(\rho\mathbf{v}))\,\mathbf{v} + \rho(\mathbf{v}\cdot\nabla)\mathbf{v} = -\mathbf{v}\frac{\partial\rho}{\partial t}$$

So the MHD equations can alternatively be written as:

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v}) = 0 \tag{7.2.1.5}$$

$$\frac{\partial\rho\mathbf{v}}{\partial t} + \nabla\cdot(\rho\mathbf{v}\mathbf{v}^T) = -\nabla p + \frac{1}{\mu}\left(\nabla\cdot(\mathbf{B}\mathbf{B}^T) - \frac{1}{2}\nabla|\mathbf{B}|^2\right) + \rho\mathbf{g} \tag{7.2.1.6}$$

$$\frac{\partial\mathbf{B}}{\partial t} = \nabla\cdot(\mathbf{B}\mathbf{v}^T - \mathbf{v}\mathbf{B}^T) + \eta\nabla^2\mathbf{B} \tag{7.2.1.7}$$

$$\nabla\cdot\mathbf{B} = 0 \tag{7.2.1.8}$$

One can also introduce a new variable $p^* = p + \frac{1}{2}\nabla|\mathbf{B}|^2$, that simplifies (7.2.1.6) a bit.

7.2.2 Derivation

The above equations can easily be derived. We have the continuity equation:

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v}) = 0$$

Navier-Stokes equations (momentum equation) with the Lorentz force on the right-hand side:

$$\rho\left(\frac{\partial\mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v}\right) = -\nabla p + \mathbf{j}\times\mathbf{B} + \rho\mathbf{g}$$

where the current density $\mathbf{j}$ is given by the Maxwell equation (we neglect the displacement current $\frac{\partial\mathbf{E}}{\partial t}$):

$$\mathbf{j} = \frac{1}{\mu}\nabla\times\mathbf{B}$$

and the Lorentz force:

$$\frac{1}{\sigma}\mathbf{j} = \mathbf{E} + \mathbf{v}\times\mathbf{B}$$

from which we eliminate $\mathbf{E}$:

$$\mathbf{E} = -\mathbf{v}\times\mathbf{B} + \frac{1}{\sigma}\mathbf{j} = -\mathbf{v}\times\mathbf{B} + \frac{1}{\sigma\mu}\nabla\times\mathbf{B}$$

and put it into the Maxwell equation:

$$\frac{\partial\mathbf{B}}{\partial t} = -\nabla\times\mathbf{E}$$

so we get:

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) - \nabla \times \left(\frac{1}{\sigma\mu} \nabla \times \mathbf{B} \right)$$

assuming the magnetic diffusivity $\eta = \frac{1}{\sigma\mu}$ is constant, we get:

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) - \eta \nabla \times (\nabla \times \mathbf{B}) = \nabla \times (\mathbf{v} \times \mathbf{B}) + \eta \left(\nabla^2 \mathbf{B} - \nabla(\nabla \cdot \mathbf{B}) \right) = \nabla \times (\mathbf{v} \times \mathbf{B}) + \eta \nabla^2 \mathbf{B}$$

where we used the Maxwell equation:

$$\nabla \cdot \mathbf{B} = 0$$

7.2.3 Finite Element Formulation

We solve the following ideal MHD equations (we use $p^* = p + \frac{1}{2}\nabla|\mathbf{B}|^2$, but we drop the star):

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} - (\mathbf{B} \cdot \nabla)\mathbf{B} + \nabla p = 0 \tag{7.2.3.1}$$

$$\frac{\partial \mathbf{B}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{B} - (\mathbf{B} \cdot \nabla)\mathbf{u} = 0 \tag{7.2.3.2}$$

$$\nabla \cdot \mathbf{u} = 0 \tag{7.2.3.3}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{7.2.3.4}$$

If the equation (7.2.3.4) is satisfied initially, then it is satisfied all the time, as can be easily proved by applying a divergence to the Maxwell equation $\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E}$ (or the equation (7.2.3.2), resp. (7.2.1.3)) and we get $\frac{\partial}{\partial t}(\nabla \cdot \mathbf{B}) = 0$, so $\nabla \cdot \mathbf{B}$ is constant, independent of time. As a consequence, we are essentially only solving equations (7.2.3.1), (7.2.3.2) and (7.2.3.3), which consist of 5 equations for 5 unknowns (components of $\mathbf{u}$, p and $\mathbf{B}$).

We discretize in time by introducing a small time step τ and we also linearize the convective terms:

$$\frac{\mathbf{u}^n - \mathbf{u}^{n-1}}{\tau} + (\mathbf{u}^{n-1} \cdot \nabla)\mathbf{u}^n - (\mathbf{B}^{n-1} \cdot \nabla)\mathbf{B}^n + \nabla p = 0 \tag{7.2.3.5}$$

$$\frac{\mathbf{B}^n - \mathbf{B}^{n-1}}{\tau} + (\mathbf{u}^{n-1} \cdot \nabla)\mathbf{B}^n - (\mathbf{B}^{n-1} \cdot \nabla)\mathbf{u}^n = 0 \tag{7.2.3.6}$$

$$\nabla \cdot \mathbf{u}^n = 0 \tag{7.2.3.7}$$

Testing (7.2.3.5) by the test functions (v_1, v_2), (7.2.3.6) by the functions (C_1, C_2) and (7.2.3.7) by the test function q, we obtain the following weak formulation:

$$\begin{aligned} \int_\Omega \frac{u_1 v_1}{\tau} + (\mathbf{u}^{n-1} \cdot \nabla) u_1 v_1 - (\mathbf{B}^{n-1} \cdot \nabla) B_1 v_1 - p \frac{\partial v_1}{\partial x} \,\mathrm{d}\mathbf{x} &= \int_\Omega \frac{u_1^{n-1} v_1}{\tau} \,\mathrm{d}\mathbf{x} \\ \int_\Omega \frac{u_2 v_2}{\tau} + (\mathbf{u}^{n-1} \cdot \nabla) u_2 v_2 - (\mathbf{B}^{n-1} \cdot \nabla) B_2 v_2 - p \frac{\partial v_2}{\partial y} \,\mathrm{d}\mathbf{x} &= \int_\Omega \frac{u_2^{n-1} v_2}{\tau} \,\mathrm{d}\mathbf{x} \end{aligned} \tag{7.2.3.8}$$

$$\begin{aligned} \int_\Omega \frac{B_1 C_1}{\tau} + (\mathbf{u}^{n-1} \cdot \nabla) B_1 C_1 - (\mathbf{B}^{n-1} \cdot \nabla) u_1 C_1 \,\mathrm{d}\mathbf{x} &= \int_\Omega \frac{B_1^{n-1} C_1}{\tau} \,\mathrm{d}\mathbf{x} \\ \int_\Omega \frac{B_2 C_2}{\tau} + (\mathbf{u}^{n-1} \cdot \nabla) B_2 C_2 - (\mathbf{B}^{n-1} \cdot \nabla) u_2 C_2 \,\mathrm{d}\mathbf{x} &= \int_\Omega \frac{B_2^{n-1} C_2}{\tau} \,\mathrm{d}\mathbf{x} \end{aligned} \tag{7.2.3.9}$$

$$\int_\Omega \frac{\partial u_1}{\partial x} q + \frac{\partial u_2}{\partial y} q \, \mathrm{d}\mathbf{x} = 0 \tag{7.2.3.10}$$

To better understand the structure of these equations, we write it using bilinear and linear forms, as well as take into account the symmetries of the forms. Then we get a particularly simple structure:

$$\begin{array}{llllll}
+A(u_1, v_1) & & -X(p, v_1) & -B(B_1, v_1) & & = \; l_1(v_1) \\
& +A(u_2, v_2) & -Y(p, v_2) & & -B(B_2, v_2) & = \; l_2(v_2) \\
+X(q, u_1) & +Y(q, u_2) & & & & = \; 0 \\
-B(u_1, C_1) & & & +A(B_1, C_1) & & = \; l_4(C_1) \\
& -B(u_2, C_2) & & & +A(B_2, C_2) & = \; l_5(C_2)
\end{array}$$

where:

$$\begin{aligned}
A(u, v) &= \int_\Omega \frac{uv}{\tau} + (\mathbf{u}^{n-1} \cdot \nabla) uv \, \mathrm{d}\mathbf{x} \\
B(u, v) &= \int_\Omega (\mathbf{B}^{n-1} \cdot \nabla) uv \, \mathrm{d}\mathbf{x} \\
X(u, v) &= \int_\Omega u \frac{\partial v}{\partial x} \, \mathrm{d}\mathbf{x} \\
Y(u, v) &= \int_\Omega u \frac{\partial v}{\partial y} \, \mathrm{d}\mathbf{x} \\
l_1(v) &= \int_\Omega \frac{u_1^{n-1} v}{\tau} \, \mathrm{d}\mathbf{x} \\
l_2(v) &= \int_\Omega \frac{u_2^{n-1} v}{\tau} \, \mathrm{d}\mathbf{x} \\
l_4(v) &= \int_\Omega \frac{B_1^{n-1} v}{\tau} \, \mathrm{d}\mathbf{x} \\
l_5(v) &= \int_\Omega \frac{B_2^{n-1} v}{\tau} \, \mathrm{d}\mathbf{x}
\end{aligned}$$

E.g. there are only 4 distinct bilinear forms. Schematically we can visualize the structure by:

A		-X	-B	
	A	-Y		-B
X	Y			
-B			A	
	-B			A

In order to solve it with Hermes, we first need to write it in the block form:

$$\begin{array}{lllllllllll}
a_{11}(u_1, v_1) & + & a_{12}(u_2, v_1) & + & a_{13}(p, v_1) & + & a_{14}(B_1, v_1) & + & a_{15}(B_2, v_1) & = & l_1(v_1) \\
a_{21}(u_1, v_2) & + & a_{22}(u_2, v_2) & + & a_{23}(p, v_2) & + & a_{24}(B_1, v_2) & + & a_{25}(B_2, v_2) & = & l_2(v_2) \\
a_{31}(u_1, q) & + & a_{32}(u_2, q) & + & a_{33}(p, q) & + & a_{34}(B_1, q) & + & a_{35}(B_2, q) & = & l_3(q) \\
a_{41}(u_1, C_1) & + & a_{42}(u_2, C_1) & + & a_{43}(p, C_1) & + & a_{44}(B_1, C_1) & + & a_{45}(B_2, C_1) & = & l_4(C_1) \\
a_{51}(u_1, C_2) & + & a_{52}(u_2, C_2) & + & a_{53}(p, C_2) & + & a_{54}(B_1, C_2) & + & a_{55}(B_2, C_2) & = & l_5(C_2)
\end{array}$$

comparing to the above, we get the following nonzero forms:

$$\begin{array}{lllllllllll}
a_{11}(u_1, v_1) & + & 0 & + & a_{13}(p, v_1) & + & a_{14}(B_1, v_1) & + & 0 & = & l_1(v_1) \\
0 & + & a_{22}(u_2, v_2) & + & a_{23}(p, v_2) & + & 0 & + & a_{25}(B_2, v_2) & = & l_2(v_2) \\
a_{31}(u_1, q) & + & a_{32}(u_2, q) & + & 0 & + & 0 & + & 0 & = & 0 \\
a_{41}(u_1, C_1) & + & 0 & + & 0 & + & a_{44}(B_1, C_1) & + & 0 & = & l_4(C_1) \\
0 & + & a_{52}(u_2, C_2) & + & 0 & + & 0 & + & a_{55}(B_2, C_2) & = & l_5(C_2)
\end{array}$$

where:

$$\begin{aligned}
a_{11}(u_1, v_1) &= A(u_1, v_1) \\
a_{22}(u_2, v_2) &= A(u_2, v_2) \\
a_{44}(B_1, C_1) &= A(B_1, C_1) \\
a_{55}(B_2, C_1) &= A(B_2, C_2) \\
a_{13}(p, v_1) &= -X(p, v_1) \\
a_{31}(u_1, q) &= X(q, u_1) \\
a_{23}(p, v_2) &= -Y(p, v_2) \\
a_{32}(u_2, q) &= Y(q, u_2) \\
a_{14}(B_1, v_1) &= -B(B_1, v_1) \\
a_{41}(u_1, C_1) &= -B(u_1, C_1) \\
a_{25}(B_2, v_2) &= -B(B_2, v_2) \\
a_{52}(u_2, C_2) &= -B(u_2, C_2)
\end{aligned}$$

and $l1, \ldots, l5$ are the same as above.

7.3 Compressible Euler Equations

7.3.1 Introduction

The compressible Euler equations are equations for perfect fluid. Perfect fluids have no heat conduction ($T^{i0} = T^{0i} = 0$) and no viscosity ($T^{ij} = p\mathbb{1}$), so in the comoving frame the stress energy tensor is:

$$T^{\alpha\beta} = \operatorname{diag}(\rho c^2, p, p, p) = \left(\rho + \frac{p}{c^2}\right) u^\alpha u^\beta + p g^{\alpha\beta}$$

(we use $g^{\mu\nu} = \operatorname{diag}(-1, 1, 1, 1)$). Relativistic Euler equations are given by the conservation of the stress energy tensor and the particle number conservation:

$$\begin{aligned}
\partial_\nu T^{\mu\nu} &= 0 \\
\partial_\mu (n u^\mu) &= 0
\end{aligned}$$

By doing the nonrelativistic limit (see *Perfect Fluids* for a detailed derivation), we get the following Euler equations:

$$\begin{aligned}
\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) &= 0 \\
\frac{\partial (\rho \mathbf{u})}{\partial t} + \nabla \cdot (\rho \mathbf{u}\mathbf{u}^T) + \nabla p - \mathbf{f} &= 0 \\
\frac{\partial E}{\partial t} + \nabla \cdot (\mathbf{u}(E + p)) &= 0
\end{aligned}$$

where

$$E = \rho e + \tfrac{1}{2}\rho u^2$$

is the total energy per unit volume, composed of the kinetic energy per unit volume ($\frac{1}{2}\rho u^2$) and the internal energy per unit volume (ρe), where e is the internal energy per unit mass ($e = \frac{U}{nM}$). The energy E doesn't contain the rest mass energy, but all other energies are hidden in the internal energy.

We use the ideal gas equations, so:

$$\begin{aligned}
e &= T c_v \\
p = \frac{n}{V}\bar{R}T = \frac{nM}{V}\frac{\bar{R}}{M}T = \rho R T = \rho R \frac{e}{c_v} &= \frac{R}{c_v}(E - \tfrac{1}{2}\rho u^2)
\end{aligned}$$

where n is the number of moles of gas, M is the molar mass of the gas (i.e. a mass of one mole of the gas, e.g. for dry air we get $M = 28.956\,\mathrm{g/mol}$, as it is mainly composed of 20% of oxygen with atomic mass 16 and 78% of nitrogen with atomic mass 14, both form diatomic molecules, so the molecular mass is twice the atomic mass giving the total of $0.2 \cdot 2 \cdot 16 + 0.78 \cdot 2 \cdot 14 = 28.24$, the rest is given by the other components and one also has to average over all isotopes), $\bar{R} = N_A k_B \doteq 8.3145 \frac{\mathrm{J}}{\mathrm{K\,mol}}$ is the ideal gas constant (N_A is the Avogadro constant and k_B is the Boltzmann constant), $R = \frac{\bar{R}}{M}$ is the specific ideal gas constant (e.g. for dry air we get $R = \frac{8.3145}{28.956} \frac{\mathrm{J}}{\mathrm{g\,K}} \doteq 287.14 \frac{\mathrm{J}}{\mathrm{kg\,K}}$), $\rho = \frac{nM}{V} = \frac{p}{RT}$ is the density of the gas (e.g. for dry air at the pressure $10^5\,\mathrm{Pa}$ and temperature $22\,^\circ\mathrm{C}$ we get $\rho = \frac{10^5}{287.14 \cdot (22+273.15)} \frac{\mathrm{kg}}{\mathrm{m}^3} = 1.18 \frac{\mathrm{kg}}{\mathrm{m}^3}$), c_v is the specific heat capacity at constant volume (i.e. the amount of energy needed to heat one kg by one Kelvin at constant volume, e.g. for dry air the experimental value is about $c_v = 717.5 \frac{\mathrm{J}}{\mathrm{kg\,K}}$), V is the volume and T is the temperature of the gas. Of those, V, n, M, R, $\bar{R}$ are constants, ρ, e, E and T are functions of (t, x, y, z).

Here are the SI units of the various terms in the Euler equations:

$$\begin{aligned}
&= \mathrm{m\,s^{-1}} \\
[\rho] &= \mathrm{kg\,m^{-3}} \\
\mathrm{N} &= \mathrm{kg\,m\,s^{-2}} \\
\mathrm{J} = \mathrm{N\,m} &= \mathrm{kg\,m^2\,s^{-2}} \\
[p] = \mathrm{N\,m^{-2}} &= \mathrm{kg\,m^{-1}\,s^{-2}} \\
[\tfrac{1}{2}\rho u^2] = [\rho][u]^2 = \mathrm{kg\,m^{-3}\,m^2\,s^{-2}} &= \mathrm{kg\,m^{-1}\,s^{-2}} \\
[E] = \mathrm{J\,m^{-3}} &= \mathrm{kg\,m^{-1}\,s^{-2}} \\
[R] = \mathrm{J\,kg^{-1}\,K^{-1}} &= \mathrm{m^2\,s^{-2}\,K^{-1}} \\
[c_v] = \mathrm{J\,kg^{-1}\,K^{-1}} &= \mathrm{m^2\,s^{-2}\,K^{-1}} \\
[e] = \frac{[E]}{[\rho]} = \frac{\mathrm{kg\,m^{-1}\,s^{-2}}}{\mathrm{kg\,m^{-3}}} &= \mathrm{m^2\,s^{-2}}
\end{aligned}$$

In order to calculate the specific heat ratio κ, we use $R = c_p - c_v$:

$$\kappa = \frac{c_p}{c_v} = \frac{c_v + R}{c_v} = 1 + \frac{R}{c_v}$$

and the speed of sound is:

$$c = \sqrt{\kappa \frac{p}{\rho}}$$

7.3.2 Dimensionless Euler Equations

We choose 3 constants l_r, u_r and ρ_r - characteristic length of the domain, velocity and density. Now we multiply the Euler equations with proper combinations of these constants as follows:

$$\begin{aligned}
\left[\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u})\right] \frac{l_r}{\rho_r u_r} &= 0 \\
\left[\frac{\partial (\rho \mathbf{u})}{\partial t} + \nabla \cdot (\rho \mathbf{u}\mathbf{u}^T) + \nabla p - \mathbf{f}\right] \frac{l_r}{\rho_r u_r^2} &= 0 \\
\left[\frac{\partial E}{\partial t} + \nabla \cdot (\mathbf{u}(E+p))\right] \frac{l_r}{\rho_r u_r^3} &= 0
\end{aligned}$$

This is equal to:

$$\frac{\partial \tilde{\rho}}{\partial \tilde{t}} + \tilde{\nabla} \cdot (\tilde{\rho}\tilde{\mathbf{u}}) = 0$$
$$\frac{\partial (\tilde{\rho}\tilde{\mathbf{u}})}{\partial \tilde{t}} + \tilde{\nabla} \cdot (\tilde{\rho}\tilde{\mathbf{u}}\tilde{\mathbf{u}}^T) + \tilde{\nabla}\tilde{p} - \tilde{\mathbf{f}} = 0$$
$$\frac{\partial \tilde{E}}{\partial \tilde{t}} + \tilde{\nabla} \cdot (\tilde{\mathbf{u}}(\tilde{E} + \tilde{p})) = 0$$

where:

$$t_r = \frac{l_r}{u_r}$$
$$\tilde{t} = \frac{t}{t_r}$$
$$\tilde{\rho} = \frac{\rho}{\rho_r}$$
$$\tilde{\mathbf{u}} = \frac{\mathbf{u}}{u_r}$$
$$\tilde{\nabla} = l_r \nabla$$
$$\tilde{E} = \frac{E}{\rho_r u_r^2}$$
$$\tilde{p} = \frac{p}{\rho_r u_r^2}$$
$$\tilde{\mathbf{f}} = \mathbf{f}\frac{l_r}{\rho_r u_r^2}$$

In particular, if $\mathbf{f} = (0, 0, -\rho g)$, then

$$\tilde{\mathbf{f}} = (0, 0, -\tilde{\rho}\tilde{g})$$
$$\tilde{g} = g\frac{l_r}{u_r^2} = g\frac{t_r^2}{l_r}$$

So the dimensionless Euler equations look exactly the same as the original ones, we just need to rescale all the quantities using the relations above.

7.3.3 Conservative Form of the Euler Equations

We can write the Euler equations as:

$$\frac{\partial \mathbf{w}}{\partial t} + \frac{\partial \mathbf{f}_x}{\partial x} + \frac{\partial \mathbf{f}_y}{\partial y} + \frac{\partial \mathbf{f}_z}{\partial z} + \mathbf{g} = 0$$

where:

$$\mathbf{w} = \begin{pmatrix} \varrho \\ \rho u_1 \\ \rho u_2 \\ \rho u_3 \\ E \end{pmatrix} = \begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix}$$

$$\mathbf{f}_x = \begin{pmatrix} \rho u_1 \\ \rho u_1^2 + p \\ \rho u_1 u_2 \\ \rho u_1 u_3 \\ u_1(E+p) \end{pmatrix} = \begin{pmatrix} w_1 \\ \frac{w_1^2}{w_0} + p \\ \frac{w_1 w_2}{w_0} \\ \frac{w_1 w_3}{w_0} \\ \frac{w_1}{w_0}(w_4 + p) \end{pmatrix}$$

$$\mathbf{f}_y = \begin{pmatrix} \rho u_2 \\ \rho u_2 u_1 \\ \rho u_2^2 + p \\ \rho u_2 u_3 \\ u_2(E+p) \end{pmatrix} = \begin{pmatrix} w_2 \\ \frac{w_2 w_1}{w_0} \\ \frac{w_2^2}{w_0} + p \\ \frac{w_2 w_3}{w_0} \\ \frac{w_2}{w_0}(w_4 + p) \end{pmatrix}$$

$$\mathbf{f}_z = \begin{pmatrix} \rho u_3 \\ \rho u_3 u_1 \\ \rho u_3 u_2 \\ \rho u_3^2 + p \\ u_3(E+p) \end{pmatrix} = \begin{pmatrix} w_3 \\ \frac{w_3 w_1}{w_0} \\ \frac{w_3 w_2}{w_0} \\ \frac{w_3^2}{w_0} + p \\ \frac{w_3}{w_0}(w_4 + p) \end{pmatrix}$$

$$\mathbf{g} = \begin{pmatrix} 0 \\ -f_x \\ -f_y \\ -f_z \\ 0 \end{pmatrix}$$

$$p = \frac{R}{c_v}\left(E - \tfrac{1}{2}\rho\left(u_1^2 + u_2^2 + u_3^3\right)\right) = \frac{R}{c_v}\left(w_4 - \frac{w_1^2 + w_2^2 + w_3^2}{2w_0}\right)$$

We solve for the unknowns w_0, w_1, w_2, w_3 and w_4 as functions of (t, x, y, z), the rest (R, c_v, f_x, f_y, f_z) are either constants or depend on the unknowns. In order to convert from the physical quantities ρ, u_1, u_2, u_3, E and p to w_0, $\ldots$, w_4, we use:

$$\begin{aligned} w_0 &= \rho \\ w_1 &= \rho u_1 \\ w_2 &= \rho u_2 \\ w_3 &= \rho u_3 \\ w_4 = E &= p\frac{c_v}{R} + \tfrac{1}{2}\rho\left(u_1^2 + u_2^2 + u_3^2\right) \end{aligned}$$

the opposite conversion is:

$$\begin{aligned}\rho &= w_0\\ u_1 &= \frac{w_1}{w_0}\\ u_2 &= \frac{w_2}{w_0}\\ u_3 &= \frac{w_3}{w_0}\\ E &= w_4\\ p &= \frac{R}{c_v}\left(w_4 - \frac{w_1^2+w_2^2+w_3^2}{2w_0}\right)\end{aligned}$$

Sometimes people also use u, v and w instead of u_1, u_2 and u_3.

Note: $\rho\mathbf{u} \equiv \mathbf{j}$, where $\mathbf{j}$ is the fluid density current (it's a 3-vector) and also $w^\mu \equiv j^\mu$ (here w^μ is the same as w_μ, e.g. we are a bit sloppy about the notation), where j^μ is the density 4-current (e.g. the first 4 components of $\mathbf{w}$ are exactly the components of the 4-current j^μ):

$$j^\mu = \rho v^\mu = \rho\gamma(c, \mathbf{u}) = \gamma \begin{pmatrix} c\rho \\ \rho u_1 \\ \rho u_2 \\ \rho u_3 \end{pmatrix}$$

where as usual $\mu = 0, 1, 2, 3$ is the relativistic index, c is the speed of light, and in the nonrelativistic limit ($c \to \infty$) we get $\gamma \to 1$ and the remaining c in j^0 will cancel with c in $\partial_0 = \frac{1}{c}\frac{\partial}{\partial t}$, so it will not be present in the final equations (that involve terms like $\partial_\mu j^\mu$). We can also just set $c = 1$ as usual in relativistic physics.

7.3.4 Weak Formulation

The Euler equations:

$$\frac{\partial \mathbf{w}}{\partial t} + \frac{\partial \mathbf{f}_x}{\partial x} + \frac{\partial \mathbf{f}_y}{\partial y} + \frac{\partial \mathbf{f}_z}{\partial z} + \mathbf{g} = 0$$

are nonlinear. The time-derivative is approximated using the implicit Euler method

$$\frac{\mathbf{w}^{n+1} - \mathbf{w}^n}{\tau} + \frac{\partial \mathbf{f}_x(\mathbf{w}^{n+1})}{\partial x} + \frac{\partial \mathbf{f}_y(\mathbf{w}^{n+1})}{\partial y} + \frac{\partial \mathbf{f}_z(\mathbf{w}^{n+1})}{\partial z} + \mathbf{g} = 0$$

The vector-valued test functions for the above system of equations have the form:

$$\begin{pmatrix} \varphi^0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \varphi^1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \varphi^2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \varphi^3 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \varphi^4 \end{pmatrix}$$

After multiplying the equation system with the test functions and integrating over the domain Ω, we obtain (here the index $i = 0, 1, 2, 3, 4$ is numbering the 5 equations, so we are *not* summing over it):

$$\int_\Omega \frac{w_i^{n+1} - w_i^n}{\tau}\varphi^i + \frac{\partial \left(\mathbf{f}_x(\mathbf{w}^{n+1})\right)_i}{\partial x}\varphi^i + \frac{\partial \left(\mathbf{f}_y(\mathbf{w}^{n+1})\right)_i}{\partial y}\varphi^i + \frac{\partial \left(\mathbf{f}_z(\mathbf{w}^{n+1})\right)_i}{\partial z}\varphi^i + g_i\varphi^i \,\mathrm{d}^3x = 0$$

Now we integrate by parts:

$$\begin{aligned}\int_\Omega \frac{w_i^{n+1} - w_i^n}{\tau}\varphi^i - \left(\mathbf{f}_x(\mathbf{w}^{n+1})\right)_i \frac{\partial \varphi^i}{\partial x} - \left(\mathbf{f}_y(\mathbf{w}^{n+1})\right)_i \frac{\partial \varphi^i}{\partial y} - \left(\mathbf{f}_z(\mathbf{w}^{n+1})\right)_i \frac{\partial \varphi^i}{\partial z} + g_i\varphi^i \,\mathrm{d}^3x + \\ + \int_{\partial\Omega} \left(\mathbf{f}_x(\mathbf{w}^{n+1})\right)_i \varphi^i \, n_x + \left(\mathbf{f}_y(\mathbf{w}^{n+1})\right)_i \varphi^i \, n_y + \left(\mathbf{f}_z(\mathbf{w}^{n+1})\right)_i \varphi^i \, n_z \,\mathrm{d}^2x = 0\end{aligned}$$

where $\mathbf{n} = (n_x, n_y, n_z)$ is the outward surface normal to $\partial\Omega$. Rearranging:

$$\int_\Omega \frac{w_i^{n+1}}{\tau}\varphi^i - \left(\mathbf{f}_x(\mathbf{w}^{n+1})\right)_i \frac{\partial\varphi^i}{\partial x} - \left(\mathbf{f}_y(\mathbf{w}^{n+1})\right)_i \frac{\partial\varphi^i}{\partial y} - \left(\mathbf{f}_z(\mathbf{w}^{n+1})\right)_i \frac{\partial\varphi^i}{\partial z}\,\mathrm{d}^3x +$$
$$+ \int_{\partial\Omega} \left(\mathbf{f}_x(\mathbf{w}^{n+1})\right)_i \varphi^i\, n_x + \left(\mathbf{f}_y(\mathbf{w}^{n+1})\right)_i \varphi^i\, n_y + \left(\mathbf{f}_z(\mathbf{w}^{n+1})\right)_i \varphi^i\, n_z\,\mathrm{d}^2x = \int_\Omega \frac{w_i^n}{\tau}\varphi^i - g_i\varphi^i\,\mathrm{d}^3x$$

We can then linearize this for example by taking the flux jacobians $\mathbf{A}_x(\mathbf{w}^{n+1})$ on the previous time level $\mathbf{A}_x(\mathbf{w}^n)$.

The finite element formulation is obtained from here by replacing in the standard way the unknown solution w^{n+1} by a piecewise-polynomial unknown function

$$w_h^{n+1} = \sum_{k=1}^{N} y_k\psi_k,$$

where ψ_k are the basis functions of the piecewise-polynomial finite element space. This turns the above weak formulation into a finite number of nonlinear algebraic equations of the form $F(Y) = 0$ that will be solved using the Newton's method.

Explicit Method

We also derive the weak formulation for the explicit method. Euler equations:

$$\frac{\partial\mathbf{w}}{\partial t} + \frac{\partial\mathbf{f}_x}{\partial x} + \frac{\partial\mathbf{f}_y}{\partial y} + \frac{\partial\mathbf{f}_z}{\partial z} + \mathbf{g} = 0$$

The time-derivative is approximated using the explicit Euler method

$$\frac{\mathbf{w}^{n+1} - \mathbf{w}^n}{\tau} + \frac{\partial\mathbf{f}_x(\mathbf{w}^n)}{\partial x} + \frac{\partial\mathbf{f}_y(\mathbf{w}^n)}{\partial y} + \frac{\partial\mathbf{f}_z(\mathbf{w}^n)}{\partial z} + \mathbf{g} = 0$$

The vector-valued test functions for the above system of equations have the form:

$$\begin{pmatrix}\varphi^0\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\\varphi^1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\\varphi^2\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\\varphi^3\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\0\\\varphi^4\end{pmatrix}$$

After multiplying the equation system with the test functions and integrating over the domain Ω, we obtain (here the index $i = 0, 1, 2, 3, 4$ is numbering the 5 equations, so we are *not* summing over it):

$$\int_\Omega \frac{w_i^{n+1} - w_i^n}{\tau}\varphi^i + \frac{\partial\left(\mathbf{f}_x(\mathbf{w}^n)\right)_i}{\partial x}\varphi^i + \frac{\partial\left(\mathbf{f}_y(\mathbf{w}^n)\right)_i}{\partial y}\varphi^i + \frac{\partial\left(\mathbf{f}_z(\mathbf{w}^n)\right)_i}{\partial z}\varphi^i + g_i\varphi^i\,\mathrm{d}^3x = 0$$

Now we integrate by parts:

$$\int_\Omega \frac{w_i^{n+1} - w_i^n}{\tau}\varphi^i - \left(\mathbf{f}_x(\mathbf{w}^n)\right)_i \frac{\partial\varphi^i}{\partial x} - \left(\mathbf{f}_y(\mathbf{w}^n)\right)_i \frac{\partial\varphi^i}{\partial y} - \left(\mathbf{f}_z(\mathbf{w}^n)\right)_i \frac{\partial\varphi^i}{\partial z} + g_i\varphi^i\,\mathrm{d}^3x +$$
$$+ \int_{\partial\Omega} \left(\mathbf{f}_x(\mathbf{w}^n)\right)_i \varphi^i\, n_x + \left(\mathbf{f}_y(\mathbf{w}^n)\right)_i \varphi^i\, n_y + \left(\mathbf{f}_z(\mathbf{w}^n)\right)_i \varphi^i\, n_z\,\mathrm{d}^2x = 0$$

where $\mathbf{n} = (n_x, n_y, n_z)$ is the outward surface normal to $\partial\Omega$. Rearranging:

$$\int_\Omega \frac{w_i^{n+1}}{\tau}\varphi^i\,\mathrm{d}^3x = \int_\Omega \frac{w_i^n}{\tau}\varphi^i + \left(\mathbf{f}_x(\mathbf{w}^n)\right)_i \frac{\partial\varphi^i}{\partial x} + \left(\mathbf{f}_y(\mathbf{w}^n)\right)_i \frac{\partial\varphi^i}{\partial y} + \left(\mathbf{f}_z(\mathbf{w}^n)\right)_i \frac{\partial\varphi^i}{\partial z} - g_i\varphi^i\,\mathrm{d}^3x +$$
$$- \int_{\partial\Omega} \left(\mathbf{f}_x(\mathbf{w}^n)\right)_i \varphi^i\, n_x + \left(\mathbf{f}_y(\mathbf{w}^n)\right)_i \varphi^i\, n_y + \left(\mathbf{f}_z(\mathbf{w}^n)\right)_i \varphi^i\, n_z\,\mathrm{d}^2x$$

7.3.5 Flux Jacobians

Now we write the spatial derivatives using the so called flux Jacobians $\mathbf{A}_x$, $\mathbf{A}_y$ and $\mathbf{A}_z$:

$$\frac{\partial \mathbf{f}_x}{\partial x} = \frac{\partial \mathbf{f}_x}{\partial \mathbf{w}} \frac{\partial \mathbf{w}}{\partial x} \equiv \mathbf{A}_x \frac{\partial \mathbf{w}}{\partial x}$$
$$\mathbf{A}_x = \mathbf{A}_x(\mathbf{w}) \equiv \frac{\partial \mathbf{f}_x}{\partial \mathbf{w}}$$

Similarly for y and z, so we get:

$$\frac{\partial \mathbf{w}}{\partial t} + \mathbf{A}_x \frac{\partial \mathbf{w}}{\partial x} + \mathbf{A}_y \frac{\partial \mathbf{w}}{\partial y} + \mathbf{A}_z \frac{\partial \mathbf{w}}{\partial z} + \mathbf{g} = 0$$

One nice thing about these particular $\mathbf{f}_x$, $\mathbf{f}_y$ and $\mathbf{f}_z$ functions is that they are homogeneous of degree 1:

$$\mathbf{f}_x(\lambda \mathbf{w}) = \lambda \, \mathbf{f}_x(\mathbf{w})$$

so the Euler equation/formula for the homogeneous function is:

$$\mathbf{w} \cdot \frac{\partial \mathbf{f}_x(\mathbf{w})}{\partial \mathbf{w}} = \mathbf{f}_x(\mathbf{w})$$
$$\mathbf{w} \cdot \mathbf{A}_x = \mathbf{f}_x(\mathbf{w})$$

So both the $\mathbf{f}_x$ and it's derivative can be nicely factored out using the flux Jacobian:

$$\mathbf{f}_x = \mathbf{A}_x \, \mathbf{w}$$
$$\frac{\partial \mathbf{f}_x}{\partial x} = \mathbf{A}_x \frac{\partial \mathbf{w}}{\partial x}$$

by differentiating the first equation and substracting the second, we get:

$$\frac{\partial \mathbf{A}_x}{\partial x} \, \mathbf{w} = 0$$

similarly for y and z. To calculate the Jacobians, we'll need:

$$\frac{\partial p}{\partial \mathbf{w}} = \frac{R}{c_v} \begin{pmatrix} \frac{w_1^2+w_2^2+w_3^2}{2w_0^2} & -\frac{w_1}{w_0} & -\frac{w_2}{w_0} & -\frac{w_3}{w_0} & 1 \end{pmatrix}$$

then we can calculate the Jacobians (and we substitute for p):

$$\mathbf{A}_x(\mathbf{w}) = \frac{\partial \mathbf{f}_x}{\partial \mathbf{w}} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ -\frac{w_1^2}{w_0^2} + \frac{R}{c_v}\frac{w_1^2+w_2^2+w_3^2}{2w_0^2} & \frac{2w_1}{w_0} - \frac{R}{c_v}\frac{w_1}{w_0} & -\frac{R}{c_v}\frac{w_2}{w_0} & -\frac{R}{c_v}\frac{w_3}{w_0} & \frac{R}{c_v} \\ -\frac{w_1 w_2}{w_0^2} & \frac{w_2}{w_0} & \frac{w_1}{w_0} & 0 & 0 \\ -\frac{w_1 w_3}{w_0^2} & \frac{w_3}{w_0} & 0 & \frac{w_1}{w_0} & 0 \\ -\frac{w_1 w_4}{w_0^2} - \frac{w_1}{w_0^2}\frac{R}{c_v}\left(w_4 - \frac{w_1^2+w_2^2+w_3^2}{2w_0}\right) + \frac{w_1}{w_0}\frac{R}{c_v}\frac{w_1^2+w_2^2+w_3^2}{2w_0^2} & \frac{w_4}{w_0} + \frac{1}{w_0}\frac{R}{c_v}\left(w_4 - \frac{w_1^2+w_2^2+w_3^2}{2w_0}\right) - \frac{R}{c_v}\frac{w_1^2}{w_0^2} & -\frac{R}{c_v}\frac{w_1 w_2}{w_0^2} & -\frac{R}{c_v}\frac{w_1 w_3}{w_0^2} & \frac{w_1}{w_0}\left(1+\frac{R}{c_v}\right) \end{pmatrix}$$

$$\mathbf{A}_y(\mathbf{w}) = \frac{\partial \mathbf{f}_y}{\partial \mathbf{w}} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ -\frac{w_2 w_1}{w_0^2} & \frac{w_2}{w_0} & \frac{w_1}{w_0} & 0 & 0 \\ -\frac{w_2^2}{w_0^2} + \frac{R}{c_v}\frac{w_1^2+w_2^2+w_3^2}{2w_0^2} & -\frac{R}{c_v}\frac{w_1}{w_0} & \frac{2w_2}{w_0} - \frac{R}{c_v}\frac{w_2}{w_0} & -\frac{R}{c_v}\frac{w_3}{w_0} & \frac{R}{c_v} \\ -\frac{w_2 w_3}{w_0^2} & 0 & \frac{w_3}{w_0} & \frac{w_2}{w_0} & 0 \\ -\frac{w_2 w_4}{w_0^2} - \frac{w_2}{w_0^2}\frac{R}{c_v}\left(w_4 - \frac{w_1^2+w_2^2+w_3^2}{2w_0}\right) + \frac{w_2}{w_0}\frac{R}{c_v}\frac{w_1^2+w_2^2+w_3^2}{2w_0^2} & -\frac{R}{c_v}\frac{w_2 w_1}{w_0^2} & \frac{w_4}{w_0} + \frac{1}{w_0}\frac{R}{c_v}\left(w_4 - \frac{w_1^2+w_2^2+w_3^2}{2w_0}\right) - \frac{R}{c_v}\frac{w_2^2}{w_0^2} & -\frac{R}{c_v}\frac{w_2 w_3}{w_0^2} & \frac{w_2}{w_0}\left(1+\frac{R}{c_v}\right) \end{pmatrix}$$

$$\mathbf{A}_z(\mathbf{w}) = \frac{\partial \mathbf{f}_z}{\partial \mathbf{w}} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ -\frac{w_3 w_1}{w_0^2} & \frac{w_3}{w_0} & 0 & \frac{w_1}{w_0} & 0 \\ -\frac{w_3 w_2}{w_0^2} & 0 & \frac{w_3}{w_0} & \frac{w_2}{w_0} & 0 \\ -\frac{w_3^2}{w_0^2} + \frac{R}{c_v}\frac{w_1^2+w_2^2+w_3^2}{2w_0^2} & -\frac{R}{c_v}\frac{w_1}{w_0} & -\frac{R}{c_v}\frac{w_2}{w_0} & \frac{2w_3}{w_0} - \frac{R}{c_v}\frac{w_3}{w_0} & \frac{R}{c_v} \\ -\frac{w_3 w_4}{w_0^2} - \frac{w_3}{w_0^2}\frac{R}{c_v}\left(w_4 - \frac{w_1^2+w_2^2+w_3^2}{2w_0}\right) + \frac{w_3}{w_0}\frac{R}{c_v}\frac{w_1^2+w_2^2+w_3^2}{2w_0^2} & -\frac{R}{c_v}\frac{w_3 w_1}{w_0^2} & -\frac{R}{c_v}\frac{w_3 w_2}{w_0^2} & \frac{w_4}{w_0} + \frac{1}{w_0}\frac{R}{c_v}\left(w_4 - \frac{w_1^2+w_2^2+w_3^2}{2w_0}\right) & \frac{w_3}{w_0}\left(1+\frac{R}{c_v}\right) \end{pmatrix}$$

7.3.6 2D Version of the Equations

$$\frac{\partial \mathbf{w}}{\partial t} + \frac{\partial \mathbf{f}_x}{\partial x} + \frac{\partial \mathbf{f}_y}{\partial y} + \mathbf{g} = 0$$

where:

$$\mathbf{w} = \begin{pmatrix} \varrho \\ \rho u_1 \\ \rho u_2 \\ E \end{pmatrix} = \begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \end{pmatrix}$$

$$\mathbf{f}_x = \begin{pmatrix} \rho u_1 \\ \rho u_1^2 + p \\ \rho u_1 u_2 \\ u_1(E+p) \end{pmatrix} = \begin{pmatrix} w_1 \\ \frac{w_1^2}{w_0} + p \\ \frac{w_1 w_2}{w_0} \\ \frac{w_1}{w_0}(w_3 + p) \end{pmatrix}$$

$$\mathbf{f}_y = \begin{pmatrix} \rho u_2 \\ \rho u_2 u_1 \\ \rho u_2^2 + p \\ u_2(E+p) \end{pmatrix} = \begin{pmatrix} w_2 \\ \frac{w_2 w_1}{w_0} \\ \frac{w_2^2}{w_0} + p \\ \frac{w_2}{w_0}(w_3 + p) \end{pmatrix}$$

$$\mathbf{g} = \begin{pmatrix} 0 \\ -f_x \\ -f_y \\ 0 \end{pmatrix}$$

$$p = \frac{R}{c_v}\left(E - \tfrac{1}{2}\rho\left(u_1^2 + u_2^2\right)\right) = \frac{R}{c_v}\left(w_3 - \frac{w_1^2 + w_2^2}{2w_0}\right)$$

Discretizing the time derivative:

$$\frac{\mathbf{w}^{n+1} - \mathbf{w}^n}{\tau} + \frac{\partial \mathbf{f}_x(\mathbf{w}^{n+1})}{\partial x} + \frac{\partial \mathbf{f}_y(\mathbf{w}^{n+1})}{\partial y} + \mathbf{g} = 0$$

The vector-valued test functions for the above system of equations have the form:

$$\begin{pmatrix} \varphi^0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \varphi^1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \varphi^2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \varphi^3 \end{pmatrix}$$

After multiplying the equation system with the test functions and integrating over the domain Ω, we obtain:

$$\int_\Omega \frac{w_i^{n+1} - w_i^n}{\tau}\varphi^i + \frac{\partial\left(\mathbf{f}_x(\mathbf{w}^{n+1})\right)_i}{\partial x}\varphi^i + \frac{\partial\left(\mathbf{f}_y(\mathbf{w}^{n+1})\right)_i}{\partial y}\varphi^i + g_i\varphi^i \,\mathrm{d}^2x = 0$$

Now we integrate by parts:

$$\int_\Omega \frac{w_i^{n+1} - w_i^n}{\tau}\varphi^i - \left(\mathbf{f}_x(\mathbf{w}^{n+1})\right)_i \frac{\partial \varphi^i}{\partial x} - \left(\mathbf{f}_y(\mathbf{w}^{n+1})\right)_i \frac{\partial \varphi^i}{\partial y} + g_i\varphi^i \,\mathrm{d}^2x + \\ + \int_{\partial\Omega} \left(\mathbf{f}_x(\mathbf{w}^{n+1})\right)_i \varphi^i \, n_x + \left(\mathbf{f}_y(\mathbf{w}^{n+1})\right)_i \varphi^i \, n_y \,\mathrm{d}x = 0$$

where $\mathbf{n} = (n_x, n_y)$ is the outward surface normal to $\partial\Omega$. Rearranging:

$$\int_\Omega \frac{w_i^{n+1}}{\tau}\varphi^i - \left(\mathbf{f}_x(\mathbf{w}^{n+1})\right)_i \frac{\partial \varphi^i}{\partial x} - \left(\mathbf{f}_y(\mathbf{w}^{n+1})\right)_i \frac{\partial \varphi^i}{\partial y} \,\mathrm{d}^2x + \\ + \int_{\partial\Omega} \left(\mathbf{f}_x(\mathbf{w}^{n+1})\right)_i \varphi^i \, n_x + \left(\mathbf{f}_y(\mathbf{w}^{n+1})\right)_i \varphi^i \, n_y \,\mathrm{d}x = \int_\Omega \frac{w_i^n}{\tau}\varphi^i - g_i\varphi^i \,\mathrm{d}^2x$$

The 2D flux Jacobians are:

$$\mathbf{A}_x(\mathbf{w}) = \frac{\partial \mathbf{f}_x}{\partial \mathbf{w}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -\frac{w_1^2}{w_0^2} + \frac{R}{c_v}\frac{w_1^2+w_2^2}{2w_0^2} & \frac{2w_1}{w_0} - \frac{R}{c_v}\frac{w_1}{w_0} & -\frac{R}{c_v}\frac{w_2}{w_0} & \frac{R}{c_v} \\ -\frac{w_1 w_2}{w_0^2} & \frac{w_2}{w_0} & \frac{w_1}{w_0} & 0 \\ -\frac{w_1 w_3}{w_0^2} - \frac{w_1}{w_0^2}\frac{R}{c_v}\left(w_3 - \frac{w_1^2+w_2^2}{2w_0}\right) + \frac{w_1}{w_0}\frac{R}{c_v}\frac{w_1^2+w_2^2}{2w_0^2} & \frac{w_3}{w_0} + \frac{1}{w_0}\frac{R}{c_v}\left(w_3 - \frac{w_1^2+w_2^2}{2w_0}\right) - \frac{R}{c_v}\frac{w_1^2}{w_0^2} & -\frac{R}{c_v}\frac{w_1 w_2}{w_0^2} & \frac{w_1}{w_0} + \frac{R}{c_v}\frac{w_1}{w_0} \end{pmatrix}$$

$$\mathbf{A}_y(\mathbf{w}) = \frac{\partial \mathbf{f}_y}{\partial \mathbf{w}} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ -\frac{w_2 w_1}{w_0^2} & \frac{w_2}{w_0} & \frac{w_1}{w_0} & 0 \\ -\frac{w_2^2}{w_0^2} + \frac{R}{c_v}\frac{w_1^2+w_2^2}{2w_0^2} & -\frac{R}{c_v}\frac{w_1}{w_0} & \frac{2w_2}{w_0} - \frac{R}{c_v}\frac{w_2}{w_0} & \frac{R}{c_v} \\ -\frac{w_2 w_3}{w_0^2} - \frac{w_2}{w_0^2}\frac{R}{c_v}\left(w_3 - \frac{w_1^2+w_2^2}{2w_0}\right) + \frac{w_2}{w_0}\frac{R}{c_v}\frac{w_1^2+w_2^2}{2w_0^2} & -\frac{R}{c_v}\frac{w_2 w_1}{w_0^2} & \frac{w_3}{w_0} + \frac{1}{w_0}\frac{R}{c_v}\left(w_3 - \frac{w_1^2+w_2^2}{2w_0}\right) - \frac{R}{c_v}\frac{w_2^2}{w_0^2} & \frac{w_2}{w_0} + \frac{R}{c_v}\frac{w_2}{w_0} \end{pmatrix}$$

7.3.7 Sea Breeze Modeling

In our 2D model we make the following assumptions:

$$f_x = 0$$
$$f_y = -\rho g = -w_0 g$$

and the boundary condition is as follows:

$$T'(x,t) = \left(\frac{A}{2}\right)\sin\left(\frac{\pi(t-t_0)}{24}\right)\left(1 + \tanh\left(\frac{S(x)}{L}\right)\right)$$
$$T(x) = T_0 + T'(x,t)$$

The weak formulation in 2D is (here $i = 0, 1, 2, 3$):

$$\int_\Omega \frac{w_i^{n+1}}{\tau}\varphi^i - \left(\mathbf{A}_x(\mathbf{w}^n)\right)_{ij} w_j^{n+1}\frac{\partial \varphi^i}{\partial x} - \left(\mathbf{A}_z(\mathbf{w}^n)\right)_{ij} w_j^{n+1}\frac{\partial \varphi^i}{\partial z}\,\mathrm{d}^2x +$$
$$+ \int_{\partial\Omega} \left(\mathbf{A}_x(\mathbf{w}^n)\right)_{ij} w_j^{n+1}\varphi^i\, n_x + \left(\mathbf{A}_z(\mathbf{w}^n)\right)_{ij} w_j^{n+1}\varphi^i\, n_z\,\mathrm{d}x = \int_\Omega \frac{w_i^n}{\tau}\varphi^i - g_i\varphi^i\,\mathrm{d}^2x$$

In order to specify the input forms for Hermes, we'll write the weak formulation as:

$$B_{00}(w_0,\varphi^0) + B_{01}(w_1,\varphi^0) + B_{02}(w_2,\varphi^0) + B_{03}(w_3,\varphi^0) = l_0(\varphi^0)$$
$$B_{10}(w_0,\varphi^1) + B_{11}(w_1,\varphi^1) + B_{12}(w_2,\varphi^1) + B_{13}(w_3,\varphi^1) = l_1(\varphi^1)$$
$$B_{20}(w_0,\varphi^2) + B_{21}(w_1,\varphi^2) + B_{22}(w_2,\varphi^2) + B_{23}(w_3,\varphi^2) = l_2(\varphi^2)$$
$$B_{30}(w_0,\varphi^3) + B_{31}(w_1,\varphi^3) + B_{32}(w_2,\varphi^3) + B_{33}(w_3,\varphi^3) = l_3(\varphi^3)$$

where the forms are (we write w_i instead of w_i^{n+1}):

$$l_0(\varphi^0) = \int_\Omega \frac{w_0^n\varphi^0}{\tau}\,\mathrm{d}^2x$$
$$l_1(\varphi^1) = \int_\Omega \frac{w_1^n\varphi^1}{\tau}\,\mathrm{d}^2x$$
$$l_2(\varphi^2) = \int_\Omega \frac{w_2^n\varphi^2}{\tau} + \rho g\varphi^2\,\mathrm{d}^2x$$
$$l_3(\varphi^3) = \int_\Omega \frac{w_3^n\varphi^3}{\tau}\,\mathrm{d}^2x$$
$$B_{ij}(w_j,\varphi^i) = \int_\Omega \frac{w_i}{\tau}\varphi^i\delta_{ij} - \left(\mathbf{A}_x(\mathbf{w}^n)\right)_{ij} w_j\frac{\partial\varphi^i}{\partial x} - \left(\mathbf{A}_z(\mathbf{w}^n)\right)_{ij} w_j\frac{\partial\varphi^i}{\partial z}\,\mathrm{d}^2x$$

In the last expression we do *not* sum over i nor j. In particular:

$$B_{00}(w_0, \varphi^0) = \int_\Omega \frac{w_0}{\tau}\varphi^0 - (\mathbf{A}_x(\mathbf{w}^n))_{00}\, w_0 \frac{\partial \varphi^0}{\partial x} - (\mathbf{A}_z(\mathbf{w}^n))_{00}\, w_0 \frac{\partial \varphi^0}{\partial z}\, \mathrm{d}^2x = \int_\Omega \frac{w_0}{\tau}\varphi^0\, \mathrm{d}^2x$$

$$B_{01}(w_1, \varphi^0) = \int_\Omega - (\mathbf{A}_x(\mathbf{w}^n))_{01}\, w_1 \frac{\partial \varphi^0}{\partial x} - (\mathbf{A}_z(\mathbf{w}^n))_{01}\, w_1 \frac{\partial \varphi^0}{\partial z}\, \mathrm{d}^2x = \int_\Omega - (\mathbf{A}_x(\mathbf{w}^n))_{01}\, w_1 \frac{\partial \varphi^0}{\partial x}\, \mathrm{d}^2x$$

$$B_{02}(w_2, \varphi^0) = \int_\Omega - (\mathbf{A}_x(\mathbf{w}^n))_{02}\, w_2 \frac{\partial \varphi^0}{\partial x} - (\mathbf{A}_z(\mathbf{w}^n))_{02}\, w_2 \frac{\partial \varphi^0}{\partial z}\, \mathrm{d}^2x = \int_\Omega - (\mathbf{A}_z(\mathbf{w}^n))_{02}\, w_2 \frac{\partial \varphi^0}{\partial z}\, \mathrm{d}^2x$$

$$B_{03}(w_3, \varphi^0) = \int_\Omega - (\mathbf{A}_x(\mathbf{w}^n))_{03}\, w_3 \frac{\partial \varphi^0}{\partial x} - (\mathbf{A}_z(\mathbf{w}^n))_{03}\, w_3 \frac{\partial \varphi^0}{\partial z}\, \mathrm{d}^2x = 0$$

$$B_{10}(w_0, \varphi^1) = \int_\Omega - (\mathbf{A}_x(\mathbf{w}^n))_{10}\, w_0 \frac{\partial \varphi^1}{\partial x} - (\mathbf{A}_z(\mathbf{w}^n))_{10}\, w_0 \frac{\partial \varphi^1}{\partial z}\, \mathrm{d}^2x$$

$$B_{11}(w_1, \varphi^1) = \int_\Omega \frac{w_1}{\tau}\varphi^1 - (\mathbf{A}_x(\mathbf{w}^n))_{11}\, w_1 \frac{\partial \varphi^1}{\partial x} - (\mathbf{A}_z(\mathbf{w}^n))_{11}\, w_1 \frac{\partial \varphi^1}{\partial z}\, \mathrm{d}^2x$$

$$\dots$$

7.3.8 Boundary Conditions

We rewrite the boundary integral by rotating coordinates, so that the flow is only in the x direction (thus we only have $\mathbf{f}_x$):

$$\int_{\partial\Omega} (\mathbf{f}_x(\mathbf{w}))_i\, \varphi^i\, n_x + (\mathbf{f}_y(\mathbf{w}))_i\, \varphi^i\, n_y + (\mathbf{f}_z(\mathbf{w}))_i\, \varphi^i\, n_z\, \mathrm{d}^2x = \\ = \int_{\partial\Omega} T^{-1}\mathbf{f}_x(T\mathbf{w})\varphi^i\, \mathrm{d}^2x$$

Now we need to approximate $\mathbf{f}_x(T\mathbf{w})$ somehow. We do that by solving the following 1D problem (Riemann problem):

$$\frac{\partial \mathbf{w}}{\partial t} + \frac{\partial}{\partial x}\mathbf{f}(\mathbf{w}) = 0$$

or:

$$\frac{\partial \mathbf{w}}{\partial t} + \mathbf{A}(\mathbf{w})\frac{\partial \mathbf{w}}{\partial x} = 0 \tag{7.3.8.1}$$

$$\mathbf{w}(x,t) = \begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix}$$

And we approximate $\mathbf{f}_x(\mathbf{w}) = \mathbf{f}(\mathbf{w}(0,t))$. The initial condition is:

$$\mathbf{w}(x,0) = \begin{cases} \mathbf{w}_L & x < 0 \\ \mathbf{w}_R & x > 0 \end{cases} = \mathbf{w}_L(1 - H(x)) + \mathbf{w}_R H(x)$$

Now we write:

$$\mathbf{w}(x,t) = \sum_i \xi^i(x,t)\mathbf{r}_i$$

$$\mathbf{w}_L = \sum_i \alpha_i \mathbf{r}_i$$

$$\mathbf{w}_R = \sum_i \beta_i \mathbf{r}_i$$

$$\xi^i(x,0) = \begin{cases} \alpha_i & x < 0 \\ \beta_i & x > 0 \end{cases}$$

and substitute into (7.3.8.1):

$$\sum_i \left(\frac{\partial \xi^i}{\partial t} + \mathbf{A}(\mathbf{w}) \frac{\partial \xi^i}{\partial x} \right) \mathbf{r}_i = 0$$
$$\sum_i \left(\frac{\partial \xi^i}{\partial t} + \lambda_i(\mathbf{w}) \frac{\partial \xi^i}{\partial x} \right) \mathbf{r}_i = 0$$

so we get:

$$\frac{\partial \xi^i}{\partial t} + \lambda_i(\mathbf{w}) \frac{\partial \xi^i}{\partial x} = 0$$

This is a nonlinear problem, that cannot be solved exactly. First, let $\mathbf{A}$ doesn't depend on $\mathbf{w}$. Then also λ_i are constants:

$$\frac{\partial \xi^i}{\partial t} + \lambda_i \frac{\partial \xi^i}{\partial x} = 0$$

and the solution is constant along the characteristic $x(t) = \lambda_i t + c$ for $t > 0$ and we get:

$$\xi_i(x,t) = \xi^i(x - \lambda_i t, 0) = \begin{cases} \alpha_i & x - \lambda_i t < 0 \\ \beta_i & x - \lambda_i t > 0 \end{cases} = \alpha_i(1 - H(x - \lambda_i t)) + \beta_i H(x - \lambda_i t)$$

and

$$\mathbf{w}(x,t) = \sum_i \xi^i(x,t)\mathbf{r}_i = \sum_i \left(\alpha_i(1 - H(x - \lambda_i t)) + \beta_i H(x - \lambda_i t)\right) \mathbf{r}_i$$
$$\mathbf{w}(0,t) = \sum_i \left(\alpha_i(1 - H(-\lambda_i t)) + \beta_i H(-\lambda_i t)\right) \mathbf{r}_i =$$
$$= \sum_i \left(\alpha_i H(\lambda_i t) + \beta_i H(-\lambda_i t)\right) \mathbf{r}_i =$$
$$= \sum_i \left(\alpha_i H(\lambda_i) + \beta_i H(-\lambda_i)\right) \mathbf{r}_i =$$
$$= \sum_{i=k+1}^{n} \alpha_i \mathbf{r}_i + \sum_{i=1}^{k} \beta_i \mathbf{r}_i$$

so:

$$\mathbf{f}(\mathbf{w}(0,t)) = \mathbf{A}\mathbf{w}(0,t) = \sum_{i=k+1}^{n} \mathbf{A}\alpha_i \mathbf{r}_i + \sum_{i=1}^{k} \mathbf{A}\beta_i \mathbf{r}_i$$
$$= \sum_{i=k+1}^{n} \lambda_i \alpha_i \mathbf{r}_i + \sum_{i=1}^{k} \lambda_i \beta_i \mathbf{r}_i =$$
$$= \mathbf{A}^+ \sum_{i=1}^{n} \alpha_i \mathbf{r}_i + \mathbf{A}^- \sum_{i=1}^{n} \beta_i \mathbf{r}_i =$$
$$= \mathbf{A}^+ \mathbf{w}_L + \mathbf{A}^- \mathbf{w}_R$$

In the nonlinear case we cannot solve it exactly, but we can approximate the solution by:

$$\mathbf{f}(\mathbf{w}(0,t)) = \mathbf{f}^+(\mathbf{w}_L) + \mathbf{f}^-(\mathbf{w}_R) =$$
$$= \mathbf{f}(\mathbf{w}_R) - \int_{\mathbf{w}_L}^{\mathbf{w}_R} \mathbf{A}^+(\mathbf{w}) \mathrm{d}\mathbf{w} =$$
$$= \mathbf{f}(\mathbf{w}_L) + \int_{\mathbf{w}_L}^{\mathbf{w}_R} \mathbf{A}^-(\mathbf{w}) \mathrm{d}\mathbf{w} \approx$$

$$\approx \mathbf{f}(\mathbf{w}_L) + \mathbf{A}^-(\mathbf{w}_R)\mathbf{w}_R - \mathbf{A}^-(\mathbf{w}_L)\mathbf{w}_L \tag{7.3.8.2}$$

Let's say the domain is for $x < 0$ and we are applying the BC condition from $x > 0$. Then $\mathbf{w}_L$ is taken from the solution and $\mathbf{w}_R$ is prescribed, for example at the bottom it could be:

$$\mathbf{w}_R = \begin{pmatrix} \rho \\ \rho u_1 \\ 0 \\ 0 \\ E \end{pmatrix}$$

Now we need to calculate $\mathbf{A}^-$. First we write:

$$\mathbf{A}_x = \mathbf{R}\mathbf{D}_x\mathbf{R}^{-1}$$
$$\mathbf{A}_x^- = \mathbf{R}\mathbf{D}_x^-\mathbf{R}^{-1}$$
$$\mathbf{D}_x(\mathbf{w}) = \frac{w_1}{w_0}\mathbb{1} + \operatorname{diag}(-c, 0, 0, 0, c) = \begin{pmatrix} u_1 - c & 0 & 0 & 0 & 0 \\ 0 & u_1 & 0 & 0 & 0 \\ 0 & 0 & u_1 & 0 & 0 \\ 0 & 0 & 0 & u_1 & 0 \\ 0 & 0 & 0 & 0 & u_1 + c \end{pmatrix}$$
$$\mathbf{D}_x(\mathbf{w})^- = \begin{cases} \operatorname{diag}(\frac{w_1}{w_0} - c, \frac{w_1}{w_0}, \frac{w_1}{w_0}, \frac{w_1}{w_0}, 0) & w_1 < 0 \\ \operatorname{diag}(\frac{w_1}{w_0} - c, 0, 0, 0, 0) & w_1 > 0 \end{cases}$$

Explicit forms of the matrices:

$$\mathbf{R} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ u-c & u & u & u & c+u \\ v & v & v-c & v & v \\ w & w & w & w-c & w \\ -cu-\frac{c^2}{1-\kappa}+\frac{1}{2}u^2+\frac{1}{2}v^2+\frac{1}{2}w^2 & \frac{1}{2}u^2+\frac{1}{2}v^2+\frac{1}{2}w^2 & -cv+\frac{1}{2}u^2+\frac{1}{2}v^2+\frac{1}{2}w^2 & -cw+\frac{1}{2}u^2+\frac{1}{2}v^2+\frac{1}{2}w^2 & cu-\frac{c^2}{1-\kappa}+\frac{1}{2}u^2+\frac{1}{2}v^2+\frac{1}{2}w^2 \end{pmatrix}$$

$$\mathbf{R}^{-1} = \frac{1}{c^2}\begin{pmatrix} \frac{1}{2}cu-\frac{1}{4}u^2-\frac{1}{4}v^2-\frac{1}{4}w^2+\frac{1}{4}\kappa u^2+\frac{1}{4}\kappa v^2+\frac{1}{4}\kappa w^2 & \frac{1}{2}u-\frac{1}{2}c-\frac{1}{2}\kappa u & \frac{1}{2}v-\frac{1}{2}\kappa v & \frac{1}{2}w-\frac{1}{2}\kappa w & -\frac{1}{2}+\frac{1}{2}\kappa \\ -cv-cw+c^2+\frac{1}{2}u^2+\frac{1}{2}v^2+\frac{1}{2}w^2-\frac{1}{2}\kappa u^2-\frac{1}{2}\kappa v^2-\frac{1}{2}\kappa w^2 & -u+\kappa u & c-v+\kappa v & c-w+\kappa w & 1-\kappa \\ cv & 0 & -c & 0 & 0 \\ cw & 0 & 0 & -c & 0 \\ -\frac{1}{2}cu-\frac{1}{4}u^2-\frac{1}{4}v^2-\frac{1}{4}w^2+\frac{1}{4}\kappa u^2+\frac{1}{4}\kappa v^2+\frac{1}{4}\kappa w^2 & \frac{1}{2}c+\frac{1}{2}u-\frac{1}{2}\kappa u & \frac{1}{2}v-\frac{1}{2}\kappa v & \frac{1}{2}w-\frac{1}{2}\kappa w & -\frac{1}{2}+\frac{1}{2}\kappa \end{pmatrix}$$

$$\mathbf{D}_x = \begin{pmatrix} u-c & 0 & 0 & 0 & 0 \\ 0 & u & 0 & 0 & 0 \\ 0 & 0 & u & 0 & 0 \\ 0 & 0 & 0 & u & 0 \\ 0 & 0 & 0 & 0 & c+u \end{pmatrix}$$

$$\mathbf{A}_x = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ -\frac{3}{2}u^2-\frac{1}{2}v^2-\frac{1}{2}w^2+\frac{1}{2}\kappa u^2+\frac{1}{2}\kappa v^2+\frac{1}{2}\kappa w^2 & 3u-\kappa u & v-\kappa v & w-\kappa w & -1+\kappa \\ -uv & v & u & 0 & 0 \\ -uw & w & 0 & u & 0 \\ \frac{-2uv^2-2uw^2+2uc^2-u\kappa^2v^2-u\kappa^2w^2+3\kappa uv^2+3\kappa uw^2-2u^3-\kappa^2u^3+3\kappa u^3}{2-2\kappa} & \frac{v^2+w^2-2c^2+3u^2-\kappa v^2-\kappa w^2-5\kappa u^2+2\kappa^2u^2}{2-2\kappa} & uv-\kappa uv & uw-\kappa uw & \kappa u \end{pmatrix}$$

For $u_1 < 0$:

$$\mathbf{A}_x^- = \begin{pmatrix} & & & & \frac{-2c^2u^2} \\ & & & & \\ \frac{-2c^3u^2-2c^3v^2-2c^3w^2+2u^3v^2+2u^3w^2-6uc^2v^2-6uc^2w^2-4\kappa u^3v^2-4\kappa u^3w^2-2\kappa c^3u^2+2cv^2w^2+2\kappa c^3v^2+2\kappa c^3w^2+2uv^2w^2+2\kappa^2u^3v^2+2\kappa^2u^3w^2+6cu^2v^2+6cu^2w^2} & & & & \end{pmatrix}$$

For $u_1 > 0$:

$$\mathbf{A}_x^- = \left(\begin{array}{c} \\ \\ \\ \frac{-2c^3u^2-2c^3v^2-2c^3w^2-2u^3v^2-2u^3w^2-2\kappa c^3u^2-2uc^2v^2-2uc^2w^2-2uv^2w^2-2\kappa^2u^3v^2-2\kappa^2u^3w^2+2cv^2w^2+2\kappa c^3v^2+2\kappa c^3w^2+4\kappa u^3v^2+4\kappa u^3w^2+6cu^2v^2+6cu^2w \end{array} \right. \frac{-10c^2u}$$

Boundary Conditions for the Sea Breeze Model

In the boundary (line) integral we prescribe w_3^{n+1} using a Dirichlet condition and calculate it at each iteration using:

$$w_3^{n+1} = E = \rho T c_v + \tfrac{1}{2}\rho u^2 = w_0 T c_v + \frac{w_1^2 + w_2^2}{2w_0}$$

where $T(t)$ is a known function of time (it changes with the day and night) and also prescribe $w_1^{n+1} = 0$ on the left and right end of the domain and $w_2^{n+1} = 0$ at the top and bottom.

All the surface integrals turn out to be zero. On the top and bottom edges we have $\mathbf{n} = (n_x, n_y) = (0, \pm 1)$ respectively and we prescribe $w_2 = 0$, so we get (remember we do not sum over i):

$$\begin{aligned} \int_{\partial\Omega} \left(\mathbf{A}_x(\mathbf{w}^n)\right)_{ij} w_j \varphi^i \, n_x + \left(\mathbf{A}_y(\mathbf{w}^n)\right)_{ij} w_j \varphi^i \, n_y \, \mathrm{d}x = \\ = \int_{\partial\Omega} \left(\mathbf{f}_x(\mathbf{w}^n)\right)_i \varphi^i \, n_x + \left(\mathbf{f}_y(\mathbf{w}^n)\right)_i \varphi^i \, n_y \, \mathrm{d}x = \\ = \pm \int_{\partial\Omega} \left(\mathbf{f}_y(\mathbf{w}^n)\right)_i \varphi^i \, \mathrm{d}x \end{aligned}$$

where:

$$\mathbf{f}_y = \begin{pmatrix} w_2 \\ \frac{w_2 w_1}{w_0} \\ \frac{w_2^2}{w_0} + p \\ \frac{w_2}{w_0}(w_3 + p) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ p \\ 0 \end{pmatrix}$$

So all the components $i \neq 3$ of the surface integral are zero, and for $i = 3$ the test function φ^3 is not there, because we prescribe the Dirichlet BC $w^3 = 0$, so the surface integral vanishes for all i.

Similarly on the left and right edges we have $\mathbf{n} = (n_x, n_y) = (\pm 1, 0)$ respectively and we prescribe $w_1 = 0$, so we get (remember we do not sum over i):

$$\begin{aligned} \int_{\partial\Omega} \left(\mathbf{A}_x(\mathbf{w}^n)\right)_{ij} w_j \varphi^i \, n_x + \left(\mathbf{A}_y(\mathbf{w}^n)\right)_{ij} w_j \varphi^i \, n_y \, \mathrm{d}x = \\ = \int_{\partial\Omega} \left(\mathbf{f}_x(\mathbf{w}^n)\right)_i \varphi^i \, n_x + \left(\mathbf{f}_y(\mathbf{w}^n)\right)_i \varphi^i \, n_y \, \mathrm{d}x = \\ = \pm \int_{\partial\Omega} \left(\mathbf{f}_x(\mathbf{w}^n)\right)_i \varphi^i \, \mathrm{d}x \end{aligned}$$

where:

$$\mathbf{f}_x = \begin{pmatrix} w_1 \\ \frac{w_1^2}{w_0} + p \\ \frac{w_1 w_2}{w_0} \\ \frac{w_1}{w_0}(w_3 + p) \end{pmatrix} = \begin{pmatrix} 0 \\ p \\ 0 \\ 0 \end{pmatrix}$$

So all the components $i \neq 1$ of the surface integral are zero, and for $i = 1$ the test function φ^1 is not there, because we prescribe the Dirichlet BC $w^1 = 0$, so the surface integral vanishes for all i.

7.3.9 Newton Method

The residual is:

$$F_{i,m}(Y^{n+1}) = \int_\Omega \frac{w_{i,m}(y_m^{n+1}) - w_{i,m}(y^n)}{\tau}\varphi_{i,m} - f_{x,m}(w(y^n))\frac{\partial \varphi_{i,m}}{\partial x} - f_{y,m}(w(y^n))\frac{\partial \varphi_{i,m}}{\partial y} + \delta_{3,m} g \varphi_{i,m}\,\mathrm{d}x\,\mathrm{d}y + \\ - \int_{\partial\Omega} f_{x,m}(w(y^n))\varphi_{i,m}\nu_x + f_{y,m}(w(y^n))\varphi_{i,m}\nu_y\,\mathrm{d}S = 0$$

where $m = 0, 1, 2, 4$ numbers the equations, $i = 1, 2, ..., M$ numbers the finite element basis functions, $N = 4M$, $Y = (y_0^1, y_1^1, y_2^1, y_3^1, y_0^2, y_1^2, ...)$. The Jacobian is:

$$J(Y^n) = \frac{\partial F_{i,m}}{\partial y_{r,s}}(Y^n) = \int_\Omega \frac{\varphi_{r,s}}{\tau}\varphi_{i,m} - A_{x,m,s}(w(y^n))\varphi_{r,s}\frac{\partial \varphi_{i,m}}{\partial x} - A_{y,m,s}(w(y^n))\varphi_{r,s}\frac{\partial \varphi_{i,m}}{\partial y}\,\mathrm{d}x\,\mathrm{d}y \\ + \int_{\partial\Omega} A_{x,m,s}(w(y^n))\varphi_{r,s}\varphi_{i,m}\nu_x + A_{y,m,s}(w(y^n))\varphi_{r,s}\varphi_{i,m}\nu_y\,\mathrm{d}S$$

And the Newton method then is:

$$J(Y^n)\delta Y^{n+1} = -F(Y^n)$$

7.3.10 Older notes

Author: Pavel Solin

Governing Equations and Boundary Conditions

$$\frac{\partial}{\partial t}\begin{pmatrix}\varrho\\U\\W\\\theta\end{pmatrix} + \frac{\partial}{\partial x}\begin{pmatrix}U\\\frac{U^2}{\varrho} + R\theta\\\frac{UW}{\varrho}\\\frac{\theta U}{\varrho}\end{pmatrix} + \frac{\partial}{\partial z}\begin{pmatrix}W\\\frac{UW}{\varrho}\\\frac{W^2}{\varrho} + R\theta\\\frac{\theta W}{\varrho}\end{pmatrix} + \begin{pmatrix}0\\0\\\varrho g\\\frac{R\theta}{c_v}\mathrm{div}\mathbf{v}\end{pmatrix} = \begin{pmatrix}0\\0\\0\\0\end{pmatrix}, \tag{7.3.10.1}$$

where ϱ is the air density, $\mathbf{v} = (u, w)$ is the velocity, $U = \varrho u$, $W = \varrho w$, T is the temperature, $\theta = \varrho T$, and g is the gravitational acceleration constant. We use the perfect gas state equation $p = \varrho RT = R\theta$ for the pressure.

Boundary conditions are prescribed as follows:

- edge a: $\partial\varrho/\partial\nu = 0$, $\partial U/\partial\nu = 0$, $W = 0$, $\theta = \tanh(x) * \sin(\pi t/86400)$
- edges b, c: $\partial\varrho/\partial\nu = 0$, $U = 0$, $\partial W/\partial\nu = 0$, $\partial\theta/\partial\nu = 0$
- edge d: $\partial\varrho/\partial\nu = 0$, $\partial U/\partial\nu = 0$, $W = 0$, $\partial\theta/\partial\nu = 0$

Initial conditions have the form

$$\begin{aligned}
p(z) &= p_0 - 11476\frac{z}{1000} + 529.54\left(\frac{z}{1000}\right)^2 - 9.38\left(\frac{z}{1000}\right)^3,\\
T(z) &= T_0 - 8.3194\frac{z}{1000} + 0.2932\left(\frac{z}{1000}\right)^2 - 0.0109\left(\frac{z}{1000}\right)^3,\\
\varrho(z) &= \frac{p(z)}{RT(z)},\\
\theta(z) &= \varrho(z)T(z),\\
U(z) &= 0,\\
W(z) &= 0.
\end{aligned}$$

Discretization and the Newton's Method

We will use the implicit Euler method in time, i.e.,

$$\frac{\partial \varrho}{\partial t} \approx \frac{\varrho^{n+1} - \varrho^n}{\tau}$$

etc. Let's discuss one equation of (7.3.10.1) at a time:

Continuity equation: The weak formulation of

$$\frac{\varrho^{n+1} - \varrho^n}{\tau} + \frac{\partial U^{n+1}}{\partial x} + \frac{\partial W^{n+1}}{\partial z} = 0$$

reads

$$F_i^\varrho(Y^{n+1}) = \int_\Omega \frac{\varrho^{n+1}}{\tau}\varphi_i^\varrho - \int_\Omega \frac{\varrho^n}{\tau}\varphi_i^\varrho + \int_\Omega \frac{\partial U^{n+1}}{\partial x}\varphi_i^\varrho + \int_\Omega \frac{\partial W^{n+1}}{\partial z}\varphi_i^\varrho = 0 \tag{7.3.10.2}$$

The global coefficient vector Y^{n+1} consists of four parts Y^ϱ, Y^U, Y^W and Y^θ corresponding to the fields ϱ, U, W and θ, respectively. The same holds for the vector function F which consists of four parts F^ϱ, F^U, F^W and F^θ. Thus the global Jacobi matrix will have a four-by-four block structure. We denote

$$\varrho^{n+1} = \sum_{k=1}^{N^\varrho} y_k^\varrho \varphi_k^\varrho, \quad U^{n+1} = \sum_{k=1}^{N^U} y_k^U \varphi_k^U, \quad W^{n+1} = \sum_{k=1}^{N^W} y_k^W \varphi_k^W, \quad \theta^{n+1} = \sum_{k=1}^{N^\theta} y_k^\theta \varphi_k^\theta. \tag{7.3.10.3}$$

It follows from (7.3.10.2) and (7.3.10.3) that

$$\frac{\partial F_i^\varrho}{\partial y_j^\varrho} = \int_\Omega \frac{\varphi_j^\varrho}{\tau}\varphi_i^\varrho, \quad \frac{\partial F_i^\varrho}{\partial y_j^U} = \int_\Omega \frac{\partial \varphi_j^U}{\partial x}\varphi_i^\varrho, \quad \frac{\partial F_i^\varrho}{\partial y_j^W} = \int_\Omega \frac{\partial \varphi_j^W}{\partial z}\varphi_i^\varrho, \quad \frac{\partial F_i^\varrho}{\partial y_j^W} = 0.$$

First momentum equation: The second equation of (7.3.10.1) has the form

$$\frac{\partial U}{\partial t} + \frac{2U}{\varrho}\frac{\partial U}{\partial x} - \frac{U^2}{\varrho^2}\frac{\partial \varrho}{\partial x} + R\frac{\partial \theta}{\partial x} + \frac{W}{\varrho}\frac{\partial U}{\partial z} + \frac{U}{\varrho}\frac{\partial W}{\partial z} - \frac{UW}{\varrho^2}\frac{\partial \varrho}{\partial z} = 0.$$

After applying the implicit Euler method, we obtain

$$\frac{\partial U^{n+1}}{\tau} - \frac{\partial U^n}{\tau} + \frac{2U^{n+1}}{\varrho^{n+1}}\frac{\partial U^{n+1}}{\partial x} - \frac{(U^{n+1})^2}{(\varrho^{n+1})^2}\frac{\partial \varrho^{n+1}}{\partial x} + R\frac{\partial \theta^{n+1}}{\partial x}$$

$$+\frac{W^{n+1}}{\varrho^{n+1}}\frac{\partial U^{n+1}}{\partial z} + \frac{U^{n+1}}{\varrho^{n+1}}\frac{\partial W^{n+1}}{\partial z} - \frac{U^{n+1}W^{n+1}}{(\varrho^{n+1})^2}\frac{\partial \varrho^{n+1}}{\partial z} = 0.$$

Thus we obtain

$$\frac{\partial F_i^U}{\partial y_j^\varrho} = -\int_\Omega \frac{2U}{\varrho^2}\frac{\partial U}{\partial x}\varphi_j^\varrho\varphi_i^U - \int_\Omega U^2\left[(-2)\frac{1}{\varrho^3}\frac{\partial \varrho}{\partial x}\varphi_j^\varrho + \frac{1}{\varrho^2}\frac{\partial \varphi_j^\varrho}{\partial x}\right]\varphi_i^U$$

$$+\int_\Omega \frac{W}{\varrho^2}\frac{\partial U}{\partial z}(-1)\varphi_j^\varrho\varphi_i^U + \int_\Omega \frac{U}{\varrho^2}\frac{\partial W}{\partial z}(-1)\varphi_j^\varrho\varphi_i^U - \int_\Omega UW\left[(-2)\frac{1}{\varrho^3}\frac{\partial \varrho}{\partial z}\varphi_j^\varrho + \frac{1}{\varrho^2}\frac{\partial \varphi_j^\varrho}{\partial z}\right]\varphi_i^U.$$

Analogously,

$$\frac{\partial F_i^U}{\partial y_j^U} = \int_\Omega \frac{\varphi_j^U}{\tau}\varphi_i^U + \int_\Omega \frac{2}{\varrho}\left[\frac{\partial U}{\partial x}\varphi_j^U + U\frac{\partial \varphi_j^U}{\partial x}\right]\varphi_i^U - \int_\Omega \frac{2U}{\varrho^2}\frac{\partial \varrho}{\partial x}\varphi_j^U\varphi_i^U$$

$$+\int_\Omega \frac{W}{\varrho}\frac{\partial \varphi_j^U}{\partial z}\varphi_i^U + \int_\Omega \frac{1}{\varrho}\frac{\partial W}{\partial z}\varphi_j^U\varphi_i^U - \int_\Omega \frac{W}{\varrho^2}\frac{\partial \varrho}{\partial z}\varphi_j^U\varphi_i^U,$$

$$\frac{\partial F_i^U}{\partial y_j^W} = \int_\Omega \frac{1}{\varrho}\frac{\partial U}{\partial z}\varphi_j^W\varphi_i^U + \int_\Omega \frac{U}{\varrho}\frac{\partial \varphi_j^W}{\partial z}\varphi_i^U - \int_\Omega \frac{U}{\varrho^2}\frac{\partial \varrho}{\partial z}\varphi_j^W\varphi_i^U,$$

$$\frac{\partial F_i^U}{\partial y_j^\theta} = \int_\Omega R\frac{\partial \varphi_j^\theta}{\partial x}\varphi_i^U.$$

Second momentum equation: The third equation of (7.3.10.1) reads

$$\frac{\partial W}{\partial t} + \frac{W}{\varrho}\frac{\partial U}{\partial x} + \frac{U}{\varrho}\frac{\partial W}{\partial x} - \frac{UW}{\varrho^2}\frac{\partial \varrho}{\partial x} + \frac{2W}{\varrho}\frac{\partial W}{\partial z} - \frac{W^2}{\varrho^2}\frac{\partial \varrho}{\partial x} + R\frac{\partial \theta}{\partial z} + \varrho g = 0.$$

After applying the implicit Euler method, we obtain

$$\frac{\partial W^{n+1}}{\tau} - \frac{\partial W^n}{\tau} + \frac{W^{n+1}}{\varrho^{n+1}}\frac{\partial U^{n+1}}{\partial x} + \frac{U^{n+1}}{\varrho^{n+1}}\frac{\partial W^{n+1}}{\partial x} - \frac{U^{n+1}W^{n+1}}{(\varrho^{n+1})^2}\frac{\partial \varrho^{n+1}}{\partial x}$$

$$+\frac{2W^{n+1}}{\varrho^{n+1}}\frac{\partial W^{n+1}}{\partial z} - \frac{(W^{n+1})^2}{(\varrho^{n+1})^2}\frac{\partial \varrho^{n+1}}{\partial x} + R\frac{\partial \theta^{n+1}}{\partial z} + \varrho^{n+1}g = 0.$$

Thus we obtain

$$\frac{\partial F_i^W}{\partial y_j^\varrho} = +\int_\Omega \frac{W}{\varrho^2}\frac{\partial U}{\partial x}(-1)\varphi_j^\varrho\varphi_i^W + \int_\Omega \frac{U}{\varrho^2}\frac{\partial W}{\partial x}(-1)\varphi_j^\varrho\varphi_i^W - \int_\Omega \frac{2W}{\varrho^2}\frac{\partial W}{\partial x}\varphi_j^\varrho\varphi_i^W$$

$$-\int_\Omega UW\left[(-2)\frac{1}{\varrho^3}\frac{\partial \varrho}{\partial x}\varphi_j^\varrho + \frac{1}{\varrho^2}\frac{\partial \varphi_j^\varrho}{\partial x}\right]\varphi_i^W - \int_\Omega W^2\left[(-2)\frac{1}{\varrho^3}\frac{\partial \varrho}{\partial z}\varphi_j^\varrho + \frac{1}{\varrho^2}\frac{\partial \varphi_j^\varrho}{\partial z}\right]\varphi_i^W + \int_\Omega g\varphi_j^\varrho\varphi_i^W.$$

Analogously,

$$\frac{\partial F_i^W}{\partial y_j^U} = \int_\Omega \frac{W}{\varrho}\frac{\partial \varphi_j^U}{\partial x}\varphi_i^W + \int_\Omega \frac{1}{\varrho}\frac{\partial W}{\partial x}\varphi_j^U\varphi_i^W - \int_\Omega \frac{W}{\varrho^2}\frac{\partial \varrho}{\partial x}\varphi_j^U\varphi_i^W,$$

$$\frac{\partial F_i^W}{\partial y_j^W} = \int_\Omega \frac{\varphi_j^W}{\tau}\varphi_i^W + \int_\Omega \frac{1}{\varrho}\frac{\partial U}{\partial x}\varphi_j^W\varphi_i^W + \int_\Omega \frac{U}{\varrho}\frac{\partial \varphi_j^W}{\partial x}\varphi_i^W - \int_\Omega \frac{U}{\varrho^2}\frac{\partial \varrho}{\partial x}\varphi_j^W\varphi_i^W$$

$$+\int_\Omega \frac{2}{\varrho}\left[\frac{\partial W}{\partial z}\varphi_j^W + W\frac{\partial \varphi_j^W}{\partial z}\right]\varphi_i^W - \int_\Omega \frac{2W}{\varrho^2}\frac{\partial \varrho}{\partial z}\varphi_j^W\varphi_i^W,$$

$$\frac{\partial F_i^W}{\partial y_j^\theta} = \int_\Omega R\frac{\partial \varphi_j^\theta}{\partial z}\varphi_i^W.$$

Internal energy equation: The last equation of (7.3.10.1) has the form

$$\frac{\partial \theta}{\partial t} + \mathrm{div}(\theta\mathbf{v}) + \frac{R\theta}{c_v}\mathrm{div}\mathbf{v} = 0$$

where $\theta = \varrho T$. This can be written equivalently as

$$\frac{\partial \theta}{\partial t} + \nabla\theta\cdot\mathbf{v} + \gamma\theta\mathrm{div}\mathbf{v} = 0.$$

Written in terms of single derivatives, this is

$$\frac{\partial \theta}{\partial t} + \frac{\partial \theta}{\partial x}\frac{U}{\varrho} + \frac{\partial \theta}{\partial z}\frac{W}{\varrho} + \gamma\theta\frac{\partial}{\partial x}\left(\frac{U}{\varrho}\right) + \gamma\theta\frac{\partial}{\partial z}\left(\frac{W}{\varrho}\right) = 0,$$

i.e.,

$$\frac{\partial \theta}{\partial t} + \frac{\partial \theta}{\partial x}\frac{U}{\varrho} + \frac{\partial \theta}{\partial z}\frac{W}{\varrho} + \gamma\frac{\theta}{\varrho}\frac{\partial U}{\partial x} - \gamma\frac{\theta U}{\varrho^2}\frac{\partial \varrho}{\partial x} + \gamma\frac{\theta}{\varrho}\frac{\partial W}{\partial z} - \gamma\frac{\theta W}{\varrho^2}\frac{\partial \varrho}{\partial z} = 0.$$

Weak formulation:

$$F_i^\theta(Y) = \int_\Omega \frac{\theta^{n+1}}{\tau}\varphi_i^\theta - \int_\Omega \frac{\theta^n}{\tau}\varphi_i^\theta + \int_\Omega \frac{\partial\theta^{n+1}}{\partial x}\frac{U^{n+1}}{\varrho^{n+1}}\varphi_i^\theta + \int_\Omega \frac{\partial\theta^{n+1}}{\partial z}\frac{W^{n+1}}{\varrho^{n+1}}\varphi_i^\theta$$

$$+ \int_\Omega \gamma\frac{\theta^{n+1}}{\varrho^{n+1}}\frac{\partial U^{n+1}}{\partial x}\varphi_i^\theta - \int_\Omega \gamma\frac{\theta^{n+1}U^{n+1}}{(\varrho^{n+1})^2}\frac{\partial\varrho^{n+1}}{\partial x}\varphi_i^\theta + \int_\Omega \gamma\frac{\theta^{n+1}}{\varrho^{n+1}}\frac{\partial W^{n+1}}{\partial z}\varphi_i^\theta - \int_\Omega \gamma\frac{\theta^{n+1}W^{n+1}}{(\varrho^{n+1})^2}\frac{\partial\varrho^{n+1}}{\partial z}\varphi_i^\theta = 0.$$

For the derivatives of the weak form we obtain:

$$\frac{\partial F_i^\theta}{\partial y_j^\varrho} = -\int_\Omega \frac{\partial\theta}{\partial x}\frac{U}{\varrho^2}\varphi_j^\varrho\varphi_i^\theta - \int_\Omega \frac{\partial\theta}{\partial z}\frac{W}{\varrho^2}\varphi_j^\varrho\varphi_i^\theta - \int_\Omega \gamma\frac{\theta}{\varrho^2}\frac{\partial U}{\partial x}\varphi_j^\varrho\varphi_i^\theta - \int_\Omega \gamma\frac{\theta}{\varrho^2}\frac{\partial W}{\partial z}\varphi_j^\varrho\varphi_i^\theta$$

$$+ \int_\Omega 2\gamma\frac{\theta U}{\varrho^3}\frac{\partial\varrho}{\partial x}\varphi_j^\varrho\varphi_i^\theta - \int_\Omega \gamma\frac{\theta U}{\varrho^2}\frac{\varphi_j^\varrho}{\partial x}\varphi_i^\theta + \int_\Omega 2\gamma\frac{\theta W}{\varrho^3}\frac{\partial\varrho}{\partial z}\varphi_j^\varrho\varphi_i^\theta - \int_\Omega \gamma\frac{\theta W}{\varrho^2}\frac{\varphi_j^\varrho}{\partial z}\varphi_i^\theta.$$

$$\frac{\partial F_i^\theta}{\partial y_j^U} = \int_\Omega \frac{\partial\theta}{\partial x}\frac{1}{\varrho}\varphi_j^U\varphi_i^\theta + \int_\Omega \gamma\frac{\theta}{\varrho}\frac{\varphi_j^U}{\partial x}\varphi_i^\theta - \int_\Omega \gamma\frac{\theta}{\varrho^2}\frac{\partial\varrho}{\partial x}\varphi_j^U\varphi_i^\theta.$$

$$\frac{\partial F_i^\theta}{\partial y_j^W} = \int_\Omega \frac{\partial\theta}{\partial z}\frac{1}{\varrho}\varphi_j^W\varphi_i^\theta + \int_\Omega \gamma\frac{\theta}{\varrho}\frac{\varphi_j^W}{\partial z}\varphi_i^\theta - \int_\Omega \gamma\frac{\theta}{\varrho^2}\frac{\partial\varrho}{\partial z}\varphi_j^W\varphi_i^\theta.$$

$$\frac{\partial F_i^\theta}{\partial y_j^\theta} = \int_\Omega \frac{1}{\tau}\varphi_j^\theta\varphi_i^\theta + \int_\Omega \frac{U}{\varrho}\frac{\varphi_j^\theta}{\partial x}\varphi_i^\theta + \int_\Omega \frac{W}{\varrho}\frac{\varphi_j^\theta}{\partial z}\varphi_i^\theta$$

$$+ \int_\Omega \frac{\gamma}{\varrho}\frac{\partial U}{\partial x}\varphi_j^\theta\varphi_i^\theta + \int_\Omega \frac{\gamma}{\varrho}\frac{\partial W}{\partial z}\varphi_j^\theta\varphi_i^\theta - \int_\Omega \frac{\gamma U}{\varrho^2}\frac{\partial\varrho}{\partial x}\varphi_j^\theta\varphi_i^\theta - \int_\Omega \frac{\gamma W}{\varrho^2}\frac{\partial\varrho}{\partial z}\varphi_j^\theta\varphi_i^\theta.$$

CHAPTER

EIGHT

QUANTUM FIELD THEORY AND QUANTUM MECHANICS

8.1 Introduction

The aim of these (work in progress) notes is to use the Standard Model of particle physics to derive all equations in quantum mechanics (and quantum field theory) that we need for our research.

We start by deriving the electroweak Standard Model from the $SU(2) \times U(1)$ symmetry and couple other (standard) assumptions in the quantum field theory. After that, we only want to derive things and make nonrelativistic limits or other approximations in order to derive everything else in quantum mechanics. In particular we show how to derive the Dirac and Schrödinger equations (as a low energy limit). We then show some particular ways to solve those equations, like perturbation theory, scattering theory, ...

The goal is to have a complete theory on about 30 or 40 pages and then lots of examples (arbitrarily long), that use the theory (but do not develop new ideas), so that one can learn how the theory works from the examples. For instance, one can ask "why is there the term $(\mathbf{p} - e\mathbf{A})^2$ in the Schrödinger equation for electromagnetic field, why this and not something else?" or "why is there the $\boldsymbol{\sigma} \cdot \mathbf{B}$ term in the Pauli equation?", to find the answer, one just finds the Pauli equation in the theory and then looks at the derivation, so in this case one quickly finds that it follows from the minimal coupling in QED, e.g. it's the easiest way how electron-foton interaction can be coupled, e.g. the $U(1)$ symmetry. Nice thing about QFT is that one can find really nice geometrical reasons why things are that way and not some other way (just open any advance book on QFT), but the problem is that basically nowhere is some easy (but correct) translation of those results to regular QM, so that everything fits into just couple dozens pages, so that it can serve as a reference.

The advantage of this top-down approach is that it is easy to see where things come from and also to understand exactly what approximations one is using when dealing with any equation in QM. However, as is well-known in physics, to be a good physicist one has to understand all the approaches, e.g. both top-down and bottom-up and all other approaches to QM and QFT, because there are no two approaches that would be 100% equivalent, so one has to use the right approach for the particular problem. So these notes do not aspire to be the right way to teach QM, but rather to serve as a reference to get quickly oriented and to find the equations to start from.

8.2 Standard Model

8.2.1 Electroweak Standard Model

Lagrangian with a global $SU(2) \times U(1)$ symmetry:

$$\mathcal{L} = i\bar{L}^{(l)}\gamma_\mu \partial^\mu L^{(l)} + i\bar{l}_R \gamma_\mu \partial^\mu l_R + \tfrac{1}{2}\partial_\mu \Phi^* \partial^\mu \Phi - m^2 \Phi^* \Phi - \frac{1}{4}\lambda(\Phi^* \Phi)^2 - h_e \bar{L}^{(l)} \Phi e_R - \text{h.c.}$$

where $l = e, \mu, \tau$ and $a = 1, 2$, $l_{L,R} = \frac{1}{2}(1 \mp \gamma_5) l$ and

$$L^{(l)} = \begin{pmatrix} \nu_{(l)L} \\ l_L \end{pmatrix}$$

Local $SU(2) \times U(1)$ symmetry: This consists of two things. First changing the partial derivatives to covariant ones:

$$\partial^\mu \to D^\mu = \partial^\mu - \frac{i}{2} g \tau_k A_k^\mu - \frac{i}{2} g' Y B^\mu$$

and second adding the kinetic terms

$$-\frac{1}{4} F_{\mu\nu}^a F^{a\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}$$

of the vector gauge particles to the lagrangian.

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g \epsilon^{abc} A_\mu^b A_\nu^c$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$$

$$\Phi = e^{\frac{i}{v}\pi^a(x)\tau^a} \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(v + H(x)) \end{pmatrix}$$

This breaks the gauge invariance. The $\partial^\mu \pi^a$ are going to be added to A_μ^a so we can set $\pi_a = 0$ now.

Higgs Terms

$$\mathcal{L}_{Higgs} = \tfrac{1}{2} \partial_\mu \Phi^* \partial^\mu \Phi - m^2 \Phi^* \Phi - \frac{1}{4} \lambda (\Phi^* \Phi)^2$$

Plugging in the covariant derivatives and Φ in U-gauge (symmetry breaking):

$$\mathcal{L}_{Higgs} = \frac{1}{2} \Phi^+ (\overleftarrow{\partial}_\mu + i g A_\mu^a \frac{\tau^a}{2} + i g' Y B_\mu)(\overrightarrow{\partial}^\mu + i g A^{a\mu} \frac{\tau^a}{2} + i g' Y B^\mu) \Phi - \lambda (\Phi^+ \Phi - \frac{v^2}{2})^2 =$$

$$= \Phi_U^+ (\overleftarrow{\partial}_\mu + i g A_\mu^a \frac{\tau^a}{2} + i g' Y B_\mu)(\overrightarrow{\partial}^\mu + i g A^{a\mu} \frac{\tau^a}{2} + i g' Y B\mu) \Phi_U - \lambda (\Phi_U^+ \Phi_U - \frac{v^2}{2})^2 =$$

$$= \frac{1}{2} \partial_\mu H \partial^\mu H - \lambda v^2 H^2 - \lambda v H^3 - \frac{1}{4} \lambda H^4 +$$

$$+ \frac{1}{8} (v + H)^2 \left(2 g^2 \frac{A_\mu^1 + i A_\mu^2}{\sqrt{2}} \frac{A^{1\mu} - i A^{2\mu}}{\sqrt{2}} + (g^2 + 4 Y^2 g'^2) \frac{g A_\mu^3 - 2 Y g' B_\mu}{\sqrt{g^2 + 4 Y^2 g'^2}} \frac{g A^{3\mu} - 2 Y g' B^\mu}{\sqrt{g^2 + 4 Y^2 g'^2}} \right) =$$

$$= \frac{1}{2} \partial_\mu H \partial^\mu H - \lambda v^2 H^2 - \lambda v H^3 - \frac{1}{4} \lambda H^4 + \frac{1}{8} (v + H)^2 \left(2 g^2 W_\mu^- W^{+\mu} + \frac{g^2}{\cos^2 \theta_W} Z_\mu Z^\mu \right) =$$

$$= \frac{1}{2} \partial_\mu H \partial^\mu H - \lambda v^2 H^2 + \frac{1}{4} g^2 v^2 W_\mu^- W^{+\mu} + \frac{g^2 v^2}{8 \cos^2 \theta_W} Z_\mu Z^\mu - \lambda v H^3 - \frac{1}{4} \lambda H^4 +$$

$$+ \frac{1}{2} v g^2 W_\mu^- W^{+\mu} H + \frac{g^2}{4 \cos \theta_W} v Z_\mu Z^\mu H + \frac{1}{4} g^2 W_\mu^- W^{+\mu} H^2 + \frac{g^2}{8 \cos \theta_W} Z_\mu Z^\mu H^2$$

Where we put

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (A_\mu^1 \mp i A_\mu^2)$$

$$Z_\mu = \frac{g}{\sqrt{g^2 + 4 Y^2 g'^2}} A_\mu^3 - \frac{2 Y g'}{\sqrt{g^2 + 4 Y^2 g'^2}} B_\mu$$

we defined θ_W by the relation

$$\cos \theta_W = \frac{g}{\sqrt{g^2 + 4 Y^2 g'^2}}$$

so that the expressions simplify a bit, e.g. we now get:

$$\sin\theta_W = \frac{2Yg'}{\sqrt{g^2+4Y^2g'^2}}$$

$$Z_\mu = \cos\theta_W A^3_\mu - \sin\theta_W B_\mu$$

$$g^2+4Y^2g'^2 = \frac{g^2}{\cos^2\theta_W}$$

Yukawa terms

$$\mathcal{L}_{Yukawa} = -h_e\bar{L}\Phi e_R - \text{h.c.} = -h_e\bar{L}\Phi_U e_R - \text{h.c.} =$$
$$= -\frac{1}{\sqrt{2}}h_e(v+H)(\bar{e}_L e_R + \bar{e}_R e_L) = -\frac{1}{\sqrt{2}}h_e(v+H)\bar{e}e =$$
$$= -\frac{1}{\sqrt{2}}h_e v\bar{e}e - \frac{1}{\sqrt{2}}h_e\bar{e}eH$$

The term $\bar{L}\Phi e_R$ is $U(1)$ (hypercharge) invariant, so

$$-Y_L + Y + Y_R = 0$$

Leptonic Terms

$$\mathcal{L} = i\bar{L}\gamma^\mu\partial_\mu L + i\bar{e}_R\gamma^\mu\partial_\mu e_R \rightarrow$$
$$\rightarrow i\bar{L}\gamma^\mu(\partial_\mu - igA^a_\mu\frac{\tau^a}{2} - ig'Y_L B_\mu)L + i\bar{e}_R\gamma^\mu(\partial_\mu - ig'Y_R B_\mu)e_R =$$
$$= i\bar{L}\gamma^\mu\partial_\mu L + i\bar{e}_R\gamma^\mu\partial_\mu e_R + g\bar{L}\gamma^\mu\frac{\tau^a}{2}LA^a_\mu + g'Y_L\bar{L}\gamma^\mu LB_\mu + g'Y_R\bar{e}_R\gamma^\mu e_R B_\mu =$$
$$= i\bar{L}\gamma^\mu\partial_\mu L + i\bar{e}_R\gamma^\mu\partial_\mu e_R + \frac{g}{\sqrt{2}}(\bar{\nu}_L\gamma^\mu e_L W^+_\mu + \text{h.c.}) + \frac{1}{2}g\bar{L}\gamma^\mu\tau^3 LA^3_\mu + g'Y_L\bar{L}\gamma^\mu LB_\mu + g'Y_R\bar{e}_R\gamma^\mu e_R B_\mu =$$
$$= i\bar{\nu}_L\gamma^\mu\partial_\mu\nu_L + i\bar{e}\gamma^\mu\partial_\mu e + \frac{g}{\sqrt{2}}(\bar{\nu}_L\gamma^\mu e_L W^+_\mu + \text{h.c.}) + \frac{1}{2}g\bar{\nu}_L\gamma^\mu\nu_L A^3_\mu - \frac{1}{2}g\bar{e}_L\gamma^\mu e_L A^3_\mu$$
$$+g'Y_L\bar{\nu}_L\gamma^\mu\nu_L B_\mu + g'Y_L\bar{e}_L\gamma^\mu e_L B_\mu + g'Y_R\bar{e}_R\gamma^\mu e_R B_\mu =$$
$$= i\bar{\nu}_L\gamma^\mu\partial_\mu\nu_L + i\bar{e}\gamma^\mu\partial_\mu e + \frac{g}{\sqrt{2}}(\bar{\nu}_L\gamma^\mu e_L W^+_\mu + \text{h.c.})$$
$$+\left[(\tfrac{1}{2}g\sin\theta_W + Y_L g'\cos\theta_W)\bar{\nu}_L\gamma^\mu\nu_L + (-\tfrac{1}{2}g\sin\theta_W + Y_L g'\cos\theta_W)\bar{e}_L\gamma^\mu e_L + Y_R g'\cos\theta_W\bar{e}_R\gamma^\mu e_R\right]A_\mu$$
$$+\left[(\tfrac{1}{2}g\cos\theta_W - Y_L g'\sin\theta_W)\bar{\nu}_L\gamma^\mu\nu_L + (-\tfrac{1}{2}g\cos\theta_W - Y_L g'\sin\theta_W)\bar{e}_L\gamma^\mu e_L - 2Y_L g'\sin\theta_W\bar{e}_R\gamma^\mu e_R\right]Z_\mu$$

Where we substituted new fields Z_μ and A_μ for the old ones A^3_μ and B_μ using the relation:

$$Z_\mu = \cos\theta_W A^3_\mu - \sin\theta_W B_\mu$$

$$A_\mu = \sin\theta_W A^3_\mu + \cos\theta_W B_\mu$$

The angle θ_W must be the same as in the Higgs sector, so that the field Z_μ is the same. We now need to make the following requirement in order to proceed further:

$$Y = -Y_L$$

This follows for example by requiring that neutrinos have zero charge, i.e. setting $\frac{1}{2}g\sin\theta_W + Y_L g'\cos\theta_W = 0$ and substituting for θ_W from the definition (see the Higgs terms), from which one gets $Y = -Y_L$. From $-Y_L + Y + Y_R = 0$ we now get

$$Y_R = 2Y_L$$

it now follows:

$$\frac{1}{2}g\sin\theta_W + Y_L g'\cos\theta_W = 0$$

$$-\frac{1}{2}g\sin\theta_W + Y_L g'\cos\theta_W = -g\sin\theta_W$$

$$Y_R g'\cos\theta_W = -g\sin\theta_W$$

$$\tan\theta_W = -2Y_L\frac{g'}{g}$$

and the Lagrangian can be further simplified:

$$\mathcal{L} = i\bar{\nu}_L\gamma^\mu\partial_\mu\nu_L + i\bar{e}\gamma^\mu\partial_\mu e + \frac{g}{\sqrt{2}}(\bar{\nu}_L\gamma^\mu e_L W^+_\mu + \text{h.c.})$$

$$-g\sin\theta_W(\bar{e}_L\gamma^\mu e_L + \bar{e}_R\gamma^\mu e_R)A_\mu$$

$$+\frac{g}{\cos\theta_W}\left[\tfrac{1}{2}\bar{\nu}_L\gamma^\mu\nu_L + (-\tfrac{1}{2} + \sin^2\theta_W)\bar{e}_L\gamma^\mu e_L + \sin^2\theta_W\bar{e}_R\gamma^\mu e_R\right]Z_\mu =$$

$$= i\bar{\nu}_L\gamma^\mu\partial_\mu\nu_L + i\bar{e}\gamma^\mu\partial_\mu e + \frac{g}{2\sqrt{2}}(\bar{\nu}\gamma^\mu(1-\gamma_5)eW^+_\mu + \text{h.c.}) - g\sin\theta_W\bar{e}\gamma^\mu eA_\mu$$

$$+\frac{g}{2\cos\theta_W}\left[\bar{\nu}\gamma^\mu(1-\gamma_5)\nu + \bar{e}\gamma^\mu(-\tfrac{1}{2} + 2\sin^2\theta_W + \tfrac{1}{2}\gamma_5)e\right]Z_\mu$$

Where we used the relations $\bar{\nu}_L\gamma^\mu e_L = \frac{1}{2}\bar{\nu}\gamma^\mu(1-\gamma_5)e$ and $\bar{\nu}_R\gamma^\mu e_R = \frac{1}{2}\bar{\nu}\gamma^\mu(1+\gamma_5)e$.

Gauge terms

$$\mathcal{L}_{Gauge} = -\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} =$$

$$= -\frac{1}{4}(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g\epsilon^{abc}A^b_\mu A^c_\nu)(\partial^\mu A^{a\nu} - \partial^\nu A^{a\mu} + g\epsilon^{ajk}A^{j\mu}A^{k\nu}) - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} =$$

$$= -\frac{1}{4}\partial_\mu A^a_\nu\partial^\mu A^{a\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{2}(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu)g\epsilon^{abc}A^{b\mu}A^{c\nu} - \frac{1}{4}g^2\epsilon^{abc}\epsilon^{ajk}A^b_\mu A^c_\nu A^{k\mu}A^{l\nu} =$$

$$= -\frac{1}{2}W^-_{\mu\nu}W^{+\mu\nu} - \frac{1}{4}A_{\mu\nu}A^{\mu\nu} - \frac{1}{4}Z_{\mu\nu}Z^{\mu\nu} - g[(\partial_\mu A^1_\nu - \partial_\nu A^1_\mu)A^{2\mu}A^{3\nu} + \text{cycl. perm. (123)}]$$

$$-\frac{1}{4}g^2[(A^a_\mu A^{a\mu})(A^b_\nu A^{b\nu}) - (A^a_\mu A^a_\nu)(A^{b\mu}A^{b\nu})] =$$

$$= -\frac{1}{2}W^-_{\mu\nu}W^{+\mu\nu} - \frac{1}{4}A_{\mu\nu}A^{\mu\nu} - \frac{1}{4}Z_{\mu\nu}Z^{\mu\nu} - g[A^1_\mu A^2_\nu\overleftrightarrow{\partial}^\mu A^{3\nu} + \text{cycl. perm. (123)}]$$

$$-\frac{1}{4}g^2[(A^a_\mu A^{a\mu})(A^b_\nu A^{b\nu}) - (A^a_\mu A^a_\nu)(A^{b\mu}A^{b\nu})] =$$

$$= -\frac{1}{2}W^-_{\mu\nu}W^{+\mu\nu} - \frac{1}{4}A_{\mu\nu}A^{\mu\nu} - \frac{1}{4}Z_{\mu\nu}Z^{\mu\nu} - ig(W^0_\mu W^-_\nu\overleftrightarrow{\partial}^\mu W^{+\nu} + \text{cycl. perm. (0-+)})$$

$$-g^2[\tfrac{1}{2}(W^+_\mu W^{-\mu})^2 - \tfrac{1}{2}(W^+_\mu W^{+\mu})(W^-_\nu W^{-\nu}) + (W^0_\mu W^{0\mu})(W^+_\nu W^{-\nu}) - (W^-_\mu W^+_\nu)(W^{0\mu}W^{0\nu}) =$$

$$= -\frac{1}{2}W^-_{\mu\nu}W^{+\mu\nu} - \frac{1}{4}A_{\mu\nu}A^{\mu\nu} - \frac{1}{4}Z_{\mu\nu}Z^{\mu\nu} + \mathcal{L}_{WW\gamma} + L_{WWZ} + L_{WW\gamma\gamma} + L_{WWWW} + L_{WWZZ} + L_{WWZ\gamma}$$

Where $W^0_\mu = A^3_\mu = \cos\theta_W Z_\mu + \sin\theta_W A_\mu$ and:

$$\mathcal{L}_{WW\gamma} = -ig\sin\theta_W(A_\mu W^-_\nu \overleftrightarrow{\partial}^\mu W^{+\nu} + \text{cycl. perm. } (A\ W^-\ W^+))$$

$$\mathcal{L}_{WWZ} = -ig\cos\theta_W(Z_\mu W^-_\nu \overleftrightarrow{\partial}^\mu W^{+\nu} + \text{cycl. perm. } (Z\ W^-\ W^+))$$

$$\mathcal{L}_{WW\gamma\gamma} = -g^2\sin^2\theta_W(W^-_\mu W^{+\mu}A_\nu A^\nu - W^-_\mu A^\mu W^+_\nu A^\nu)$$

$$\mathcal{L}_{WWWW} = \tfrac{1}{2}g^2(W^-_\mu W^{-\mu}W^+_\nu W^{+\nu} - W^-_\mu W^{+\mu}W^-_\nu W^{+\nu})$$

$$\mathcal{L}_{WWZZ} = -g^2\cos^2\theta_W(W^-_\mu W^{+\mu}Z_\nu Z^\nu - W^-_\mu Z^\mu W^+_\nu Z^\nu)$$

$$\mathcal{L}_{WWZ\gamma} = g^2\sin\theta_W\cos\theta_W(-2W^-_\mu W^{+\mu}A_\nu Z^\nu + W^-_\mu Z^\mu W^+_\nu A^\nu + W^-_\mu A^\mu W^+_\nu Z^\nu)$$

GWS Lagrangian

Plugging everything together we get the GWS Lagrangian:

$$\mathcal{L} = \frac{1}{2}\partial_\mu H\partial^\mu H - \lambda v^2H^2 + \frac{1}{4}g^2v^2W^-_\mu W^{+\mu} + \frac{g^2v^2}{8\cos^2\theta_W}Z_\mu Z^\mu - \lambda vH^3 - \frac{1}{4}\lambda H^4 +$$

$$+\frac{1}{2}vg^2W^-_\mu W^{+\mu}H + \frac{g^2}{4\cos\theta_W}vZ_\mu Z^\mu H + \frac{1}{4}g^2W^-_\mu W^{+\mu}H^2 + \frac{g^2}{8\cos\theta_W}Z_\mu Z^\mu H^2$$

$$-\frac{1}{\sqrt{2}}h_e v\bar{e}e - \frac{1}{\sqrt{2}}h_e\bar{e}eH$$

$$-\frac{1}{2}W^-_{\mu\nu}W^{+\mu\nu} - \frac{1}{4}A_{\mu\nu}A^{\mu\nu} - \frac{1}{4}Z_{\mu\nu}Z^{\mu\nu} + \mathcal{L}_{WW\gamma} + L_{WWZ} + L_{WW\gamma\gamma} + L_{WWWW} + L_{WWZZ} + L_{WWZ\gamma}$$

$$+i\bar{\nu}_L\gamma^\mu\partial_\mu\nu_L + i\bar{e}\gamma^\mu\partial_\mu e + \frac{g}{2\sqrt{2}}(\bar{\nu}\gamma^\mu(1-\gamma_5)eW^+_\mu + \text{h.c.}) - g\sin\theta_W\bar{e}\gamma^\mu eA_\mu$$

$$+\frac{g}{2\cos\theta_W}\left[\bar{\nu}\gamma^\mu(1-\gamma_5)\nu + \bar{e}\gamma^\mu(-\tfrac{1}{2} + 2\sin^2\theta_W + \tfrac{1}{2}\gamma_5)e\right]Z_\mu$$

$$+(e,\nu_e,h_e \leftrightarrow \mu,\nu_\mu,h_\mu) + (e,\nu_e,h_e \leftrightarrow \tau,\nu_\tau,h_\tau)$$

The free parameters are g, θ_W, v, λ, h_e, h_μ, h_τ.

Particle Masses

The particle masses are deduced from the terms

$$\mathcal{L} = -\frac{1}{2}m^2_H H^2 + m^2_W W^-_\mu W^{+\mu} + \frac{1}{2}m^2_Z Z_\mu Z^\mu - m_e\bar{e}e + \cdots$$

comparing to the above:

$$\mathcal{L} = -\lambda v^2H^2 + \frac{1}{4}g^2v^2W^-_\mu W^{+\mu} + \frac{g^2v^2}{8\cos^2\theta_W}Z_\mu Z^\mu - \frac{1}{\sqrt{2}}h_e v\bar{e}e + \cdots$$

we get

$$
\begin{aligned}
m_W &= \tfrac{1}{2} g v \\
m_Z = \frac{g v}{2 \cos\theta_W} &= \frac{m_W}{\cos\theta_W} \\
m_H &= v\sqrt{2\lambda} \\
m_e &= \frac{1}{\sqrt{2}} h_e v \\
m_\mu &= \frac{1}{\sqrt{2}} h_\mu v \\
m_\tau &= \frac{1}{\sqrt{2}} h_\tau v
\end{aligned}
$$

Note that those are the bare masses (e.g. in order to obtain the real mesaured masses of the particles, one has to renormalize them by calculating the higher order corrections given by the loop diagrams).

Parameters of the Standard Model

The free parameters are g, θ_W, v, λ, then three masses of the charged leptons h_e, h_μ, h_τ, six quark masses and four parameters of the CKM mixing matrix, which gives 4 + 3 + 6 + 4 = 17 free parameters (if one allows for three neutrino masses and the corresponding four mixings parameters, one gets 17 + 3 + 4 = 24 free parameters).

They can be traded for other physical parameters (see below), but their numerical values are not predicted by the theory, so they have to be measured and their experimental values are approximately:

$$
\begin{aligned}
g &= 0.631 \\
\theta_W &= 28.67^\circ \\
v &= 246.218\,\mathrm{GeV} \\
0.2 < \lambda &< 4.0 \\
h_e &= 2.929 \cdot 10^{-6}\,\mathrm{eV} \\
h_\mu &= 6.065 \cdot 10^{-4}\,\mathrm{eV} \\
h_\tau &= 1.021 \cdot 10^{-2}\,\mathrm{eV}
\end{aligned}
$$

All the parameters have been measured quite exactly, except λ.

Other physical constants can then be calculated using the formulas:

$$
\begin{aligned}
m_W &= \tfrac{1}{2} g v = 77.7\,\mathrm{GeV} \\
m_Z = \frac{m_W}{\cos\theta_W} &= 88.6\,\mathrm{GeV} \\
m_H = v\sqrt{2\lambda} &= \text{from } 150\,\mathrm{GeV} \text{ to } 700\,\mathrm{GeV} \\
m_e = \frac{1}{\sqrt{2}} h_e v &= 511\,\mathrm{KeV} \\
m_\mu = \frac{1}{\sqrt{2}} h_\mu v &= 105.6\,\mathrm{MeV} \\
m_\tau = \frac{1}{\sqrt{2}} h_\tau v &= 1.777\,\mathrm{GeV} \\
G_F = \frac{1}{\sqrt{2} v^2} &= (1.16639 \pm 0.00001) \times 10^{-5}\,\mathrm{GeV}^{-2} \\
e = g \sin\theta_W &= 0.3 \\
\alpha = \frac{1}{4\pi} g^2 \sin^2\theta_W &\doteq \frac{1}{137}
\end{aligned}
$$

Code:

```
>>> from math import pi, sin, cos, sqrt
>>> eV = 1
>>> KeV = 1e3
>>> MeV = 1e6
>>> GeV = 1e9
>>> g = 0.631
>>> theta_W = 28.67 * pi / 180
>>> v = 246.218 * GeV
>>> h_e = 2.935 * 1e-6 * eV
>>> h_mu = 6.065 * 1e-4 * eV
>>> h_tau = 1.021 * 1e-2 * eV
>>> g*v/2 / GeV
77.681779
>>> g*v/2/cos(theta_W) / GeV
88.5365869768
>>> h_e * v / sqrt(2) / KeV
510.99059521630568
>>> h_mu * v / sqrt(2) / MeV
105.59311618353983
>>> h_tau * v / sqrt(2) / GeV
1.7775856821664329
>>> 1./sqrt(2)/v**2 / (1e-5 * GeV**-2)
1.1663943402665491
>>> g*sin(theta_W)
0.30273118431564783
>>> 1. / (g**2*sin(theta_W)**2/(4*pi))
137.11833915409719
```

Quarks

$$\mathcal{L}_{fermion} += \sum_{q=d,s,b} i\bar{L}_0^{(q)}\gamma^\mu\partial_\mu L_0^{(q)} + \sum_{q=d,u,s,c,b,t} i\bar{q}_{0R}\gamma^\mu\partial_\mu q_{0R}$$

$$\mathcal{L}_{Yukawa} += -\sum_{\substack{q=d,s,b\\q'=d,s,b}} h_{qq'} i\bar{L}_0^{(q)}\Phi q'_{0R} + \text{h.c.} - \sum_{\substack{q=d,s,b\\q'=u,c,t}} \tilde{h}_{qq'} i\bar{L}_0^{(q)}\tilde{\Phi} q'_{0R} + \text{h.c.}$$

8.2.2 QFT

Field Operators

The free (non-interacting) fields in the interaction picture are expressed using the creation and anihilation operators below, also the corresponding non-interacting Hamiltonian is shown.

The general idea behind the machinery is that the field operator $\hat{\psi}(\mathbf{x}) = \sum_k \psi_k(\mathbf{x})c_k$ is constructed as a sum (or an integral, depending on if the index k is discrete or continuous) of single-particle wave functions (i.e. solutions of the noninteracting equation of motion) multiplied by the creation/anihilation operators (c_k or $c_k^\dagger$) that create/destroy the particle in the given single-particle state. Note that the noninteracting equation of motion usually means that we set all potentials (interactions) as zero, but in principle it can be any equation that we can solve exactly.

The coefficients $\psi_k(\mathbf{x})$ don't depend on time (so neither the field operators in the Schrödinger picture), but we work in the interaction picture, where the creation/anihilation operators depend on time, and the time dependence is put into the exponentials below (but the integration is still done over the spatial components of p only).

Scalar bosons:

$$\phi_I(x) = \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \left(a_{\mathbf{p}} e^{-ip\cdot x} + a^\dagger_{\mathbf{p}} e^{ip\cdot x}\right)$$

$$\pi_I(x) = \partial_t \phi_I(x) = \int \frac{\mathrm{d}^3 p}{(2\pi)^3} (-i) \sqrt{\frac{E_{\mathbf{p}}}{2}} \left(a_{\mathbf{p}} e^{-ip\cdot x} - a^\dagger_{\mathbf{p}} e^{ip\cdot x}\right)$$

where:

$$\left[a_{\mathbf{p}}, a^\dagger_{\mathbf{q}}\right] = (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{q})$$

(all other commutators are equal to zero). The equal-time commutation relations for ϕ and π are then:

$$[\phi(\mathbf{x}), \pi(\mathbf{y})] = i\delta^{(3)}(\mathbf{x} - \mathbf{y})$$

(all other commutators are equal to zero).

The Hamiltonian is

$$H = \int \frac{\mathrm{d}^3 p}{(2\pi)^3} E_{\mathbf{p}} a^\dagger_{\mathbf{p}} a_{\mathbf{p}}$$

Fermions:

$$\psi_I(x) = \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \sum_{s=1}^{2} \left(b^s_{\mathbf{p}} u^s(\mathbf{p}) e^{-ip\cdot x} + d^{s\dagger}_{\mathbf{p}} v^s(\mathbf{p}) e^{ip\cdot x}\right)$$

$$\bar{\psi}_I(x) = \psi^\dagger_I(x)\gamma^0 = \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \sum_{s=1}^{2} \left(d^s_{\mathbf{p}} \bar{v}^s(\mathbf{p}) e^{-ip\cdot x} + b^{s\dagger}_{\mathbf{p}} \bar{u}^s(\mathbf{p}) e^{ip\cdot x}\right)$$

where

$$u^s(\mathbf{p}) = \begin{pmatrix} \sqrt{\mathbf{p}\cdot\sigma}\xi^s \\ \sqrt{\mathbf{p}\cdot\bar{\sigma}}\xi^s \end{pmatrix}$$

$$v^s(\mathbf{p}) = \begin{pmatrix} \sqrt{\mathbf{p}\cdot\sigma}\eta^s \\ -\sqrt{\mathbf{p}\cdot\bar{\sigma}}\eta^s \end{pmatrix}$$

$$\sum_{s=1}^{2} u^s(\mathbf{p})\bar{u}^s(\mathbf{p}) = \not{p} + m$$

$$\sum_{s=1}^{2} v^s(\mathbf{p})\bar{v}^s(\mathbf{p}) = \not{p} - m$$

$$\left\{b^r_{\mathbf{p}}, b^{s\dagger}_{\mathbf{q}}\right\} = \left\{d^r_{\mathbf{p}}, d^{s\dagger}_{\mathbf{q}}\right\} = (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{q})\delta^{rs}$$

(all other anticommutators are equal to zero). The equal-time anticommutation relations for ψ and $\psi^\dagger$ are then:

$$\left\{\psi_a(\mathbf{x}), \psi^\dagger_b(\mathbf{y})\right\} = \delta^{(3)}(\mathbf{x} - \mathbf{y})\delta_{ab}$$

$$\{\psi_a(\mathbf{x}), \psi_b(\mathbf{y})\} = \left\{\psi^\dagger_a(\mathbf{x}), \psi^\dagger_b(\mathbf{y})\right\} = 0$$

The Hamiltonian is

$$H = \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \sum_{s=1}^{2} E_{\mathbf{p}} \left(b^{s\dagger}_{\mathbf{p}} b^s_{\mathbf{p}} + d^{s\dagger}_{\mathbf{p}} d^s_{\mathbf{p}}\right)$$

and the total charge:

$$Q = \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \sum_{s=1}^{2} \left(b^{s\dagger}_{\mathbf{p}} b^s_{\mathbf{p}} - d^{s\dagger}_{\mathbf{p}} d^s_{\mathbf{p}}\right)$$

So the b-type particles and d-type particles are identical except the charge. In QED, we identify the b-type particles as electrons and the d-type particles as positrons.

Vector bosons:

$$A_\mu(x) = \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \sum_{r=0}^{3} \left(a^r_{\mathbf{p}} \epsilon^r_\mu(\mathbf{p}) e^{-ip\cdot x} + a^{r\dagger}_{\mathbf{p}} \epsilon^{r*}_\mu(\mathbf{p}) e^{ip\cdot x} \right)$$

where

$$\left[a^r_{\mathbf{p}}, a^{s\dagger}_{\mathbf{q}} \right] = (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{q}) \delta^{rs}$$

The equal-time commutation relations for A_μ are then:

$$\left[A_\mu(\mathbf{x}), A^\dagger_\nu(\mathbf{y}) \right] = \delta^{(3)}(\mathbf{x} - \mathbf{y}) \delta_{\mu\nu}$$

Calculating Scattering Amplitudes using Green Functions

We are interested in calculating the following scattering amplitudes:

$$\langle f | i \rangle$$

where the initial $|i\rangle$ and final $|f\rangle$ states are created by creation operators of the fields from the previous section. For example

$$|i\rangle = b^\dagger_1 b^\dagger_2 |\Omega\rangle$$
$$|f\rangle = b^\dagger_{1'} b^\dagger_{2'} |\Omega\rangle$$

Depending on the particular creation and anihilation operators, it can be shown that they can be replaced by:

$$a^\dagger_{\mathbf{k}\text{ in}} \to i \int \mathrm{d}^4 x e^{ikx} \left(\partial^2 + m^2 \right) \phi(x) = \frac{k^2 - m^2}{i} \tilde{\phi}(-k) = \frac{1}{\tilde{D}(k)} \tilde{\phi}(-k)$$

$$a_{\mathbf{k}\text{ out}} \to i \int \mathrm{d}^4 x e^{-ikx} \left(\partial^2 + m^2 \right) \phi(x) = \frac{k^2 - m^2}{i} \tilde{\phi}(k) = \frac{1}{\tilde{D}(k)} \tilde{\phi}(k)$$

$$b^{s\dagger}_{\mathbf{k}\text{ in}} \to i \int \mathrm{d}^4 x \bar{\psi}(x) \left(i \overleftarrow{\partial\!\!\!/} + m \right) u^s(\mathbf{k}) e^{ikx} = \tilde{\bar{\psi}}(-k) \frac{-k\!\!\!/ - m}{i} u^s(\mathbf{k}) = \tilde{\bar{\psi}}(-k) \frac{1}{\tilde{S}(-k)} u^s(\mathbf{k})$$

$$b^s_{\mathbf{k}\text{ out}} \to i \int \mathrm{d}^4 x e^{-ikx} \bar{u}^s(\mathbf{k}) \left(-i\partial\!\!\!/ + m \right) \psi(x) = \bar{u}^s(\mathbf{k}) \frac{k\!\!\!/ - m}{i} \psi(k) = \bar{u}^s(\mathbf{k}) \frac{1}{\tilde{S}(k)} \tilde{\psi}(k)$$

$$d^{s\dagger}_{\mathbf{k}\text{ in}} \to -i \int \mathrm{d}^4 x e^{ikx} \bar{v}^s(\mathbf{k}) \left(-i\partial\!\!\!/ + m \right) \psi(x) = -\bar{v}^s(\mathbf{k}) \frac{k\!\!\!/ - m}{i} \psi(-k) = -\bar{v}^s(\mathbf{k}) \frac{1}{\tilde{S}(k)} \tilde{\psi}(-k)$$

$$d^s_{\mathbf{k}\text{ out}} \to -i \int \mathrm{d}^4 x \bar{\psi}(x) \left(i \overleftarrow{\partial\!\!\!/} + m \right) v^s(\mathbf{k}) e^{-ikx} = -\tilde{\bar{\psi}}(k) \frac{-k\!\!\!/ - m}{i} v^s(\mathbf{k}) = -\tilde{\bar{\psi}}(k) \frac{1}{\tilde{S}(-k)} v^s(\mathbf{k})$$

$$a^{r\dagger}_{\mathbf{k}\text{ in}} \to i \epsilon^{r*}_\mu(\mathbf{k}) \int \mathrm{d}^4 x e^{ikx} \partial^2 A^\mu(x) = \epsilon^{r*}_\mu(\mathbf{k}) \frac{k^2}{i} \tilde{A}^\mu(-k)$$

$$a^r_{\mathbf{k}\text{ out}} \to i \epsilon^r_\mu(\mathbf{k}) \int \mathrm{d}^4 x e^{-ikx} \partial^2 A^\mu(x) = \epsilon^r_\mu(\mathbf{k}) \frac{k^2}{i} \tilde{A}^\mu(k)$$

where the "in" is the operator for $t \to -\infty$ and "out" for $t \to \infty$. The fields $\phi(x)$, $\psi(x)$, $\bar{\psi}(x)$ and $A^\mu(x)$ have to be time ordered. On the left hand side is a position space representation, the two expressions on the right hand side are the momentum representation (the last expression is written using the propagators), e.g. a Fourier transform, which is

essentially just the following substitutions:

$$\begin{aligned}
\partial^2 &\to -k^2 \\
i\partial\!\!\!/ &\to k\!\!\!/ \\
e^{\pm ikx}\phi(x) &\to \tilde{\phi}(\mp k) \\
\frac{k^2 - m^2}{i} &\to \frac{1}{\tilde{D}(k)} \\
\frac{\pm k\!\!\!/ - m}{i} &\to \frac{1}{\tilde{S}(\pm k)}
\end{aligned}$$

both representations are of course equivalent (but the momentum one is easier to use, since the formulas are shorter).

For our example we get in the position space:

$$\begin{aligned}
\langle f|i\rangle = \langle\Omega|b_{\mathbf{p}_{2'}} b_{\mathbf{p}_{1'}} b^\dagger_{\mathbf{p}_1} b^\dagger_{\mathbf{p}_2}|\Omega\rangle = \langle\Omega|T\, b_{\mathbf{p}_{2'}} b_{\mathbf{p}_{1'}} b^\dagger_{\mathbf{p}_1} b^\dagger_{\mathbf{p}_2}|\Omega\rangle = \\
= i^4 \int \mathrm{d}^4x_1 \mathrm{d}^4x_2 \mathrm{d}^4x_{1'} \mathrm{d}^4x_{2'} \\
e^{-ip_{1'}x_{1'}} \left[\bar{u}^{s_{1'}}(\mathbf{k}_{1'})\left(-i\partial\!\!\!/_{1'} + m\right)\right]_{\alpha_{1'}} \\
e^{-ip_{2'}x_{2'}} \left[\bar{u}^{s_{2'}}(\mathbf{k}_{2'})\left(-i\partial\!\!\!/_{2'} + m\right)\right]_{\alpha_{2'}} \\
\langle\Omega|T\, \psi_{\alpha_{2'}}(x_{2'})\psi_{\alpha_{1'}}(x_{1'})\bar{\psi}_{\alpha_1}(x_1)\bar{\psi}_{\alpha_2}(x_2)|\Omega\rangle \\
\left[\left(i\overleftarrow{\partial\!\!\!/}_1 + m\right) u^{s_1}(\mathbf{p}_1)\right]_{\alpha_1} e^{ip_1x_1} \\
\left[\left(i\overleftarrow{\partial\!\!\!/}_2 + m\right) u^{s_2}(\mathbf{p}_2)\right]_{\alpha_2} e^{ip_2x_2}
\end{aligned}$$

where the α_1, α_2, $\alpha_{1'}$ and $\alpha_{2'}$ spinor indices were introduced to show how the matrices should be multiplied. The vacuum amplitude is called a 4 point interacting Green function in position space:

$$G^{(4)}_{\alpha_{1'}\alpha_{2'}\alpha_1\alpha_2}(x_{1'}, x_{2'}, x_1, x_2) = \langle\Omega|T\, \psi_{\alpha_{2'}}(x_{2'})\psi_{\alpha_{1'}}(x_{1'})\bar{\psi}_{\alpha_1}(x_1)\bar{\psi}_{\alpha_2}(x_2)|\Omega\rangle$$

we can also take a Fourier transform to get the Green function in momentum space:

$$\tilde{G}^{(n)}(p_1, \ldots, p_n) = \int \prod_{i=1}^{n} \mathrm{d}^4x_i e^{-ip_ix_i}\, G^{(n)}(x_1, \ldots, x_n)$$

then the scattering amplitude becomes (resuming the previous calculation):

$$\begin{aligned}
\langle f|i\rangle = \cdots = i^4 \\
\left[\bar{u}^{s_{1'}}(\mathbf{k}_{1'})\left(-p\!\!\!/_{1'} + m\right)\right]_{\alpha_{1'}} \\
\left[\bar{u}^{s_{2'}}(\mathbf{k}_{2'})\left(-p\!\!\!/_{2'} + m\right)\right]_{\alpha_{2'}} \\
\tilde{G}^{(4)}_{\alpha_{1'}\alpha_{2'}\alpha_1\alpha_2}(p_{1'}, p_{2'}, -p_1, -p_2) \\
\left[\left(p\!\!\!/_1 + m\right) u^{s_1}(\mathbf{p}_1)\right]_{\alpha_1} \\
\left[\left(p\!\!\!/_2 + m\right) u^{s_2}(\mathbf{p}_2)\right]_{\alpha_2}
\end{aligned}$$

We can get the same result much faster if we use the momentum space from the beginning:

$$\langle f|i\rangle = \langle\Omega|b_{\mathbf{p}_{2'}} b_{\mathbf{p}_{1'}} b^\dagger_{\mathbf{p}_1} b^\dagger_{\mathbf{p}_2}|\Omega\rangle = \langle\Omega|T\, b_{\mathbf{p}_{2'}} b_{\mathbf{p}_{1'}} b^\dagger_{\mathbf{p}_1} b^\dagger_{\mathbf{p}_2}|\Omega\rangle =$$
$$= \langle\Omega|T\, \bar{u}^{s_{2'}}(\mathbf{p}_{2'})\frac{1}{\tilde{S}(\mathbf{p}_{2'})}\tilde{\psi}(\mathbf{p}_{2'})\bar{u}^{s_{1'}}(\mathbf{p}_{1'})\frac{1}{\tilde{S}(\mathbf{p}_{1'})}\tilde{\psi}(\mathbf{p}_{1'})\tilde{\bar{\psi}}(-\mathbf{p}_1)\frac{1}{\tilde{S}(-\mathbf{p}_1)}u^{s_1}(\mathbf{p}_1)\tilde{\bar{\psi}}(-\mathbf{p}_2)\frac{1}{\tilde{S}(-\mathbf{p}_2)}u^{s_2}(\mathbf{p}_2)|\Omega\rangle =$$
$$= \left[\bar{u}^{s_{2'}}(\mathbf{p}_{2'})\frac{1}{\tilde{S}(\mathbf{p}_{2'})}\right]_{\alpha_{2'}}\left[\bar{u}^{s_{1'}}(\mathbf{p}_{1'})\frac{1}{\tilde{S}(\mathbf{p}_{1'})}\right]_{\alpha_{1'}}$$
$$\langle\Omega|T\,\tilde{\psi}_{\alpha_{2'}}(\mathbf{p}_{2'})\tilde{\psi}_{\alpha_{1'}}(\mathbf{p}_{1'})\tilde{\bar{\psi}}_{\alpha_1}(-\mathbf{p}_1)\tilde{\bar{\psi}}_{\alpha_2}(-\mathbf{p}_2)|\Omega\rangle$$
$$\left[\frac{1}{\tilde{S}(-\mathbf{p}_1)}u^{s_1}(\mathbf{p}_1)\right]_{\alpha_1}\left[\frac{1}{\tilde{S}(-\mathbf{p}_2)}u^{s_2}(\mathbf{p}_2)\right]_{\alpha_1}$$

This is called Lehmann-Symanzik-Zimmermann (LSZ) reduction formula. One obtains similar expressions for other fields as well (if there were different creation operators between the initial and final states). All that remains is to calculate the interacting Green functions (for which we need to know the interaction Lagrangian). But first couple more examples:

Example 1

ν_e - e elastic scattering:

$$\nu_e(k) + e(p) \to \nu_e(k') + e(p')$$

So the initial and final states are:

$$|i\rangle = b^\dagger_{\mathbf{k}} b^\dagger_{\mathbf{p}}\,|\Omega\rangle$$
$$|f\rangle = b^\dagger_{\mathbf{k}'} b^\dagger_{\mathbf{p}'}\,|\Omega\rangle$$

and we get:

$$\langle f|i\rangle = \langle\Omega|b_{\mathbf{p}'} b_{\mathbf{k}'} b^\dagger_{\mathbf{k}} b^\dagger_{\mathbf{p}}|\Omega\rangle = \langle\Omega|T\, b_{\mathbf{p}'} b_{\mathbf{k}'} b^\dagger_{\mathbf{k}} b^\dagger_{\mathbf{p}}|\Omega\rangle =$$
$$= \left[\bar{u}(\mathbf{p}')\frac{1}{\tilde{S}(\mathbf{p}')}\right]\left[\bar{u}(\mathbf{k}')\frac{1}{\tilde{S}(\mathbf{k}')}\right]$$
$$\langle\Omega|T\,\tilde{\psi}(\mathbf{p}')\tilde{\psi}(\mathbf{k}')\tilde{\bar{\psi}}(-\mathbf{k})\tilde{\bar{\psi}}(-\mathbf{p})|\Omega\rangle$$
$$\left[\frac{1}{\tilde{S}(-\mathbf{k})}u(\mathbf{k})\right]\left[\frac{1}{\tilde{S}(-\mathbf{p})}u(\mathbf{p})\right]$$

We only multiply the matrices with the same momentum, i.e. $\left[\bar{u}^s(\mathbf{p}')\frac{1}{\tilde{S}(\mathbf{p}')}\right]$ with $\tilde{\psi}(\mathbf{p}')$, $\left[\bar{u}^s(\mathbf{k}')\frac{1}{\tilde{S}(\mathbf{k}')}\right]$ with $\tilde{\psi}(\mathbf{k}')$ and so on. Also we don't write the spin anymore, e.g. $u(\mathbf{k})$ should in fact be $u^{s_k}(\mathbf{k})$ and so on.

Example 2

Muon decay:

$$\mu(P) \to e(p) + \bar{\nu}_e(k') + \nu_\mu(k)$$

So the initial and final states are:

$$|i\rangle = b^\dagger_{\mathbf{P}}\,|\Omega\rangle$$
$$|f\rangle = b^\dagger_{\mathbf{p}} d^\dagger_{\mathbf{k}'} b^\dagger_{\mathbf{k}}\,|\Omega\rangle$$

and we get:

$$\begin{aligned}\langle f|i\rangle &= \langle\Omega|b_{\mathbf{k}}d_{\mathbf{k}'}b_{\mathbf{p}}b^{\dagger}_{\mathbf{P}}|\Omega\rangle \\ &= \left[\bar{u}(\mathbf{k})\frac{1}{\tilde{S}(\mathbf{k})}\right]\left[\bar{u}(\mathbf{p})\frac{1}{\tilde{S}(\mathbf{p})}\right] \\ &\langle\Omega|T\,\tilde{\psi}(\mathbf{k})\tilde{\psi}(\mathbf{p})\tilde{\bar{\psi}}(\mathbf{k}')\tilde{\bar{\psi}}(-\mathbf{P})|\Omega\rangle \\ &\left[-\frac{1}{\tilde{S}(-\mathbf{k}')}v(\mathbf{k}')\right]\left[\frac{1}{\tilde{S}(-\mathbf{P})}u(\mathbf{P})\right]\end{aligned}$$

Example 3

$e^+ + e^-$ scattering:

$$e^-(p_1) + e^+(p_2) \to e^-(k_1) + e^+(k_2)$$

Initial and final states:

$$\begin{aligned}|i\rangle &= b^{t\dagger}_{\mathbf{p}_1}d^{u\dagger}_{\mathbf{p}_2}|\Omega\rangle \\ |f\rangle &= b^{r\dagger}_{\mathbf{k}_1}d^{s\dagger}_{\mathbf{k}_2}|\Omega\rangle\end{aligned}$$

And we get:

$$\begin{aligned}\langle f|i\rangle &= \langle\Omega|b^r_{\mathbf{k}_1}d^s_{\mathbf{k}_2}b^{t\dagger}_{\mathbf{p}_1}d^{u\dagger}_{\mathbf{p}_2}|\Omega\rangle = \\ &= \langle\Omega|Tb^r_{\mathbf{k}_1}d^s_{\mathbf{k}_2}b^{t\dagger}_{\mathbf{p}_1}d^{u\dagger}_{\mathbf{p}_2}|\Omega\rangle = \\ &= \langle\Omega|T\left[\bar{u}^r(\mathbf{k}_1)\frac{1}{\tilde{S}(k_1)}\tilde{\psi}(k_1)\right]\left[-\tilde{\bar{\psi}}(k_2)\frac{1}{\tilde{S}(-k_2)}v^s(\mathbf{k}_2)\right]\left[\tilde{\bar{\psi}}(-p_1)\frac{1}{\tilde{S}(-p_1)}u^t(\mathbf{p}_1)\right]\left[-\bar{v}^u(\mathbf{p}_2)\frac{1}{\tilde{S}(p_2)}\tilde{\psi}(-p_2)\right]|\Omega\rangle = \\ &= \left[\bar{u}^r(\mathbf{k}_1)\frac{1}{\tilde{S}(k_1)}\right]\left[\bar{v}^u(\mathbf{p}_2)\frac{1}{\tilde{S}(p_2)}\right] \\ &\langle\Omega|T\tilde{\psi}(k_1)\tilde{\bar{\psi}}(k_2)\tilde{\bar{\psi}}(-p_1)\tilde{\psi}(-p_2)|\Omega\rangle \\ &\left[\frac{1}{\tilde{S}(-k_2)}v^s(\mathbf{k}_2)\right]\left[\frac{1}{\tilde{S}(-p_1)}u^t(\mathbf{p}_1)\right]\end{aligned}$$

Example 4

$H(p) \to Z(k) + Z(l)$ decay. Initial and final states:

$$\begin{aligned}|i\rangle &= a^{\dagger}_{\mathbf{p}}|\Omega\rangle \\ |f\rangle &= a^{r\dagger}_{\mathbf{k}}a^{s\dagger}_{\mathbf{l}}|\Omega\rangle\end{aligned}$$

and we get:

$$\begin{aligned}\langle f|i\rangle &= \langle\Omega|a^r_{\mathbf{k}}a^s_{\mathbf{l}}a^{\dagger}_{\mathbf{p}}|\Omega\rangle = \langle\Omega|Ta^r_{\mathbf{k}}a^s_{\mathbf{l}}a^{\dagger}_{\mathbf{p}}|\Omega\rangle = \\ &= \langle\Omega|T\epsilon^{r*}_{\mu}(\mathbf{k})\frac{k^2}{i}\tilde{A}^{\mu}(k)\epsilon^{s*}_{\nu}(\mathbf{l})\frac{l^2}{i}\tilde{A}^{\nu}(l)\frac{1}{\tilde{D}(p)}\tilde{\phi}(-p)|\Omega\rangle = \\ &= \frac{\epsilon^{r*}_{\mu}(\mathbf{k})\epsilon^{s*}_{\nu}(\mathbf{l})}{\frac{i}{k^2}\frac{i}{l^2}\tilde{D}(p)}\langle\Omega|T\tilde{A}^{\mu}(k)\tilde{A}^{\nu}(l)\tilde{\phi}(-p)|\Omega\rangle\end{aligned}$$

Example 5

$e^+e^- \to W^+W^-$ scattering:

$$e^-(k) + e^+(-l) \to W^-(p) + W^+(r)$$

So the initial and final states are:

$$|i\rangle = b^\dagger_{\mathbf{k}} d^\dagger_{-\mathbf{l}} |\Omega\rangle$$
$$|f\rangle = a^{\lambda\dagger}_{\mathbf{p}} a^{\mu\dagger}_{\mathbf{r}} |\Omega\rangle$$

and we get:

$$\langle f|i\rangle = \langle\Omega|a^\mu_{\mathbf{r}} a^\lambda_{\mathbf{p}} b^\dagger_{\mathbf{k}} d^\dagger_{-\mathbf{l}}|\Omega\rangle$$
$$= \epsilon^\mu_\alpha(\mathbf{r})\frac{r^2}{i}\epsilon^\lambda_\beta(\mathbf{p})\frac{p^2}{i}\left[-\bar{v}(-\mathbf{l})\frac{1}{\tilde{S}(-\mathbf{l})}\right]$$
$$\langle\Omega|T\,\tilde{A}^\alpha(\mathbf{r})\tilde{A}^\beta(\mathbf{p})\tilde{\bar{\psi}}(-\mathbf{k})\tilde{\psi}(\mathbf{l})|\Omega\rangle$$
$$\left[\frac{1}{\tilde{S}(-\mathbf{k})}u(\mathbf{k})\right]$$

Example 6

$W^+W^- \to W^+W^-$ scattering:

$$W^-(k) + W^+(p) \to W^-(r) + W^+(l)$$

So the initial and final states are:

$$|i\rangle = a^{\mu\dagger}_{\mathbf{k}} a^{\nu\dagger}_{\mathbf{p}} |\Omega\rangle$$
$$|f\rangle = a^{\alpha\dagger}_{\mathbf{r}} a^{\beta\dagger}_{\mathbf{l}} |\Omega\rangle$$

and we get:

$$\langle f|i\rangle = \langle\Omega|a^\beta_{\mathbf{l}} a^\alpha_{\mathbf{r}} a^{\mu\dagger}_{\mathbf{k}} a^{\nu\dagger}_{\mathbf{p}}|\Omega\rangle$$
$$= \epsilon^{\mu*}_\rho(\mathbf{k})\frac{k^2}{i}\epsilon^{\nu*}_\sigma(\mathbf{p})\frac{p^2}{i}\,\langle\Omega|T\,\tilde{A}^\kappa(\mathbf{r})\tilde{A}^\lambda(\mathbf{l})\tilde{A}^\sigma(-\mathbf{p})\tilde{A}^\rho(-\mathbf{k})|\Omega\rangle\,\epsilon^\alpha_\kappa(\mathbf{r})\frac{r^2}{i}\epsilon^\beta_\lambda(\mathbf{l})\frac{l^2}{i}$$

Evaluation of the Interacting Green Functions

The interacting Green functions can be evaluated using the formula:

$$G^{(n)}(x_1,\ldots,x_n) = \langle\Omega|T\,\phi_H(x_1)\ldots\phi_H(x_n)|\Omega\rangle =$$
$$= \frac{\langle 0|T\,\phi_I(x_1)\ldots\phi_I(x_n)S|0\rangle}{\langle 0|S|0\rangle}$$

where

$$S = U_I(\infty,-\infty) = T\exp\left(-\frac{i}{\hbar}\int_{-\infty}^{\infty} H_1(t)\mathrm{d}t\right) = T\exp\left(-\frac{i}{\hbar}\int \mathrm{d}^4x\mathcal{H}_1(x)\right)$$

ϕ_H is a field in the Heisenberg picture ($\phi(\mathbf{x},t) = e^{iHt}\phi(\mathbf{x},0)e^{-iHt}$) and ϕ_I is a field in the interaction picture ($\phi(\mathbf{x},t) = e^{iH_0t}\phi(\mathbf{x},0)e^{-iH_0t}$), where the Hamiltonian is $H = H_0 + H_1$ and the vacua (ground states) are $H_0|0\rangle = 0$ and $H|\Omega\rangle = 0$.

This can be proven by evaluating the right hand side:

$$\frac{\langle 0|T\,\phi_I(x_1)\dots\phi_I(x_n)S|0\rangle}{\langle 0|S|0\rangle} = \frac{\langle 0|T\,\phi_I(x_1)\dots\phi_I(x_n)U_I(\infty,-\infty)|0\rangle}{\langle 0|U_I(\infty,0)U_I(0,-\infty)|0\rangle}$$
$$= \frac{\langle 0|U_I(\infty,t_1)\phi_I(x_1)U_I(t_1,t_2)\dots U_I(t_{n-1},t_n)\phi_I(x_n)U_I(t_n,-\infty)|0\rangle}{\langle 0|U_I(\infty,0)U_I(0,-\infty)|0\rangle}$$
$$= \frac{\langle 0|U_I(\infty,0)\phi_H(x_1)\dots\phi_H(x_n)U_I(0,-\infty)|0\rangle}{\langle 0|U_I(\infty,0)U_I(0,-\infty)|0\rangle} =$$
$$= \frac{\langle 0|\Omega\rangle\langle\Omega|T\phi_H(x_1)\dots\phi_H(x_n)|\Omega\rangle\langle\Omega|0\rangle}{\langle 0|\Omega\rangle\langle\Omega|\Omega\rangle\langle\Omega|0\rangle} =$$
$$= \frac{\langle\Omega|T\phi_H(x_1)\dots\phi_H(x_n)|\Omega\rangle}{\langle\Omega|\Omega\rangle} =$$
$$= \langle\Omega|T\phi_H(x_1)\dots\phi_H(x_n)|\Omega\rangle$$

where we used the following relations:

$$U_I(t_{k-1},t_k)\phi_I(x_k)U_I(t_k,t_{k+1}) = U_I(t_{k-1},0)U_I^\dagger(t_k,0)\phi_I(x_k)U_I(t_k,0)U_I(0,t_{k+1}) = U_I(t_{k-1},0)\phi_H(x_k)U_I(0,t_{k+1})$$

$$U_I(0,-\infty)\,|0\rangle = U_I(0,-\infty)\left[|\Omega\rangle\langle\Omega| + \sum_{n\neq 0}|n\rangle\langle n|\right]|0\rangle = |\Omega\rangle\langle\Omega|0\rangle + \lim_{t\to-\infty}\sum_{n\neq 0}e^{iE_nt}\,|n\rangle\langle n|0\rangle = |\Omega\rangle\langle\Omega|0\rangle$$

$$\langle\Omega|\Omega\rangle = 1$$

Evolution Operator, S-Matrix Elements

The evolution operator U is defined by the equations:

$$|\phi(t_2)\rangle = U(t_2,t_1)\,|\phi(t_1)\rangle$$

$$i\hbar\frac{\partial U(t,t_1)}{\partial t} = H(t)U(t,t_1)$$

$$U(t_1,t_1) = 1$$

We are interested in calculating the S matrix elements:

$$\langle f|U(\infty,-\infty)|i\rangle = \langle f|S|i\rangle = S_{fi}$$

so we first calculate $U(\infty,-\infty)$. Integrating the equation for the evolution operator:

$$U(t_2,t_1) = U(t_1,t_1) - \frac{i}{\hbar}\int_{t_1}^{t_2}H(t)U(t,t_1)\mathrm{d}t = 1 - \frac{i}{\hbar}\int_{t_1}^{t_2}H(t)U(t,t_1)\mathrm{d}t$$

Now:

$$S = U(\infty,-\infty) = 1 - \frac{i}{\hbar}\int_{-\infty}^{\infty}H(t')U(t',-\infty)\mathrm{d}t' =$$

$$= 1 + \left(-\frac{i}{\hbar}\right)\int_{-\infty}^{\infty}H(t')U(t',-\infty)\mathrm{d}t' + \left(-\frac{i}{\hbar}\right)^2\int_{-\infty}^{\infty}\int_{-\infty}^{t'}H(t')H(t'')U(t'',-\infty)\mathrm{d}t'\mathrm{d}t'' =$$

$$= \dots = \sum_{n=0}^{\infty}\left(-\frac{i}{\hbar}\right)^n\frac{1}{n!}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots T\{H(t_1)H(t_2)\cdots\}\mathrm{d}t_1\mathrm{d}t_2\cdots =$$

$$= T\exp\left(-\frac{i}{\hbar}\int_{-\infty}^{\infty} H(t)\mathrm{d}t\right) = T\exp\left(-\frac{i}{\hbar}\int_{-\infty}^{\infty}\mathrm{d}^4x\mathcal{H}(x)\right)$$

If $\mathcal{L}$ doesn't contain derivatives of the fields, then $\mathcal{H} = -\mathcal{L}$ so:

$$U(\infty,-\infty) = T\exp\left(\frac{i}{\hbar}\int_{-\infty}^{\infty}\mathrm{d}^4x\mathcal{L}(x)\right)$$

Let's write $S = 1 + iT$ and $|i\rangle = |k_1\cdots k_m\rangle$, $|f\rangle = |p_1\cdots p_n\rangle$. As a first step now, let's investigate a scalar field, e.g. $\mathcal{L} = -\frac{\lambda}{4}\phi^4$ (e.g. a Higgs self interaction term above), we'll look at other fields later:

$$\langle f|S|i\rangle = \langle f|iT|i\rangle = \langle p_1\cdots p_n|iT|k_1\cdots k_m\rangle = \frac{1}{\tilde{D}(k_1)\cdots\tilde{D}(k_m)}\frac{1}{\tilde{D}(p_1)\cdots\tilde{D}(p_n)}$$

$$\int\mathrm{d}^4x_1\cdots\mathrm{d}^4x_m e^{-i(k_1x_1+\cdots+k_mx_m)}\int\mathrm{d}^4y_1\cdots\mathrm{d}^4y_n e^{+i(p_1y_1+\cdots+p_ny_n)}G(x_1,\cdots,x_m,y_1,\cdots,y_m)$$

where

$$G(x_1,\cdots,x_n) = \langle\Omega|T\{\phi(x_1)\cdots\phi(x_n)\}|\Omega\rangle =$$

$$\frac{\langle 0|T\{\phi_I(x_1)\cdots\phi_I(x_n)\exp\left(\frac{i}{\hbar}\int_{-\infty}^{\infty}\mathrm{d}^4x\mathcal{L}(x)\right)\}|0\rangle}{\langle 0|T\exp\left(\frac{i}{\hbar}\int_{-\infty}^{\infty}\mathrm{d}^4x\mathcal{L}(x)\right)|0\rangle}$$

This is called the LSZ formula. Now we use the Wick contraction, get some terms like $D_{23}D_{34}$ integrate things out, this will give the delta function and $\tilde{D}(p)$'s and that's it.

Let's see how it goes for $\mathcal{L} = -\frac{\lambda}{4}\phi^4$ for the process $k_1 + k_2 \to p_1 + p_2$:

$$\langle p_1p_2|S|k_1k_2\rangle = \frac{\int\mathrm{d}^4x_1\mathrm{d}^4x_2e^{-i(k_1x_1+k_2x_2)}\int\mathrm{d}^4y_1\mathrm{d}^4y_2e^{i(p_1y_1+p_2y_2)}}{\tilde{D}(k_1)\tilde{D}(k_2)\tilde{D}(p_1)\tilde{D}(p_2)}$$

$$\frac{\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(y_1)\phi_I(y_2)\exp\left(-\frac{i\lambda}{4\hbar}\int\mathrm{d}^4x\phi_I^4(x)\right)\}|0\rangle}{\langle 0|T\exp\left(-\frac{i\lambda}{4\hbar}\int\mathrm{d}^4x\phi_I^4(x)\right)|0\rangle} =$$

$$= \frac{\int\mathrm{d}^4x_1\mathrm{d}^4x_2e^{-i(k_1x_1+k_2x_2)}\int\mathrm{d}^4y_1\mathrm{d}^4y_2e^{i(p_1y_1+p_2y_2)}}{\tilde{D}(k_1)\tilde{D}(k_2)\tilde{D}(p_1)\tilde{D}(p_2)}$$

$$\left[\frac{\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(y_1)\phi_I(y_2)\}|0\rangle}{\langle 0|T\exp\left(-\frac{i\lambda}{4\hbar}\int\mathrm{d}^4x\phi_I^4(x)\right)|0\rangle} +\right.$$

$$+\frac{\left(-\frac{i\lambda}{4\hbar}\right)\int\mathrm{d}^4x\,\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(y_1)\phi_I(y_2)\phi_I^4(x)\}|0\rangle}{\langle 0|T\exp\left(-\frac{i\lambda}{4\hbar}\int\mathrm{d}^4x\phi_I^4(x)\right)|0\rangle}+$$

$$\left.+\frac{\left(-\frac{i\lambda}{4\hbar}\right)^2\int\mathrm{d}^4x\,\mathrm{d}^4y\,\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(y_1)\phi_I(y_2)\phi_I^4(x)\phi_I^4(y)\}|0\rangle}{\langle 0|T\exp\left(-\frac{i\lambda}{4\hbar}\int\mathrm{d}^4x\phi_I^4(x)\right)|0\rangle}+\cdots\right] =$$

$$= \frac{1}{\tilde{D}(k_1)\tilde{D}(k_2)\tilde{D}(p_1)\tilde{D}(p_2)}$$

$$\Big[(2\pi)^4\delta^{(4)}(p_1+p_2)(2\pi)^4\delta^{(4)}(k_1+k_2)\tilde{D}(p_1)\tilde{D}(k_1)+$$

$$(-i\lambda)6(2\pi)^4\delta^{(4)}(p_1+p_2-k_1-k_2)\tilde{D}(k_1)\tilde{D}(k_2)\tilde{D}(p_1)\tilde{D}(p_2)+$$

$$(-i\lambda)(\text{disconnected terms with not enough }\tilde{D}(\cdots)\text{s}) + (-i\lambda)^2(\cdots)+\cdots\Big] =$$

$$= (2\pi)^4\delta^{(4)}(p_1+p_2-k_1-k_2)\left[6(-i\lambda)+3(-i\lambda)^2\int\frac{\mathrm{d}^4k}{(2\pi)^4}\tilde{D}(k)\tilde{D}(p1+p2-k)+(-i\lambda)^3(\cdots)+\cdots\right]$$

The denominator cancels with the disconnected terms. We used the Wick contractions (see below for a thorough explanation+derivation):

$$\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(y_1)\phi_I(y_2)\}|0\rangle = D(x_1-x_2)D(y_1-y_2)+D(x_2-y_1)D(x_1-y_2)+D(x_2-y_2)D(x_1-y_1)$$

$$\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(y_1)\phi_I(y_2)\phi_I^4(x)\}|0\rangle = D(x_1-x)D(x_2-x)D(y_1-x)D(y_2-x)+\text{disconnected}$$

$$\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(y_1)\phi_I(y_2)\phi_I^4(x)\phi_I^4(y)\}|0\rangle = D(x_1-x)D(x_2-x)D(y_1-y)D(y_2-y)D(x-y)D(x-y)$$
$$+\text{disconnected}$$

Where the "disconnected" terms are $D(x_1-y_1)D(x_2-y_2)D(x-x)D(x-x)$ and similar. When they are integrated over, they do not generate enough $\tilde{D}(p_1)$ propagators to cancel the propagators from the LSZ formula, which will cause the terms to vanish.

For the $\mathcal{L}=\phi^3(x)$ theory, one also needs the following contractions:

$$\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(y_1)\phi_I(y_2)\phi_I^3(x)\}|0\rangle = 0$$

$$\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(y_1)\phi_I(y_2)\phi_I^3(x)\phi_I^3(y)\}|0\rangle = D(x_1-x)D(x_2-x)D(x-y)D(y_1-y)D(y_2-y)$$

Thus it is clear that the only difference from the above is the factor $D(x-y)$ which after integrating changes to $\tilde{D}(p_1+p_2)$ and this ends up in the final result.

One always gets the delta function in the result, so we define the matrix element $\mathcal{M}_{fi}$ by:

$$S_{fi} = (2\pi)^4\delta^{(4)}(p_1+p_2+\cdots-k_1-k_2-\cdots)i\mathcal{M}_{fi}$$

Propagators for Scalar Bosons, Fermions and Vector Bosons

The only nonzero contractions that can occur are the propagators below. All other contractions are zero.

Propagator for a scalar boson is:

$$\langle 0|T\{\phi_I(x)\phi_I(y)\}|0\rangle \equiv D(x-y) = \int \frac{\mathrm{d}^4p}{(2\pi)^4}\tilde{D}(p)e^{-ip(x-y)}$$

with

$$\tilde{D}(p) = \frac{i}{p^2-m^2+i\epsilon}$$

For fermions (Feynman propagator):

$$\langle 0|T\{\psi_I(x)\bar{\psi}_I(y)\}|0\rangle \equiv S(x-y) = \int \frac{\mathrm{d}^4p}{(2\pi)^4}\tilde{S}(p)e^{-ip(x-y)}$$

with

$$\tilde{S}(p) = \frac{i}{\not{p}-m+i\epsilon} = \frac{i(\not{p}+m)}{p^2-m^2+i\epsilon}$$

For vector bosons:

$$\langle 0|T\{A_\mu(x)A_\nu(y)\}|0\rangle \equiv D_{\mu\nu}(x-y) = \int \frac{\mathrm{d}^4p}{(2\pi)^4}\tilde{D}_{\mu\nu}(p)e^{-ip(x-y)}$$

with

$$\tilde{D}_{\mu\nu}(p) = i\frac{-g_{\mu\nu}+\frac{p_\mu p_\nu}{m^2}}{p^2-m^2+i\epsilon}$$

For massless bosons:

$$\tilde{D}_{\mu\nu}(p) = i\frac{-g_{\mu\nu}}{p^2+i\epsilon}$$

Wick Theorem

As seen above, we need to be able to calculate

$$\langle 0|T\{\phi_I(x_1)\cdots\phi_I(x_n)\}|0\rangle$$

The Wick theorem says, that this is equal to all possible contractions of fields (all fields need to be contracted), where a contraction is defined as:

$$\langle 0|T\{\phi_I(x)\phi_I(y)\}|0\rangle \equiv D(x-y) = \int \frac{\mathrm{d}^4p}{(2\pi)^4}\tilde{D}(p)e^{-ip(x-y)}$$

with

$$\tilde{D}(p) = \frac{i}{p^2 - m^2 + i\epsilon}$$

A few lowest possibilities:

$$\langle 0|T\{\phi_I(x_1)\}|0\rangle = 0$$

$$\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\}|0\rangle = D_{12}$$

$$\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\}|0\rangle = 0$$

$$\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)\}|0\rangle = \text{disconnected}$$

$$\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)\phi_I(x)\}|0\rangle = 0$$

$$\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)\phi_I^2(x)\}|0\rangle = \text{disconnected}$$

$$\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)\phi_I^3(x)\}|0\rangle = 0$$

$$\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)\phi_I^4(x)\}|0\rangle = 4!\, D(x_1-x)D(x_2-x)D(x_3-x)D(x_4-x) + \text{disconnected}$$

$$\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)\phi_I^3(x)\phi_I^3(y)\}|0\rangle =$$

$$= D(x_1-x)D(x_2-x)D(x-y)D(x_3-y)D(x_4-y) + \text{disconnected}$$

$$\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)\phi_I^4(x)\phi_I^4(y)\}|0\rangle =$$

$$= D(x_1-x)D(x_2-x)D(x-y)D(x-y)D(x_3-y)D(x_4-y) + \text{disconnected}$$

For the last two equations, not all possibilities of the connected graphs are listed (and also the combinatorial factor is omitted).

Nonrelativistic Field Operators

One difference in nonrelativistic quantum mechanics is that the noninteracting solutions to the equation of motion (Schrödinger equation in this case) can be numbered using a discrete index, so for example the momentum $\mathbf{q}$ is not continuous, thus the (anti)commutation relations for creation and anihilation operators contain the Kronecker delta (instead of a delta function) and integrals over the index are replaced by sums. The reason for that is that we usually employ boundary conditions (like a lattice, or one particle potential due to nuclei, etc.) that make the spectrum discrete.

For bosons the field operators are given by:

$$\hat{\psi}(\mathbf{x}) = \sum_k \psi_k(\mathbf{x})c_k$$

$$\hat{\psi}^\dagger(\mathbf{x}) = \sum_k \psi_k^*(\mathbf{x})c_k^\dagger$$

where the coefficients are the single-particle wave functions.

$$\begin{aligned} &= \delta_{kl} \\ [c_k, c_l] = [c_k^\dagger, c_l^\dagger] &= 0 \end{aligned}$$

so the commutation relations for $\hat{\psi}$ and $\hat{\psi}^\dagger$ are:

$$\begin{aligned} &= \delta^{(3)}(\mathbf{x} - \mathbf{y}) \\ [\hat{\psi}(\mathbf{x}), \hat{\psi}(\mathbf{y})] = [\hat{\psi}^\dagger(\mathbf{x}), \hat{\psi}^\dagger(\mathbf{y})] &= 0 \end{aligned}$$

For fermions:

$$\begin{aligned} \hat{\psi}(\mathbf{x}) &= \sum_k \sum_{s=1}^{2} \psi_k^s(\mathbf{x}) c_k \\ \hat{\psi}^\dagger(\mathbf{x}) &= \sum_k \sum_{s=1}^{2} \psi_k^{s*}(\mathbf{x}) c_k^\dagger \end{aligned}$$

where

$$\begin{aligned} \{c_k, c_l^\dagger\} &= \delta_{kl} \\ \{c_k, c_l\} = \{c_k^\dagger, c_l^\dagger\} &= 0 \end{aligned}$$

so the commutation relations for $\hat{\psi}$ and $\hat{\psi}^\dagger$ are:

$$\begin{aligned} \{\hat{\psi}(\mathbf{x}), \hat{\psi}^\dagger(\mathbf{y})\} &= \delta^{(3)}(\mathbf{x} - \mathbf{y}) \\ \{\hat{\psi}(\mathbf{x}), \hat{\psi}(\mathbf{y})\} = \{\hat{\psi}^\dagger(\mathbf{x}), \hat{\psi}^\dagger(\mathbf{y})\} &= 0 \end{aligned}$$

The (interacting) Hamiltonian for both bosons and fermions is

$$\begin{aligned} i\hbar\partial_t |\Psi(t)\rangle &= \hat{H} |\Psi(t)\rangle \\ \hat{H} = \hat{T} + \hat{V} = \sum_{ij} c_i^\dagger \langle i|T|j\rangle c_j + \tfrac{1}{2} \sum_{ijkl} c_i^\dagger c_j^\dagger \langle ij|V|kl\rangle c_l c_k \end{aligned}$$

Note the ordering of the final two destruction operators $c_l c_k$, which is opposite that of the last two single-particle wave functions in the matrix elements of the potential $\langle ij|V|kl\rangle$ (for bosons it doesn't matter, for fermions it changes a sign).

Nonrelativistic Propagator

Nonrelativistic limits of the propagators are obtained by assuming $|\mathbf{p}|/mc \ll 1$, and then expressing the propagator using the nonrelativistic energy ω (total energy minus the rest mass energy) by using the well-known relations:

$$\begin{aligned} p_0 c = E &= mc^2 + \omega \\ \sqrt{\mathbf{p}^2 c^2 + m^2 c^4} = E_{\mathbf{p}} &= mc^2 + \frac{\mathbf{p}^2}{2m} + O\left(\frac{1}{mc^2}\left(\frac{p^2}{m}\right)^2\right) \end{aligned}$$

we use them to simplify $E^2 - E_{\mathbf{p}}^2$ in the limit $c \to \infty$:

$$\frac{1}{c^2}(E^2 - E_{\mathbf{p}}^2) = \frac{1}{c^2}(E - E_{\mathbf{p}})(E + E_{\mathbf{p}}) =$$
$$= \frac{1}{c^2}\left(\omega - \frac{\mathbf{p}^2}{2m} + O\left(\frac{1}{mc^2}\left(\frac{p^2}{m}\right)^2\right)\right)\left(2mc^2 + \omega + \frac{\mathbf{p}^2}{2m} + O\left(\frac{1}{mc^2}\left(\frac{p^2}{m}\right)^2\right)\right) =$$
$$= 2m\left(\omega - \frac{\mathbf{p}^2}{2m} + \frac{\omega}{2mc^2}\left(\omega - \frac{\mathbf{p}^2}{2m}\right) + O\left(\frac{1}{mc^2}\left(\frac{p^2}{m}\right)^2\right)\right) \to$$
$$\to 2m\left(\omega - \frac{\mathbf{p}^2}{2m}\right)$$

Where E is the total energy, ω is the nonrelativistic energy, $E_{\mathbf{p}}$ is the relativistic energy of a noninteracting particle (kinetic energy). Now we can rewrite the propagator of a scalar boson:

$$\tilde{D}(p) = \frac{i}{p^2 - m^2c^2 + i\epsilon} = \frac{i}{p_0^2 - \mathbf{p}^2 - m^2c^2 + i\epsilon} = \frac{i}{\frac{1}{c^2}\left(p_0^2c^2 - \mathbf{p}^2c^2 - m^2c^4\right) + i\epsilon} =$$
$$= \frac{i}{\frac{1}{c^2}\left(E^2 - E_{\mathbf{p}}^2\right) + i\epsilon} \to \frac{i}{2m\left(\omega - \frac{\mathbf{p}^2}{2m}\right) + i\epsilon} =$$
$$= \frac{1}{2m}\frac{i}{\omega - \frac{\mathbf{p}^2}{2m} + i\frac{\epsilon}{2m}} = \frac{1}{2m}\frac{i}{\omega - \frac{\mathbf{p}^2}{2m} + i\epsilon'}$$

As you can see, we are interested in the behavior of the propagator in the vicinity of its positive frequency pole $\omega \approx \frac{\mathbf{p}^2}{2m}$.

Similarly for fermions (we set $c = 1$):

$$\tilde{S}(p) = \frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon} = \frac{i(p^0\gamma_0 - p^j\gamma_j + m)}{p^2 - m^2 + i\epsilon} \approx \frac{1}{2m}\frac{i(p^0\gamma_0 - p^j\gamma_j + m)}{\omega - \frac{\mathbf{p}^2}{2m} + i\epsilon'} =$$
$$= \frac{1}{2m}\frac{i((\omega + m)\gamma_0 - p^j\gamma_j + m)}{\omega - \frac{\mathbf{p}^2}{2m} + i\epsilon'} \approx \frac{1}{2m}\frac{i(m\gamma_0 - p^j\gamma_j + m)}{\omega - \frac{\mathbf{p}^2}{2m} + i\epsilon'} =$$
$$= \frac{i\left(\frac{1}{2}(\gamma_0 + 1) - \frac{p^j\gamma_j}{2m}\right)}{\omega - \frac{\mathbf{p}^2}{2m} + i\epsilon'} \tag{8.2.2.1}$$

The first term

$$\tfrac{1}{2}(\gamma_0 + 1) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

selects the two upper components of a given bispinor. The second term

$$-\frac{p^j\gamma_j}{2m} = \begin{pmatrix} 0 & -\frac{p^j\sigma_j}{2m} \\ \frac{p^j\sigma_j}{2m} & 0 \end{pmatrix}$$

mixes the upper and lower components of the bispinor and the contribution of this term is quadratic in $\frac{\mathbf{p}}{m}$ so it can be neglected. The numerator of (8.2.2.1) reduces to a unit matrix (in spin space):

$$\tilde{S}(p) \approx \frac{i\left(\frac{1}{2}(\gamma_0 + 1) - \frac{p^j\gamma_j}{2m}\right)}{\omega - \frac{\mathbf{p}^2}{2m} + i\epsilon} \approx \frac{i\mathbb{1}}{\omega - \frac{\mathbf{p}^2}{2m} + i\epsilon} = \mathbb{1}G_0^+(\mathbf{p}, \omega)$$

where $G_0^+(\mathbf{p},\omega)$ is the nonrelativistic retarded propagator defined by:

$$G_0^+(x-y) = i\int \frac{\mathrm{d}^3 p}{(2\pi)^3}\int \frac{\mathrm{d}\omega}{2\pi} G_0^+(\mathbf{p},\omega) e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} e^{-i\omega(t_x-t_y)}$$

(For the other pole $p_0 = -\sqrt{\mathbf{p}^2+m^2}$, we define $\omega = -p_0 - m$ and we would see that the antiparticles' propagator reduces to the advanced Green's function in the nonrelativistic limit.)

As shown above, the nonrelativistic free propagator is defined by:

$$G_0^+(x-y) = i\int \frac{\mathrm{d}^3 p}{(2\pi)^3}\int \frac{\mathrm{d}\omega}{2\pi} G_0^+(\mathbf{p},\omega) e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} e^{-i\omega(t_x-t_y)}$$

with:

$$G_0^+(\mathbf{p},\omega) = \frac{i}{\omega - \frac{\mathbf{p}^2}{2m} + i\epsilon}$$

If we use the energies of the noninteracting particles $E_k \equiv \epsilon_k = \frac{\hbar^2 k^2}{2m} = \frac{k^2}{2m}$, we can write it as:

$$G_0^+(\mathbf{p},\omega) = \frac{i}{\omega - \frac{\mathbf{p}^2}{2m} + i\epsilon} = \frac{i}{\omega - E_k + i\epsilon}$$

so

$$G_0^+(k,\omega) = \frac{i}{\omega - E_k + i\epsilon}$$

using $E = \hbar\omega$ we can also write:

$$G_0^+(k,E) = \frac{i}{E - E_k + i\epsilon}$$

Other equivalent ways of representing the propagator:

$$\begin{aligned} G_0^+(x-y) = G_0^+(\mathbf{x},t_x,\mathbf{y},t_y) &= i\int \frac{\mathrm{d}^3 p \mathrm{d}E}{(2\pi\hbar)^4} G_0^+(\mathbf{p},E) e^{\frac{i}{\hbar}\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} e^{-\frac{i}{\hbar}E(t_x-t_y)} = \\ &= i\int \frac{\mathrm{d}^3 k \mathrm{d}\omega}{(2\pi)^4} G_0^+(k,\omega) e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})} e^{-i\omega(t_x-t_y)} \end{aligned}$$

Sometimes it's useful to calculate the mixed representation $G_0^+(k,t)$:

$$G_0^+(k,t) = \int \frac{\mathrm{d}\omega}{2\pi} e^{-i\omega t} G_0^+(k,\omega) = \int \frac{\mathrm{d}\omega}{2\pi} e^{-i\omega t} \frac{i}{\omega - E_k + i\epsilon} = \cdots = \theta_t e^{-i(E_k - i\epsilon)t}$$

(The "$\cdots$" means to use the Residue Theorem and distinguish two cases $t<0$ and $t>0$, thus getting the Heaviside step function θ_t in the result.)

Very often, in practice, one just needs to work with $G_0^+(k,t)$ and $G_0^+(k,\omega)$, here is how to convert between those:

$$\begin{aligned} G_0^+(k,t) &= \int_{-\infty}^{\infty} \frac{\mathrm{d}\omega}{2\pi} e^{-i\omega t} G_0^+(k,\omega) \\ G_0^+(k,\omega) &= \int_{-\infty}^{\infty} \mathrm{d}t\, e^{i\omega t} G_0^+(k,t) \end{aligned}$$

The relation to the contraction of operators is:

$$G_0^+(\mathbf{k},t_2-t_1) = -i\theta_{t_2-t_1} \left\langle \Psi_0 | c_{\mathbf{k}}(t_2) c_{\mathbf{k}}^\dagger(t_1) | \Psi_0 \right\rangle$$

where $|\Psi_0\rangle$ is the ground state wavefunction and:

$$c_{\mathbf{k}}(t) = e^{iH_0 t} c_{\mathbf{k}} e^{-iH_0 t}$$

so to understand the meaning of $G_0^+(\mathbf{k}, t_2 - t_1)$, we write it as:

$$G_0^+(\mathbf{k}, t_2 - t_1) = -i\theta_{t_2-t_1} \langle \Psi_0 | c_{\mathbf{k}}(t_2) c_{\mathbf{k}}^\dagger(t_1) | \Psi_0 \rangle = -i\theta_{t_2-t_1} \langle \Psi_0 | e^{iH_0 t_2} c_{\mathbf{k}} e^{-iH_0(t_2-t_1)} c_{\mathbf{k}}^\dagger e^{-iH_0 t_1} | \Psi_0 \rangle =$$
$$= -i\theta_{t_2-t_1} \left(e^{-iH_0 t_2} |\Psi_0\rangle \right)^\dagger \left(c_{\mathbf{k}} e^{-iH_0(t_2-t_1)} c_{\mathbf{k}}^\dagger e^{-iH_0 t_1} |\Psi_0\rangle \right)$$

which describes the probability amplitude of adding a bare particle at time t_1, removing at time t_2 and regaining the original many-body system (that in the meantime evolved into $e^{-iH_0 t_2} |\Psi_0\rangle$).

Evaluating the Interacting Green Functions

The Green functions below can either be evaluated using the Wick theorem, or using Feynman diagrams and the corresponding Feynman rules.

Example 1

$\mathcal{L}_{ZZH} = \lambda Z_\mu Z^\mu H$, in the first order:

$$\langle \Omega | T \tilde{A}^\mu(k) \tilde{A}^\nu(l) \tilde{\phi}(-p) | \Omega \rangle = i\lambda (2\pi)^4 \delta(k + l - p) \tilde{D}^\mu{}_\alpha(k) \tilde{D}^{\nu\alpha}(l) \tilde{D}(p)$$

Example 2

$\mathcal{L}_{ee\gamma} = -\lambda \bar{e} \gamma^\mu e A_\mu$, in the second order:

$$\langle \Omega | T \tilde{\psi}(k_1) \tilde{\bar{\psi}}(k_2) \tilde{\bar{\psi}}(-p_1) \tilde{\psi}(-p_2) | \Omega \rangle =$$
$$= (-i\lambda)^2 (2\pi)^4 \delta(k_1 + k_2 - p_1 - p_2) \Big[\tilde{S}(k_1) \gamma^\mu \tilde{S}(-k_2) D_{\mu\nu}(k_1 + k_2) \tilde{S}(p_2) \gamma^\nu \tilde{S}(-p_1) +$$
$$+ \tilde{S}(k_1) \gamma^\mu \tilde{S}(-p_1) D_{\mu\nu}(k_1 - p_1) \tilde{S}(p_2) \gamma^\nu \tilde{S}(-k_2) \Big]$$

Example 3

$\mathcal{L}_{\text{int}} = \frac{g}{2\sqrt{2}}\bar{\nu}_e\gamma^\mu(1-\gamma_5)eW_\mu^+ + \text{h.c.}$, in the second order:

$$\langle\Omega|T\tilde{A}^\alpha(r)\tilde{A}^\beta(p)\tilde{\bar{\psi}}(-k)\tilde{\psi}(-l)|\Omega\rangle =$$
$$= \int \mathrm{d}^4x\,\mathrm{d}^4y\,\langle 0|\,T\tilde{A}^\alpha(r)\tilde{A}^\beta(p)\tilde{\bar{\psi}}(-k)\tilde{\psi}(-l)$$
$$i\frac{g}{2\sqrt{2}}\bar{\nu}_e(x)\gamma^\mu(1-\gamma_5)e(x)W_\mu^+(x)$$
$$i\frac{g}{2\sqrt{2}}\bar{\nu}_e(y)\gamma^\nu(1-\gamma_5)e(y)W_\nu^+(y)$$
$$|0\rangle =$$
$$= \left(i\frac{g}{2\sqrt{2}}\right)^2 \int \mathrm{d}^4x\,\mathrm{d}^4y\,\mathrm{d}\hat{r}\,\mathrm{d}\hat{p}\,\mathrm{d}\hat{k}\,\mathrm{d}\hat{l}\,e^{i(\hat{r}r+\hat{p}p-\hat{k}k-\hat{l}l)}$$
$$\langle 0|\,TA^\alpha(\hat{r})A^\beta(\hat{p})\bar{\psi}(\hat{k})\psi(\hat{l})$$
$$\bar{\nu}_e(x)\gamma^\mu(1-\gamma_5)e(x)W_\mu^+(x)$$
$$\bar{\nu}_e(y)\gamma^\nu(1-\gamma_5)e(y)W_\nu^+(y)$$
$$|0\rangle =$$
$$= \left(i\frac{g}{2\sqrt{2}}\right)^2 \int \mathrm{d}^4x\,\mathrm{d}^4y\,\mathrm{d}\hat{r}\,\mathrm{d}\hat{p}\,\mathrm{d}\hat{k}\,\mathrm{d}\hat{l}\,e^{i(\hat{r}r+\hat{p}p-\hat{k}k-\hat{l}l)}$$
$$D^\alpha{}_\mu(\hat{r}-x)D^\mu{}_\nu(\hat{p}-y)S(\hat{l}-x)\gamma^\mu(1-\gamma_5)S(x-y)\gamma^\nu(1-\gamma_5)S(\hat{k}-y) =$$
$$= \left(i\frac{g}{2\sqrt{2}}\right)^2 \int \mathrm{d}^4x\,\mathrm{d}^4y\,\mathrm{d}\hat{r}\,\mathrm{d}\hat{p}\,\mathrm{d}\hat{k}\,\mathrm{d}\hat{l}\,e^{i((\hat{r}+x)r+(\hat{p}+y)p-(\hat{k}+y)k-(\hat{l}+x)l)}$$
$$D^\alpha{}_\mu(\hat{r})D^\mu{}_\nu(\hat{p})S(\hat{l})\gamma^\mu(1-\gamma_5)S(x-y)\gamma^\nu(1-\gamma_5)S(\hat{k}) =$$
$$= \left(i\frac{g}{2\sqrt{2}}\right)^2 \int \mathrm{d}^4x\,\mathrm{d}^4y\,e^{i(xr+yp-yk-xl)}$$
$$\tilde{D}^\alpha{}_\mu(r)\tilde{D}^\mu{}_\nu(p)\tilde{S}(l)\gamma^\mu(1-\gamma_5)S(x-y)\gamma^\nu(1-\gamma_5)\tilde{S}(k) =$$
$$= \left(i\frac{g}{2\sqrt{2}}\right)^2 \int \mathrm{d}^4x\,\mathrm{d}^4y\,e^{i((x+y)r+yp-yk-(x+y)l)}$$
$$\tilde{D}^\alpha{}_\mu(r)\tilde{D}^\mu{}_\nu(p)\tilde{S}(l)\gamma^\mu(1-\gamma_5)S(x)\gamma^\nu(1-\gamma_5)\tilde{S}(k) =$$
$$= \left(i\frac{g}{2\sqrt{2}}\right)^2 \int \mathrm{d}^4y\,e^{i(yr+yp-yk-yl)}$$
$$\tilde{D}^\alpha{}_\mu(r)\tilde{D}^\mu{}_\nu(p)\tilde{S}(l)\gamma^\mu(1-\gamma_5)\tilde{S}(r-l)\gamma^\nu(1-\gamma_5)\tilde{S}(k) =$$
$$= \left(i\frac{g}{2\sqrt{2}}\right)^2 (2\pi)^4\delta(r+p-k-l)\tilde{D}^\alpha{}_\mu(r)\tilde{D}^\beta{}_\nu(p)\tilde{S}(l)\gamma^\mu(1-\gamma_5)\tilde{S}(r-l)\gamma^\nu(1-\gamma_5)\tilde{S}(k)$$

ZZH interaction

Let's calculate the $\mathcal{L}_{ZZH} = \lambda Z_\mu Z^\mu H$ interaction in the SM, where $\lambda = \frac{g^2}{4\cos\theta_W}$. Consider $H(p) \to Z(k) + Z(l)$:

$$\begin{aligned}
\langle f|S|i\rangle = \langle f|iT|i\rangle = \langle kl|iT|p\rangle = \langle\Omega|a^r_{\mathbf{k}} a^s_{\mathbf{l}} a^\dagger_{\mathbf{p}}|\Omega\rangle = \langle\Omega|T a^r_{\mathbf{k}} a^s_{\mathbf{l}} a^\dagger_{\mathbf{p}}|\Omega\rangle = \\
= \langle\Omega|T\epsilon^{r*}_\mu(\mathbf{k})\frac{k^2}{i}\tilde{A}^\mu(k)\epsilon^{s*}_\nu(\mathbf{l})\frac{l^2}{i}\tilde{A}^\nu(l)\frac{1}{\tilde{D}(p)}\tilde{\phi}(-p)|\Omega\rangle = \\
= \frac{\epsilon^{r*}_\mu(\mathbf{k})\epsilon^{s*}_\nu(\mathbf{l})}{\frac{i}{k^2}\frac{i}{l^2}\tilde{D}(p)}\langle\Omega|T\tilde{A}^\mu(k)\tilde{A}^\nu(l)\tilde{\phi}(-p)|\Omega\rangle \\
= \frac{\epsilon^{r*}_\mu(\mathbf{k})\epsilon^{s*}_\nu(\mathbf{l})}{\frac{i}{k^2}\frac{i}{l^2}\tilde{D}(p)} i\lambda(2\pi)^4\delta(k+l-p)\tilde{D}^\mu{}_\alpha(k)\tilde{D}^{\nu\alpha}(l)\tilde{D}(p) \\
= \frac{\epsilon^{r*}_\mu(\mathbf{k})\epsilon^{s*}_\nu(\mathbf{l})}{\frac{i}{k^2}\frac{i}{l^2}\tilde{D}(p)} i\lambda(2\pi)^4\delta(k+l-p)\frac{-ig^\mu{}_\alpha}{k^2}\frac{-ig^{\nu\alpha}}{l^2}\tilde{D}(p) \\
= \epsilon^{r*}_\mu(\mathbf{k})\epsilon^{s*}_\nu(\mathbf{l}) i\lambda(2\pi)^4\delta(k+l-p)g^\mu{}_\alpha g^{\nu\alpha} \\
= \epsilon^{r*}_\mu(\mathbf{k})\epsilon^{s*}_\nu(\mathbf{l}) i\lambda(2\pi)^4\delta(k+l-p)g^{\mu\nu} \\
= i\lambda(2\pi)^4\delta(k+l-p)\epsilon^{r*}_\mu(\mathbf{k})\epsilon^{s\mu*}(\mathbf{l})
\end{aligned}$$

where we used the fact, that the first order contribution of the $\lambda Z_\mu Z^\mu H$ interaction to the interacting Green function is:

$$\langle\Omega|T\tilde{A}^\mu(k)\tilde{A}^\nu(l)\tilde{\phi}(-p)|\Omega\rangle = i\lambda(2\pi)^4\delta(k+l-p)\tilde{D}^\mu{}_\alpha(k)\tilde{D}^{\nu\alpha}(l)\tilde{D}(p)$$

eeH interaction

This is only approximate, it will be fixed soon.

Let's calculate the $\mathcal{L}_{eeH} = -\lambda\bar{e}eH$ interaction in the SM, where $\lambda = \frac{h_e}{\sqrt{2}}$. Consider $H(p) \to e^-(k) + e^+(l)$:

$$\langle f|S|i\rangle = \langle f|iT|i\rangle = \langle kl|iT|p\rangle = \frac{\bar{u}(k)v(l)}{\tilde{S}(k)\tilde{S}(l)}\frac{1}{\tilde{D}(p)}$$

$$\int \mathrm{d}^4x_1 e^{-ipx_1}\int \mathrm{d}^4y_1\mathrm{d}^4y_2 e^{+i(ky_1+ly_2)}\langle 0|T\{\bar{e}(y_1)e(y_2)H(x_1)\}|0\rangle =$$

$$= \frac{\bar{u}(k)v(l)}{\tilde{S}(k)\tilde{S}(l)}\frac{1}{\tilde{D}(p)}$$

$$\int \mathrm{d}^4x_1 e^{-ipx_1}\int \mathrm{d}^4y_1\mathrm{d}^4y_2 e^{+i(ky_1+ly_2)}\int \mathrm{d}^4x(-i\lambda)S(y_1-x)S(y_2-x)D(x_1-x) =$$

$$= (-i\lambda)(2\pi)^4\delta^{(4)}(p-k-l)\bar{u}(k)v(l)$$

where we used the fact, that the only nonzero element of the Green function is

$$\int \mathrm{d}^4x\,\langle 0|T\{\bar{e}(y_1)e(y_2)H(x_1)\bar{e}(x)e(x)H(x)\}|0\rangle$$

ee gamma interaction

This is only approximate, it will be fixed soon.

Let's calculate the $\mathcal{L}_{ee\gamma} = -\lambda \bar{e}\gamma^\mu e A_\mu$ interaction in the SM, where $\lambda = g\sin\theta_W$. Consider $\gamma(p) \to e^-(k) + e^+(l)$:

$$\langle f|S|i\rangle = \langle f|iT|i\rangle = \langle kl|iT|p\rangle = \frac{\bar{u}(k)v(l)}{\tilde{S}(k)\tilde{S}(l)}\frac{\varepsilon_\mu(p)}{\tilde{D}_{\alpha\beta}(p)}$$

$$\int \mathrm{d}^4x_1 e^{-ipx_1}\int \mathrm{d}^4y_1\mathrm{d}^4y_2 e^{+i(ky_1+ly_2)}\,\langle 0|T\{\bar{e}(y_1)e(y_2)A^\mu(x_1)\}|0\rangle =$$

$$= \frac{\bar{u}(k)v(l)}{\tilde{S}(k)\tilde{S}(l)}\frac{\varepsilon_\mu(p)}{\tilde{D}_{\alpha\beta}(p)}$$

$$\int \mathrm{d}^4x_1 e^{-ipx_1}\int \mathrm{d}^4y_1\mathrm{d}^4y_2 e^{+i(ky_1+ly_2)}\int \mathrm{d}^4x(-i\lambda)S(y_2-x)\gamma^\mu S(y_1-x)D^\alpha_\mu(x_1-x) =$$

$$= (2\pi)^4\delta^{(4)}(p-k-l)\bar{u}(k)(-i\lambda)\gamma^\mu v(l)\varepsilon_\mu(p)$$

where we used the fact, that the only nonzero element of the Green function is

$$\int \mathrm{d}^4x\,\langle 0|T\{\bar{e}(y_1)e(y_2)A^\alpha(x_1)\bar{e}(x)\gamma^\mu e(x)A_\mu(x)\}|0\rangle =$$

$$= \pm S(y_2-x)\gamma^\mu S(y_1-x)D^\alpha_\mu(x_1-x)$$

eeee interaction

Let's calculate the $\mathcal{L}_{ee\gamma} = -\lambda \bar{e}\gamma^\mu e A_\mu$ interaction in the SM, where $\lambda = g\sin\theta_W$. Consider $e^-(p_1) + e^+(p_2) \to \gamma(q) \to e^-(k_1) + e^+(k_2)$:

$$\langle f|S|i\rangle = \langle f|iT|i\rangle = \langle k_1k_2|iT|p_1p_2\rangle = \langle\Omega|b^r_{\mathbf{k}_1}d^s_{\mathbf{k}_2}b^{t\dagger}_{\mathbf{p}_1}d^{u\dagger}_{\mathbf{p}_2}|\Omega\rangle =$$

$$= \langle\Omega|Tb^r_{\mathbf{k}_1}d^s_{\mathbf{k}_2}b^{t\dagger}_{\mathbf{p}_1}d^{u\dagger}_{\mathbf{p}_2}|\Omega\rangle =$$

$$= \langle\Omega|T\left[\bar{u}^r(\mathbf{k}_1)\frac{1}{\tilde{S}(k_1)}\tilde{\psi}(k_1)\right]\left[-\bar{\tilde{\psi}}(k_2)\frac{1}{\tilde{S}(-k_2)}v^s(\mathbf{k}_2)\right]\left[\bar{\tilde{\psi}}(-p_1)\frac{1}{\tilde{S}(-p_1)}u^t(\mathbf{p}_1)\right]\left[-\bar{v}^u(\mathbf{p}_2)\frac{1}{\tilde{S}(p_2)}\tilde{\psi}(-p_2)\right]|\Omega\rangle =$$

$$= \left[\bar{u}^r(\mathbf{k}_1)\frac{1}{\tilde{S}(k_1)}\right]\left[\bar{v}^u(\mathbf{p}_2)\frac{1}{\tilde{S}(p_2)}\right]$$

$$\langle\Omega|T\tilde{\psi}(k_1)\bar{\tilde{\psi}}(k_2)\bar{\tilde{\psi}}(-p_1)\tilde{\psi}(-p_2)|\Omega\rangle$$

$$\left[\frac{1}{\tilde{S}(-k_2)}v^s(\mathbf{k}_2)\right]\left[\frac{1}{\tilde{S}(-p_1)}u^t(\mathbf{p}_1)\right] =$$

$$= \left[\bar{u}^r(\mathbf{k}_1)\frac{1}{\tilde{S}(k_1)}\right]\left[\bar{v}^u(\mathbf{p}_2)\frac{1}{\tilde{S}(p_2)}\right]$$

$$(-i\lambda)^2(2\pi)^4\delta(k_1+k_2-p_1-p_2)\Big[\tilde{S}(k_1)\gamma^\mu\tilde{S}(-k_2)D_{\mu\nu}(k_1+k_2)\tilde{S}(p_2)\gamma^\nu\tilde{S}(-p_1)+$$

$$+\tilde{S}(k_1)\gamma^\mu\tilde{S}(-p_1)D_{\mu\nu}(k_1-p_1)\tilde{S}(p_2)\gamma^\nu\tilde{S}(-k_2)\Big]$$

$$\left[\frac{1}{\tilde{S}(-k_2)}v^s(\mathbf{k}_2)\right]\left[\frac{1}{\tilde{S}(-p_1)}u^t(\mathbf{p}_1)\right] =$$

$$= -\lambda^2(2\pi)^4\delta(k_1+k_2-p_1-p_2)\left[\bar{u}^r(\mathbf{k}_1)\gamma^\mu v^s(\mathbf{k}_2)\frac{1}{(k_1+k_2)^2}\bar{v}^u(\mathbf{p}_2)\gamma_\mu u^t(\mathbf{p}_1)+\right.$$

$$\left.+\bar{u}^r(\mathbf{k}_1)\gamma^\mu u^t(\mathbf{p}_1)\frac{1}{(k_1-p_1)^2}\bar{v}^u(\mathbf{p}_2)\gamma_\mu v^s(\mathbf{k}_2)\right]$$

where we used the fact, that the interacting Green function is in the lowest nonzero order equal to:

$$\langle\Omega|T\tilde{\psi}(k_1)\tilde{\bar{\psi}}(k_2)\tilde{\bar{\psi}}(-p_1)\tilde{\psi}(-p_2)|\Omega\rangle =$$
$$= (-i\lambda)^2(2\pi)^4\delta(k_1+k_2-p_1-p_2)\Big[\tilde{S}(k_1)\gamma^\mu\tilde{S}(-k_2)D_{\mu\nu}(k_1+k_2)\tilde{S}(p_2)\gamma^\nu\tilde{S}(-p_1)+$$
$$+\tilde{S}(k_1)\gamma^\mu\tilde{S}(-p_1)D_{\mu\nu}(k_1-p_1)\tilde{S}(p_2)\gamma^\nu\tilde{S}(-k_2)\Big]$$

8.2.3 Low energy theories

Fermi-type theory

This is a low energy ($m_W^2 \gg m_\mu m_e$) model for the EW interactions, that can be derived for example from the muon decay:

$$\mu^- \to e^- + \nu_\mu + \bar{\nu}_e$$

From the SM the relevant Lagrangian is

$$\mathcal{L} = \frac{g}{2\sqrt{2}}(\bar{e}\gamma^\mu(1-\gamma_5)\nu_e W_\mu^-) + \frac{g}{2\sqrt{2}}(\bar{\mu}\gamma^\mu(1-\gamma_5)\nu_\mu W_\mu^-)$$

and one gets the diagram $\mu^- + \bar{\nu}_\mu + \to e^- + \bar{\nu}_e$ and the corresponding matrix element:

$$iM = -i\frac{g^2}{8}[\bar{u}\gamma_\mu(1-\gamma_5)u]\frac{-g^{\mu\nu}+\frac{q^\mu q^\nu}{m_W^2}}{q^2-m_W^2}[\bar{u}\gamma_\nu(1-\gamma_5)v]$$

which when the momentum transfer q is much less than m_w becomes

$$iM = -i\frac{g^2}{8m_W^2}[\bar{u}\gamma^\mu(1-\gamma_5)u][\bar{u}\gamma_\mu(1-\gamma_5)v]$$

but this matrix element can be derived directly from the Lagrangian:

$$\mathcal{L} = -\frac{G_\mu}{\sqrt{2}}[\bar{\psi}_{\nu_\mu}\gamma^\mu(1-\gamma_5)\psi_\mu][\bar{\psi}_e\gamma^\mu(1-\gamma_5)\psi_{\nu_e}]$$

with

$$\frac{G_\mu}{\sqrt{2}} = \frac{g^2}{8m_W^2}$$

This is the universal V-A theory Lagrangian (after adding the h.c. term). Note that the Fermi constant G_F is equal to G_μ.

For the beta decay we get:

$$\mathcal{L} = -\frac{G_\beta}{\sqrt{2}}[\bar{\psi}_p\gamma^\mu(1-f\gamma_5)\psi_n][\bar{\psi}_e\gamma^\mu(1-\gamma_5)\psi_{\nu_e}]$$

where $G_\beta = G_F\cos\theta_C$, $\theta_C = 13^\circ$ is the Cabibbo angle and $f \doteq 1.26$.

8.3 Quantum Electrodynamics (QED)

8.3.1 Local Gauge Invariance

We use a metric with signature +2 in this section.

The Dirac equation for an electron is:

$$\mathcal{L} = \bar{\psi}(i\hbar c\gamma^\mu \partial_\mu - mc^2)\psi$$

Physical quantities like a charge density ($\bar{\psi}\psi$) or a current ($\bar{\psi}\gamma^\mu\psi$), are all invariant if we add a local phase $\Lambda(x)$ to the field (this is called a local U(1) gauge transformation):

$$\begin{aligned}\psi(x) &\to e^{iq\Lambda(x)/\hbar}\psi(x)\\ \bar{\psi}(x) &\to \bar{\psi}(x)e^{-iq\Lambda(x)/\hbar}\end{aligned}$$

Where q is a parameter that measures the strength of the phase transformation (this will be later interpreted as a charge, for example for electrons $q = -|e|$) and $\hbar$ is the Planck constant. And so we require that the Lagrangian is also invariant under the local gauge transformation, because there is no experiment that would change if this local gauge transformation is applied on the wave functions. By putting this gauge transformation into the Lagrangian density, we otain:

$$\begin{aligned}\mathcal{L} \to \bar{\psi}e^{-iq\Lambda(x)/\hbar}(i\hbar c\gamma^\mu\partial_\mu - mc^2)e^{iq\Lambda(x)/\hbar}\psi =\\ = \bar{\psi}(i\hbar c\gamma^\mu(\partial_\mu + iq\partial_\mu\Lambda(x)/\hbar) - mc^2)\psi\end{aligned}$$

The reason the Lagrangian is not invariant is due to the derivative, which does not transform covariantly under a local gauge transformation:

$$\begin{aligned}\bar{\psi}\partial_\mu\psi \to \bar{\psi}e^{-iq\Lambda(x)/\hbar}\partial_\mu e^{iq\Lambda(x)/\hbar}\psi =\\ = \bar{\psi}(\partial_\mu + iq\partial_\mu\Lambda(x)/\hbar)\psi \neq \bar{\psi}\partial_\mu\psi\end{aligned}$$

In order to make the derivative transform covariantly (and thus the Lagrangian gauge invariant), we have to introduce a gauge field, in this case a vector field $A_\mu(x)$, as follows:

$$D_\mu = \partial_\mu - \frac{i}{\hbar}qA_\mu \tag{8.3.1.1}$$

and the field A_μ must transform as $A_\mu \to A_\mu + \partial_\mu\Lambda(x)$. At this level, we are free to choose either plus or minus sign in (8.3.1.1), since the sign change can be absorbed in the definition of the A_μ field without loss of generality (if we change the sign, the field transformation then changes to $A_\mu \to A_\mu - \partial_\mu\Lambda(x)$). In the +2 metric signature we chose a minus sign, so that A_μ coincides with the usual definition of the electromagnetic 4-potential:

$$\begin{aligned}D_\mu &= \partial_\mu - \frac{i}{\hbar}qA_\mu\\ -i\hbar D_\mu &= -i\hbar\partial_\mu - qA_\mu\\ m\hat{v}_\mu &= \hat{p}_\mu - qA_\mu\\ m\hat{\mathbf{v}} &= \hat{\mathbf{p}} - q\mathbf{A}\end{aligned}$$

With signature -2, we must choose a plus sign and the identification goes as follows:

$$\begin{aligned}D_\mu &= \partial_\mu + \frac{i}{\hbar}qA_\mu\\ i\hbar D_\mu &= i\hbar\partial_\mu - qA_\mu\\ m\hat{v}_\mu &= \hat{p}_\mu - qA_\mu\\ m\hat{\mathbf{v}} &= \hat{\mathbf{p}} - q\mathbf{A}\end{aligned}$$

And we obtain the same final equation. So the kinematic momentum is equal to canonical momentum minus charge times the gauge field. The last expression is independent of a metric signature, and that is what is e.g. in the kinetic term of a Schrödinger or Pauli equation (with the minus sign in $\hat{\mathbf{p}} - q\mathbf{A}$). We derive the non-relativistic limit rigorously later, but it gives the same result. At this level we just have to make sure we choose the correct sign in (8.3.1.1), depending on the metric signature, otherwise we would get the electromagnetic 4-potential with the opposite sign (the sign of A_μ is ultimately just a convention, but later we want to get the same equations as everybody else).

Another unrelated convention is in choosing the sign of the parameter q. We have choosen it to coincide with an electric charge (negative for electrons). Some authors choose q to be positive for electrons, then one must flip the sign in (8.3.1.1).

We will continue using the +2 signature in the rest of the section.

The operator $D_\mu = \partial_\mu - \frac{i}{\hbar} q A_\mu$ is called a covariant derivative, because it does not change a form (is invariant) under a local gauge transformation:

$$\begin{aligned}
\bar{\psi} D_\mu \psi &= \bar{\psi}(\partial_\mu - \frac{i}{\hbar} q A_\mu)\psi \\
\rightarrow \bar{\psi} e^{-iq\Lambda(x)/\hbar}(\partial_\mu - \frac{i}{\hbar} q(A_\mu + \partial_\mu \Lambda(x))) e^{iq\Lambda(x)/\hbar}\psi &= \\
= \bar{\psi}(\partial_\mu - \frac{i}{\hbar} q A_\mu - \frac{i}{\hbar} q \partial_\mu \Lambda(x) + iq\partial_\mu \Lambda(x)/\hbar)\psi &= \\
= \bar{\psi}(\partial_\mu - \frac{i}{\hbar} q A_\mu)\psi = \bar{\psi} D_\mu \psi &
\end{aligned}$$

Then the Lagrangian

$$\mathcal{L} = \bar{\psi}(i\hbar c \gamma^\mu(\partial_\mu - iqA_\mu/\hbar) - mc^2)\psi \tag{8.3.1.2}$$

is also gauge invariant:

$$\begin{aligned}
\mathcal{L} \rightarrow \bar{\psi} e^{-iq\Lambda(x)/\hbar}(i\hbar c \gamma^\mu(\partial_\mu - iqA_\mu/\hbar - iq\partial_\mu \Lambda(x)/\hbar) - mc^2) e^{iq\Lambda(x)/\hbar}\psi &= \\
= \bar{\psi}(i\hbar c \gamma^\mu(\partial_\mu - iqA_\mu/\hbar - iq\partial_\mu \Lambda(x)/\hbar + iq\partial_\mu \Lambda(x)/\hbar) - mc^2)\psi &= \\
= \bar{\psi}(i\hbar c \gamma^\mu(\partial_\mu - iqA_\mu/\hbar) - mc^2)\psi &
\end{aligned}$$

The Lagrangian (8.3.1.2) can also be written as:

$$\begin{aligned}
\mathcal{L} = \bar{\psi}(i\hbar c \gamma^\mu(\partial_\mu - \frac{i}{\hbar} q A_\mu) - mc^2)\psi &= \\
= \bar{\psi}(i\hbar c \gamma^\mu \partial_\mu - mc^2)\psi + qc\bar{\psi}\gamma^\mu \psi A_\mu &
\end{aligned}$$

We can see that the condition of a local gauge invariance requires an interaction with a vector field A_μ. Now we need to add the kinetic term for the field A_μ:

$$-\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$$

The mass term $\frac{1}{2} m^2 A_\mu A^\mu$ is not gauge invariant, and so we have to set $m = 0$. Here is the full Lagrangian:

$$\mathcal{L} = \bar{\psi}(i\hbar c \gamma^\mu \partial_\mu - mc^2)\psi + qc\bar{\psi}\gamma^\mu \psi A_\mu - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}$$

This is a Lagrangian for an electron and a massless vector boson (photon) of spin 1. We can introduce a current $j^\mu = c\bar{\psi}\gamma^\mu\psi$, then the Lagrangian density becomes:

$$\mathcal{L} = \bar{\psi}(i\hbar c \gamma^\mu \partial_\mu - mc^2)\psi + qj^\mu A_\mu - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}$$

For an electron, we can set $q = -e$, where e is the elementary charge (e is positive).

8.3.2 QED Lagrangian

We use a metric with signature -2 in this section.

The QED Lagrangian density is

$$\mathcal{L} = \bar{\psi}(i\hbar c\gamma^\mu D_\mu - mc^2)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

where

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}$$

and we must choose a plus sign in (8.3.1.1) since we use the -2 signature:

$$D_\mu = \partial_\mu + \frac{i}{\hbar}eA_\mu$$

e is the charge (negative for electrons $e = -|e|$).

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

is the electromagnetic field tensor. It's astonishing, that this simple Lagrangian can account for all phenomena from macroscopic scales down to something like $10^{-13}\,\mathrm{cm}$. So it's not a surprise that Feynman, Schwinger and Tomonaga received the 1965 Nobel Prize in Physics for such a fantastic achievement.

Plugging this Lagrangian into the Euler-Lagrange equation of motion for a field, we get:

$$(i\hbar c\gamma^\mu D_\mu - mc^2)\psi = 0$$

$$\partial_\nu F^{\nu\mu} = -ec\bar{\psi}\gamma^\mu\psi$$

The first equation is the Dirac equation in the electromagnetic field and the second equation is a set of Maxwell equations ($\partial_\nu F^{\nu\mu} = -ej^\mu$) with a source $j^\mu = c\bar{\psi}\gamma^\mu\psi$, which is a 4-current comming from the Dirac equation.

8.3.3 Magnetic moment of an electron

In this section we derive the order-α correction to the magnetic moment of an electron.

We start by computing the electron vertex function for the process $\gamma(q) \to e^+(p) + e^-(p')$:

$$iM = ie^2\left(\bar{u}(p')\Gamma^\mu(p',p)u(p)\right)\frac{1}{q^2}\left(\bar{u}(k')\gamma_\mu u(k)\right)$$

where k corresponds to some heavy target. If A^{cl}_μ is a fixed classical potential, we get:

$$iM2\pi\delta(p^{0'} - p^0) = -ie\bar{u}(p')\Gamma^\mu(p',p)u(p)A^{\mathrm{cl}}_\mu$$

Using general arguments (Lorentz invariance, parity-conservation, Ward identity) we can always write Γ^μ as:

$$\Gamma^\mu(p',p) = \gamma^\mu F_1(q^2) + \frac{i\sigma^{\mu\nu}q_\nu}{2m}F_2(q^2)$$

where F_1 and F_2 ar unknown functions of $q^2 = (p'-p)^2 = -2p'\cdot p + 2m^2$ called form factors. As we will see below, in the lowest order we get $F_1 = 1$ and $F_2 = 0$.

We can calculate the amplitude for elastic Coulomb scattering of a nonrelativistic electron from a region of nonzero electrostatic potential by setting $A^{\rm cl}_\mu(x) = (\phi(\mathbf{x}), 0)$, then:

$$\begin{aligned} A^{\rm cl}_\mu(q) &= (2\pi\delta(q^0)\tilde{\phi}(\mathbf{q}), 0) \\ iM 2\pi\delta(p^{0'} - p^0) &= -ie\bar{u}(p')\Gamma^0(p', p)u(p)2\pi\delta(q^0)\tilde{\phi}(\mathbf{q}) \\ iM &= -ie\bar{u}(p')\Gamma^0(p', p)u(p)\tilde{\phi}(\mathbf{q}) \end{aligned}$$

If the electrostatic field is very slowly varying over a large (even macroscopic) region, $\tilde{\phi}(\mathbf{q})$ will be concentrated about $\mathbf{q} = 0$, then we can take the limit $\mathbf{q} \to 0$:

$$\begin{aligned} iM &= -ie\bar{u}(p')\Gamma^0(p', p)u(p)\tilde{\phi}(\mathbf{q}) \\ iM &= -ie\bar{u}(p')\left(\gamma^0 F_1(q^2) + \frac{i\sigma^{0\nu}q_\nu}{2m}F_2(q^2)\right)u(p)\tilde{\phi}(\mathbf{q}) \\ iM &= -ie\bar{u}(p')\gamma^0 u(p)F_1(0)\tilde{\phi}(\mathbf{q}) \\ iM &= -ie2m\xi^{'\dagger}\xi F_1(0)\tilde{\phi}(\mathbf{q}) \\ iM &= -i\left(eF_1(0)\tilde{\phi}(\mathbf{q})\right)2m\xi^{'\dagger}\xi \end{aligned}$$

This corresponds to the Born approximation for scattering from a potential

$$V(\mathbf{x}) = eF_1(0)\phi(\mathbf{x})$$

Thus $F_1(0)$ is the electric charge of the electron, in units of e. Since $F_1(0) = 1$ already in the first order of perturbation theory, radiative corrections to $F_1(q^2)$ must vanish at $q^2 = 0$.

Now we calculate the scattering from a static vector potential by setting $A^{\rm cl}_\mu(x) = (0, \mathbf{A}^{\rm cl}_\mu(\mathbf{x}))$, then:

$$\begin{aligned} A^\mu_{\rm cl}(q) &= (0, 2\pi\delta(q^i)\tilde{A}^i_{\rm cl}(\mathbf{q})) \\ iM 2\pi\delta(p^{'i} - p^i) &= ie\bar{u}(p')\Gamma^i(p', p)u(p)2\pi\delta(q^i)\tilde{A}^i_{\rm cl}(\mathbf{q}) \\ iM &= ie\bar{u}(p')\Gamma^i(p', p)u(p)\tilde{A}^i_{\rm cl}(\mathbf{q}) \\ iM &= ie\bar{u}(p')\left(\gamma^i F_1(q^2) + \frac{i\sigma^{i\nu}q_\nu}{2m}F_2(q^2)\right)u(p)\tilde{A}^i_{\rm cl}(\mathbf{q}) \end{aligned}$$

In the limit $q \to 0$ this becomes:

$$\begin{aligned} iM &= ie2m\xi^{'\dagger}\left(-i\epsilon^{ijk}\frac{q^j\sigma^k}{2m}(F_1(0) + F_2(0))\right)\xi\tilde{A}^i_{\rm cl}(\mathbf{q}) \\ iM &= -ie2m\xi^{'\dagger}\left(-\frac{\sigma^k}{2m}(F_1(0) + F_2(0))\right)\xi\left(-i\epsilon^{ijk}q^j\tilde{A}^i_{\rm cl}(\mathbf{q})\right) \\ iM &= -ie2m\xi^{'\dagger}\left(-\frac{\sigma^k}{2m}(F_1(0) + F_2(0))\right)\xi\tilde{B}^k(\mathbf{q}) \\ iM &= -i\left(-\frac{e}{m}(F_1(0) + F_2(0))2m\xi^{'\dagger}\frac{\sigma^k}{2}\xi\tilde{B}^k(\mathbf{q})\right) \end{aligned}$$

where

$$\tilde{B}^k(\mathbf{q}) = \left(-i\epsilon^{ijk}q^j\tilde{A}^i_{\rm cl}(\mathbf{q})\right)$$

is the Fourier transform of the magnetic field produced by $\mathbf{A}^{\rm cl}(\mathbf{x})$.

This corresponds to the Born approximation for scattering from a potential

$$\begin{aligned} V(\mathbf{x}) &= -\frac{e}{m}(F_1(0) + F_2(0))\xi^{'\dagger}\frac{\sigma^k}{2}\xi B^k(\mathbf{x}) \\ V(\mathbf{x}) &= -\frac{e}{m}(F_1(0) + F_2(0))\xi^{'\dagger}\frac{\boldsymbol{\sigma}}{2}\xi\cdot\mathbf{B}(\mathbf{x}) \\ V(\mathbf{x}) &= -<\boldsymbol{\mu}>\cdot\mathbf{B}(\mathbf{x}) \end{aligned}$$

where

$$< \boldsymbol{\mu} >= \frac{e}{m}(F_1(0) + F_2(0))\xi'^{\dagger}\frac{\boldsymbol{\sigma}}{2}\xi$$
$$< \boldsymbol{\mu} >= g\frac{e}{2m}\mathbf{S}$$

where

$$g = 2(F_1(0) + F_2(0))$$
$$\mathbf{S} = \xi'^{\dagger}\frac{\boldsymbol{\sigma}}{2}\xi$$

The coefficient g is called the Landé g-factor, and since the leading order of perturbation theory gives $F_2(0) = 0$ (and we know that $F_1(0) = 1$ to all orders), we get:

$$g = 2(F_1(0) + F_2(0)) = 2 + 2F_2(0) = 2 + O(\alpha)$$

This is the standard prediction of the Dirac equation. The anomalous magnetic moment is then:

$$a_e = \frac{g-2}{2} = F_2(0)$$

To calculate that, we need to evaluate the one-loop correction to the vertex function, so we start by deriving the appropriate Green function for the process $\gamma(q) + e^+(p) \to e^+(p')$:

$$|i\rangle = a^{r\dagger}_{\mathbf{q}} b^{t\dagger}_{\mathbf{p}} |\Omega\rangle$$
$$|f\rangle = b^{s\dagger}_{\mathbf{p}'} |\Omega\rangle$$
$$\langle f|i\rangle = \langle\Omega| b^s_{\mathbf{p}'} a^{r\dagger}_{\mathbf{q}} b^{t\dagger}_{\mathbf{p}} |\Omega\rangle =$$
$$= \langle\Omega| T b^s_{\mathbf{p}'} a^{r\dagger}_{\mathbf{q}} b^{t\dagger}_{\mathbf{p}} |\Omega\rangle =$$
$$= \langle\Omega| T\bar{u}^s(\mathbf{p}')\frac{1}{\tilde{S}(p')}\tilde{\psi}(p')\epsilon^{r*}_{\mu}(\mathbf{q})\frac{q^2}{i}\tilde{A}^{\mu}(-q)\tilde{\bar{\psi}}(-p)\frac{1}{\tilde{S}(-p)}u^t(\mathbf{p}) |\Omega\rangle =$$
$$= \bar{u}^s(\mathbf{p}')\frac{1}{\tilde{S}(p')}\epsilon^{r*}_{\mu}(\mathbf{q})\frac{q^2}{i}\langle\Omega| T\tilde{\psi}(p')\tilde{A}^{\mu}(-q)\tilde{\bar{\psi}}(-p) |\Omega\rangle \frac{1}{\tilde{S}(-p)}u^t(\mathbf{p}) =$$
$$= \bar{u}^s(\mathbf{p}')\frac{1}{\tilde{S}(p')}\epsilon^{r*}_{\mu}(\mathbf{q})\frac{q^2}{i}\tilde{G}(p,p',q)\frac{1}{\tilde{S}(-p)}u^t(\mathbf{p}) =$$

where:

$$\tilde{G}(p,p',q) = \langle\Omega| T\tilde{\psi}(p')\tilde{A}^{\mu}(-q)\tilde{\bar{\psi}}(-p) |\Omega\rangle$$

is the interacting Green function for the Lagrangian $-\lambda\bar{e}\gamma^{\mu}eA_{\mu}$. In the first order:

$$\tilde{G}(p,p',q) = \langle\Omega| T\tilde{\psi}(p')\tilde{A}^{\mu}(-q)\tilde{\bar{\psi}}(-p) |\Omega\rangle =$$
$$= \int \mathrm{d}^4x \langle 0| T\tilde{\psi}(p')\tilde{A}^{\mu}(-q)\tilde{\bar{\psi}}(-p)(-\lambda)\bar{e}(x)\gamma^{\rho}e(x)A_{\rho}(x) |0\rangle =$$
$$= (-\lambda)\int \mathrm{d}^4x\mathrm{d}\hat{p}'\mathrm{d}\hat{q}\mathrm{d}\hat{p}e^{i\hat{p}'p'-\hat{q}q-\hat{p}p} \langle 0| T\psi(\hat{p}')A^{\mu}(\hat{q})\bar{\psi}(\hat{p})\bar{e}(x)\gamma^{\rho}e(x)A_{\rho}(x) |0\rangle =$$
$$= (-\lambda)\int \mathrm{d}^4x\mathrm{d}\hat{p}'\mathrm{d}\hat{q}\mathrm{d}\hat{p}e^{i\hat{p}'p'-\hat{q}q-\hat{p}p}D^{\mu}_{\rho}(\hat{q}-x)S(\hat{p}'-x)\gamma^{\rho}S(\hat{p}-x) =$$
$$= (-\lambda)(2\pi)^4\delta(p'-q-p)\tilde{D}^{\mu}_{\rho}(q)\tilde{S}(p')\gamma^{\rho}\tilde{S}(p)$$

so the amplitude is:

$$\langle f|i\rangle = \bar{u}^s(\mathbf{p}')\frac{1}{\tilde{S}(p')}\epsilon^{r*}_{\mu}(\mathbf{q})\frac{q^2}{i}(-\lambda)(2\pi)^4\delta(p'-q-p)\tilde{D}^{\mu}_{\rho}(q)\tilde{S}(p')\gamma^{\rho}\tilde{S}(p)\frac{1}{\tilde{S}(-p)}u^t(\mathbf{p}) =$$
$$= (-\lambda)(2\pi)^4\delta(p'-q-p)\epsilon^{r*}_{\mu}(\mathbf{q})u^s(\mathbf{p}')\gamma^{\mu}u^t(\mathbf{p})$$

and we got $\Gamma^\mu = \gamma^\mu$, so $F_1 = 1$ and $F_2 = 0$ in the lowest order. In the next order we get:

$$\tilde{G}(p, p', q) = (-\lambda)(2\pi)^4 \delta(p' - q - p)\tilde{D}^\mu_\rho(q)\tilde{S}(p')\delta\Gamma^\rho\tilde{S}(p)$$
$$\delta\Gamma^\mu = \int \frac{\mathrm{d}^4 k}{(2\pi)^4}\tilde{D}_{\nu\rho}(k-p)(-ie\gamma^\nu)\tilde{S}(k')\gamma^\mu\tilde{S}(k)(-ie\gamma^\rho)$$

Now we can write:

$$\bar{u}(p')\Gamma^\mu(p', p)u(p) = \bar{u}(p')(\gamma^\mu +$$
$$\bar{u}(p')\delta\Gamma^\mu(p', p)u(p) = \int \frac{\mathrm{d}^4 k}{(2\pi)^4}\tilde{D}_{\nu\rho}(k-p)\bar{u}(p')(-ie\gamma^\nu)\tilde{S}(k')\gamma^\mu\tilde{S}(k)(-ie$$
$$= \int \frac{\mathrm{d}^4 k}{(2\pi)^4}\frac{-ig_{\nu\rho}}{(k-p)^2 + i\epsilon}\bar{u}(p')(-ie\gamma^\nu)\frac{i(\not{k}' + m)}{k'^2 - m^2 + i\epsilon}\gamma^\mu\frac{i(\not{k} + m)}{k^2 - m^2 + i\epsilon}(-ie$$
$$= 2ie^2\int \frac{\mathrm{d}^4 k}{(2\pi)^4}\frac{\bar{u}(p')\left(\not{k}\gamma^m u\not{k}' + m^2\gamma^\mu - 2m(k + k')\right.}{((k-p)^2 + i\epsilon)(k'^2 - m^2 + i\epsilon)(k^2 - m$$

$$= 2ie^2\int \frac{\mathrm{d}^4 l}{(2\pi)^4}\int_0^1 \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\delta(x + y + z - 1)\frac{2}{D^3}\bar{u}(p')\left(\gamma^\mu(-\tfrac{1}{2}l^2 + (1-x)(1-y)q^2 + (1 - 4z + z^2)m^2) + \frac{i\sigma^{\mu\nu}q_\nu}{2m}(2m^2z(1-z\right.$$
$$= \frac{\alpha}{2\pi}\int_0^1 \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\delta(x + y + z - 1)\bar{u}(p')\left(\gamma^\mu\left[\log\frac{z\Lambda^2}{\Delta} + \frac{1}{\Delta}\left((1-x)(1-y)q^2 + (1 - 4z + z^2)m^2\right)\right] + \frac{i\sigma^{\mu\nu}q_\nu}{2m}\left[\frac{1}{\Delta}2m^2z(1-\right.$$

where

$$k' = k + q$$
$$D = l^2 - \Delta + i\epsilon$$
$$\Delta = -xyq^2 + (1-z)^2m^2 > 0$$

So the expressions for the form factors are:

$$F_1(q^2) = 1 + \frac{\alpha}{2\pi}\int_0^1 \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\delta(x + y + z - 1)\left[\log\frac{z\Lambda^2}{\Delta} + \frac{1}{\Delta}\left((1-x)(1-y)q^2 + (1 - 4z + z^2)m^2\right)\right] + O(\alpha^2)$$
$$F_2(q^2) = \frac{\alpha}{2\pi}\int_0^1 \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\delta(x + y + z - 1)\left[\frac{1}{\Delta}2m^2z(1-z)\right] + O(\alpha^2) =$$
$$= \frac{\alpha}{2\pi}\int_0^1 \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\delta(x + y + z - 1)\left[\frac{2m^2z(1-z)}{m^2(1-z)^2 - q^2xy}\right] + O(\alpha^2)$$

F_1 contains both ultraviolet and infrared divergencies. To cure the infrared divergence, we add a term $\mu^2 z$ to Δ. To cure the ultraviolet divergence, we make the substitution:

$$F_1(q^2) \to F_1(q^2) - \delta F_1(0)$$

where δF_1 is the first order (in α) correction to F_1 (i.e. $F_1 = 1 + \delta F_1 + O(\alpha^2)$):

$$\delta F_1(0) = \frac{\alpha}{2\pi}\int_0^1 \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\delta(x + y + z - 1)\left[\log\frac{z\Lambda^2}{\Delta(q^2 = 0)} + \frac{1}{\Delta(q^2 = 0)}(1 - 4z + z^2)m^2\right]$$

so the corrected F_1 is:

$$F_1(q^2) = 1 + \frac{\alpha}{2\pi}\int_0^1 \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\delta(x + y + z - 1)\left[\log\frac{z\Lambda^2}{\Delta} + \frac{1}{\Delta}\left((1-x)(1-y)q^2 + (1 - 4z + z^2)m^2\right) + \right.$$
$$\left. - \log\frac{z\Lambda^2}{\Delta(q^2 = 0)} - \frac{1}{\Delta(q^2 = 0)}(1 - 4z + z^2)m^2\right] + O(\alpha^2) =$$
$$= 1 + \frac{\alpha}{2\pi}\int_0^1 \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\delta(x + y + z - 1)\left[\log\frac{m^2(1-z)^2}{m^2(1-z)^2 - q^2xy} + \left(\frac{(1-x)(1-y)q^2 + (1 - 4z + z^2)m^2}{m^2(1-z)^2 - q^2xy + \mu^2z}\right) + \right.$$
$$\left. - \frac{(1 - 4z + z^2)m^2}{m^2(1-z)^2 + \mu^2z}\right] + O(\alpha^2)$$

Neither the ultraviolet nor the infrared divergence affects $F_2(q^2)$, so we just set $q = 0$:

$$F_2(0) = \frac{\alpha}{2\pi}\int_0^1 \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\delta(x+y+z-1)\left[\frac{2m^2 z(1-z)}{m^2(1-z)^2}\right] + O(\alpha^2) =$$
$$= \frac{\alpha}{2\pi}\int_0^1 \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\delta(x+y+z-1)\frac{2z}{1-z} + O(\alpha^2) =$$
$$= \frac{\alpha}{2\pi}\int_0^1 \mathrm{d}y\int_0^1 \mathrm{d}z\,\theta(1-(1-y-z))\theta((1-y-z)-0)\frac{2z}{1-z} + O(\alpha^2) =$$
$$= \frac{\alpha}{2\pi}\int_0^1 \mathrm{d}y\int_0^1 \mathrm{d}z\,\theta(y+z)\theta(1-y-z)\frac{2z}{1-z} + O(\alpha^2) =$$
$$= \frac{\alpha}{2\pi}\int_0^1 \mathrm{d}y\int_0^1 \mathrm{d}z\,\theta(1-y-z)\frac{2z}{1-z} + O(\alpha^2) =$$
$$= \frac{\alpha}{2\pi}\int_0^1 \mathrm{d}z\int_0^{1-z} \mathrm{d}y\frac{2z}{1-z} + O(\alpha^2) =$$
$$= \frac{\alpha}{2\pi}\int_0^1 \mathrm{d}z(1-z)\frac{2z}{1-z} + O(\alpha^2) =$$
$$= \frac{\alpha}{2\pi}\int_0^1 \mathrm{d}z 2z + O(\alpha^2) =$$
$$= \frac{\alpha}{2\pi} + O(\alpha^2)$$

Thus we get the correction to the g-factor of the electron:

$$a_e = \frac{g-2}{2} = F_2(0) = \frac{\alpha}{2\pi} \approx 0.0011614$$

Code:

```
>>> from math import pi
>>> alpha = 1/137.035999049
>>> a_e = alpha / (2*pi)
>>> a_e
0.0011614097331824923
```

Experiments give $a_e = 0.00115965218073 \pm 0.00000000000028$ (arXiv:1412.8284, eq. (1)).

Higher order corrections from QED can also be calculated:

$$a_e = A_1\left(\frac{\alpha}{\pi}\right) + A_2\left(\frac{\alpha}{\pi}\right)^2 + A_3\left(\frac{\alpha}{\pi}\right)^3 + A_4\left(\frac{\alpha}{\pi}\right)^4 + \cdots$$

we already know that $A_1 = \frac{1}{2}$. See for example hep-ph/9410248 for the expression for A_2:

$$A_2 = \frac{197}{144} + \frac{3}{4}\zeta(3) - \frac{1}{2}\pi^2\log(2) + \frac{1}{12}\pi^2 =$$
$$= -0.328478965579\ldots$$

Code:

```
>>> from sympy import zeta, S, log
>>> A_2 = S(197)/144 + zeta(2)/2 + 3*zeta(3)/4 - 3*zeta(2) * log(2)
>>> A_2.n()
-0.328478965579194
```

See hep-ph/9602417 for the A_3 term:

$$A_3 = \frac{28259}{5184} - \frac{215}{24}\zeta(5) + \frac{100}{3}\left(\sum_{n=1}^{\infty}\frac{1}{2^n n^4} - \frac{1}{24}\pi^2\log^2(2) + \frac{1}{24}\log^4(2)\right) + \\ + \frac{139}{18}\zeta(3) - \frac{298}{9}\pi^2\log(2) + \frac{83}{72}\pi^2\zeta(3) + \frac{17101}{810}\pi^2 - \frac{239}{2160}\pi^4 = \\ = 1.181241456\ldots$$

Code:

```
>>> from sympy import pi, zeta, S, log, sum, var, oo
>>> var("n")
n
>>> a4 = sum(1/(2**n * n**4), (n, 1, oo))
>>> A_3 = 83*pi**2*zeta(3)/72 - 215*zeta(5)/24 + 100*(a4 + log(2)**4/24 - \
...          pi**2*log(2)**2/24)/3 - \
...          239*pi**4/2160 + 139*zeta(3)/18 - 298 * pi**2 * log(2)/9 + \
...          17101 * pi**2 / 810 + S(28259)/5184
>>> A_3.n()
1.18124145658720
```

Higher terms are only known numerically. The A_4 and A_5 terms can be found in arXiv:1412.8284:

$$A_4 = -1.91298(84)$$
$$A_5 = 7.795(336)$$

We can now sum a_e up to a given order by the following script:

```
from sympy import pi, zeta, S, log, summation, var, oo
var("n")
a4 = summation(1/(2**n * n**4), (n, 1, oo))
A1 = S(1)/2
A2 = S(197)/144 + zeta(2)/2 + 3*zeta(3)/4 - 3*zeta(2) * log(2)
A3 = 83*pi**2*zeta(3)/72 - 215*zeta(5)/24 + 100*(a4 + log(2)**4/24 - \
                pi**2*log(2)**2/24)/3 - \
                239*pi**4/2160 + 139*zeta(3)/18 - 298 * pi**2 * log(2)/9 + \
                17101 * pi**2 / 810 + S(28259)/5184
A4 = -1.91298
A5 = 7.795
alpha = 1/137.035999049
a_e_exp = 0.00115965218073
a_e_exp_err = 0.00000000000028
a_e_other = 0.00000000000448
A = [A1, A2, A3, A4, A5]
a_e= []
for i in range(len(A)):
        a_e.append((A[i]*(alpha/pi)**(i+1)).n())
print "========== ================"
print "Order      :math:`a_e`"
print "========== ================"
for i in range(len(A)):
        print "%d          %16.14f" % (i+1, sum(a_e[:i+1]))
print "Other      %16.14f" % a_e_other
print "Total      %16.14f" % (sum(a_e) + a_e_other)
print "Experiment %16.14f" % a_e_exp
print "Difference %16.14f" % abs(sum(a_e) + a_e_other - a_e_exp)
print "Exp. err   %16.14f" % a_e_exp_err
print "========== ================"
```

and obtain the following table:

Order	a_e
1	0.00116140973318
2	0.00115963742812
3	0.00115965223232
4	0.00115965217663
5	0.00115965217716
Other	0.00000000000448
Total	0.00115965218164
Experiment	0.00115965218073
Difference	0.00000000000091
Exp. err	0.00000000000028

The "Other" line are contributions from the dependence on the muon and tau particle masses, the hadronic vacuum-polarization, the hadronic light-by-light-scattering and the electroweak contribution (see arXiv:1412.8284). The "Difference" line is the difference from the theory (the "Total" line) and experiment. The "Exp. err" line is the experimental error.

At this level of accuracy, the uncertainty of the exact value of α is the primary cause of the difference from experiment, and one can use this result to predict a more accurate value for α, assuming that QED and the standard model are valid.

8.4 Quantum Mechanics

We use a metric with signature -2 in this section.

8.4.1 From QED to Quantum Mechanics

The QED Lagrangian density is

$$\mathcal{L} = \bar{\psi}(i\hbar c\gamma^\mu D_\mu - mc^2)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

Plugging this Lagrangian into the Euler-Lagrange equation of motion for a field, we get:

$$(i\hbar c\gamma^\mu D_\mu - mc^2)\psi = 0$$

$$\partial_\nu F^{\nu\mu} = -ec\bar{\psi}\gamma^\mu\psi$$

The first equation is the Dirac equation in the electromagnetic field and the second equation is a set of Maxwell equations ($\partial_\nu F^{\nu\mu} = -ej^\mu$) with a source $j^\mu = c\bar{\psi}\gamma^\mu\psi$, which is a 4-current comming from the Dirac equation.

The fields ψ and A^μ are quantized. The first approximation is that we take ψ as a wavefunction, that is, it is a classical 4-component field. It can be shown that this corresponds to taking the tree diagrams in the perturbation theory.

We multiply the Dirac equation by γ^0 from left to get:

$$0 = \gamma^0(i\hbar c\gamma^\mu D_\mu - mc^2)\psi = \gamma^0(i\hbar c\gamma^0(\partial_0 + \frac{i}{\hbar}eA_0) + ic\gamma^i(\partial_i + \frac{i}{\hbar}eA_i) - mc^2)\psi =$$

$$= (i\hbar c\partial_0 + i\hbar c\gamma^0\gamma^i\partial_i - \gamma^0 mc^2 - ceA_0 - ce\gamma^0\gamma^i A_i)\psi$$

and we make the following substitutions (it's just a formalism, nothing more): $\beta = \gamma^0$, $\alpha^i = \gamma^0\gamma^i$, $p_j = i\hbar\partial_j$, $\partial_0 = \frac{1}{c}\frac{\partial}{\partial t}$ to get

$$(i\hbar\frac{\partial}{\partial t} + c\alpha^i p_i - \beta mc^2 - ceA_0 - ce\alpha^i A_i)\psi = 0\,.$$

or:

$$i\hbar\frac{\partial\psi}{\partial t} = (c\alpha^i(-p_i + eA_i) + \beta mc^2 + ceA_0)\psi\,.$$

This can be written as:

$$i\frac{\partial\psi}{\partial t} = H\psi\,,$$

where the Hamiltonian is given by:

$$H = c\alpha^i(-p_i + eA_i) + \beta mc^2 + ceA_0\,,$$

or introducing the electrostatic potential $\phi = cA_0$ and writing the momentum as a vector (see the appendix for all the details regarding signs):

$$H = c\boldsymbol{\alpha}\cdot(\mathbf{p} - e\mathbf{A}) + \beta mc^2 + e\phi\,.$$

The right hand side of the Maxwell equations is the 4-current, so it's given by:

$$j^\mu = c\bar{\psi}\gamma^\mu\psi$$

Now we make the substitution $\psi = e^{-imc^2t}\varphi$, which states, that we separate the largest oscillations of the wavefunction and we get

$$j^0 = c\bar{\psi}\gamma^0\psi = c\psi^\dagger\psi = c\varphi^\dagger\varphi$$

$$j^i = c\bar{\psi}\gamma^i\psi = c\psi^\dagger\alpha^i\psi = c\varphi^\dagger\alpha^i\varphi$$

Derivation of the Pauli Equation

We start from the Dirac equation:

$$H\psi = W\psi$$

where:

$$H = c\boldsymbol{\alpha}\cdot(\mathbf{p} - e\mathbf{A}) + \beta mc^2 + e\phi\,.$$
$$W = E + mc^2$$

W is the relativistic energy, E is the nonrelativistic energy, $e\phi = V$ is the potential. The matrices $\boldsymbol{\alpha}$ and β are given by:

$$\alpha^i = \gamma^0\gamma^i$$
$$\beta = \gamma^0$$

so written explicitly:

$$\alpha = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}$$
$$\beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}$$

And the Dirac equation is:

$$\begin{pmatrix} V + mc^2 & c\boldsymbol{\sigma}\cdot(\mathbf{p} - e\mathbf{A}) \\ c\boldsymbol{\sigma}\cdot(\mathbf{p} - e\mathbf{A}) & V - mc^2 \end{pmatrix}\begin{pmatrix}\psi^L \\ \psi^S\end{pmatrix} = W\begin{pmatrix}\psi^L \\ \psi^S\end{pmatrix}$$

After introducing E we get:

$$\begin{pmatrix} V & c\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A}) \\ c\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A}) & V-2mc^2 \end{pmatrix}\begin{pmatrix}\psi^L\\ \psi^S\end{pmatrix} = E\begin{pmatrix}\psi^L\\ \psi^S\end{pmatrix}$$

We put everything on the left hand side:

$$\begin{pmatrix} V-E & c\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A}) \\ c\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A}) & V-E-2mc^2 \end{pmatrix}\begin{pmatrix}\psi^L\\ \psi^S\end{pmatrix} = 0$$

We put c next to ψ^S:

$$\begin{pmatrix} V-E & \boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A}) \\ c\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A}) & \frac{V-E}{c}-2mc \end{pmatrix}\begin{pmatrix}\psi^L\\ c\psi^S\end{pmatrix} = 0$$

And we divide the second equation by c:

$$\begin{pmatrix} V-E & \boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A}) \\ \boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A}) & \frac{V-E}{c^2}-2m \end{pmatrix}\begin{pmatrix}\psi^L\\ c\psi^S\end{pmatrix} = 0$$

Now we express $c\psi^S$ from the second equation:

$$c\psi^S = \frac{\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A})\psi^L}{2m-\frac{V-E}{c^2}}$$

And substitute into the first equation:

$$\left(V-E+\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A})\frac{1}{2m-\frac{V-E}{c^2}}\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A})\right)\psi^L = 0$$

So we get the following equation (so far this is an exact equation for the first two components of the Dirac equation, no approximation has been made):

$$\left(\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A})\frac{1}{2m-\frac{V-E}{c^2}}\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A})+V\right)\psi^L = E\psi^L$$

Note that the first operator $\mathbf{p}$ (on the left hand side) acts among other things on the V in the denominator. By doing the nonrelativistic approximation $\frac{V-E}{c^2} \ll 2m$ we obtain the Pauli equation:

$$\left(\frac{\left(\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A})\right)^2}{2m}+V\right)\psi^L = E\psi^L \tag{8.4.1.1}$$

We can see, that the quantity

$$M = m-\frac{V-E}{2c^2}$$

can be interpreted as relativistic mass.

Using the relations between the Pauli matrices, we can further simplify:

$$\begin{aligned}
\left(\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A})\right)^2 &= (\mathbf{p}-e\mathbf{A})^2 + i\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A})\times(\mathbf{p}-e\mathbf{A}) = \\
&= (\mathbf{p}-e\mathbf{A})^2 + i\boldsymbol{\sigma}\cdot\left(\mathbf{p}\times\mathbf{p}-e\mathbf{A}\times\mathbf{p}-e\mathbf{p}\times\mathbf{A}+e^2\mathbf{A}\times\mathbf{A}\right) = \\
&= (\mathbf{p}-e\mathbf{A})^2 - ie\boldsymbol{\sigma}\cdot(\mathbf{A}\times\mathbf{p}+\mathbf{p}\times\mathbf{A}) = \\
&= (\mathbf{p}-e\mathbf{A})^2 - ie\boldsymbol{\sigma}\cdot(\mathbf{A}\times\mathbf{p}-\mathbf{A}\times\mathbf{p}-i\hbar(\nabla\times\mathbf{A})) = \\
&= (\mathbf{p}-e\mathbf{A})^2 - e\hbar\boldsymbol{\sigma}\cdot(\nabla\times\mathbf{A}) = \\
&= (\mathbf{p}-e\mathbf{A})^2 - e\hbar\boldsymbol{\sigma}\cdot\mathbf{B}
\end{aligned}$$

At the end, we have introduced the magnetic field $\mathbf{B} = \nabla \times \mathbf{A}$. In the above, one has to be careful, because $\mathbf{p}$ and $\mathbf{A}$ don't commute and also the operator $\mathbf{p}$ acts on everything on the right. We used the formula $\mathbf{p} \times \mathbf{A} = -\mathbf{A} \times \mathbf{p} - i\hbar(\nabla \times \mathbf{A})$, that can be proven by:

$$\begin{aligned}(\mathbf{p} \times \mathbf{A}\psi)_i = \epsilon_{ijk} p_j A_k \psi = \\ = -i\hbar\epsilon_{ijk}\partial_j(A_k\psi) = \\ = -i\hbar\epsilon_{ijk}((\partial_j A_k)\psi + A_k\partial_j\psi) = \\ = -i\hbar\epsilon_{ijk}((\partial_j A_k)\psi - A_j\partial_k\psi) = \\ = \epsilon_{ijk}(-i\hbar(\partial_j A_k)\psi - A_j p_k\psi) = \\ = -i\hbar((\nabla \times \mathbf{A})\psi)_i - (\mathbf{A} \times \mathbf{p}\psi)_i\end{aligned}$$

Putting this into the Pauli equation (8.4.1.1), we get:

$$\left(\frac{(\mathbf{p} - e\mathbf{A})^2}{2m} + V - \frac{e\hbar}{2m}\boldsymbol{\sigma} \cdot \mathbf{B}\right)\psi^L = E\psi^L \tag{8.4.1.2}$$

We can expand $(\mathbf{p} - e\mathbf{A})^2$ as follows:

$$\begin{aligned}(\mathbf{p} - e\mathbf{A})^2 = p^2 - e(\mathbf{p} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{p}) + e^2A^2 = \\ = p^2 - 2e\mathbf{A} \cdot \mathbf{p} + e^2A^2 + ie\hbar(\nabla \cdot \mathbf{A})\end{aligned}$$

where we used:

$$\mathbf{p} \cdot \mathbf{A}\psi = -i\hbar\partial_i A^i\psi = -i\hbar A^i\partial_i\psi - i\hbar(\partial_i A^i)\psi = \mathbf{A} \cdot \mathbf{p}\psi - i\hbar(\nabla \cdot \mathbf{A})\psi$$

and (8.4.1.2) becomes:

$$\left(\frac{p^2}{2m} - \frac{e}{m}\mathbf{A} \cdot \mathbf{p} + \frac{e^2}{2m}A^2 + \frac{ie\hbar}{2m}(\nabla \cdot \mathbf{A}) + V - \frac{e\hbar}{2m}\boldsymbol{\sigma} \cdot \mathbf{B}\right)\psi^L = E\psi^L$$

Using $\mathbf{p} = -i\hbar\nabla$ we get:

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + \frac{ie\hbar}{m}\mathbf{A} \cdot \nabla + \frac{e^2}{2m}A^2 + \frac{ie\hbar}{2m}(\nabla \cdot \mathbf{A}) + V - \frac{e\hbar}{2m}\boldsymbol{\sigma} \cdot \mathbf{B}\right)\psi^L = E\psi^L \tag{8.4.1.3}$$

Note that $V = e\phi$ and $e = -|e|$ is the (negative) electron charge. In the Coulomb gauge the term $\nabla \cdot \mathbf{A} = 0$. Sometimes one can neglect the quadratic term e^2A^2.

Example: velocity and length gauges

Let us assume a given spatially homogeneous time dependent electric field $\mathbf{E}(t)$ and no magnetic field $\mathbf{B} = 0$. Then one choice (gauge) of the electromagnetic potentials is:

$$\begin{aligned}\phi = 0 \\ \mathbf{A}(t) = -\int_0^t \mathbf{E}(t')\mathrm{d}t'\end{aligned}$$

The vector potential $\mathbf{A}$ is time dependent, but spatially homogeneous (constant). Let us first check that we are getting the correct $\mathbf{E}$ and $\mathbf{B}$:

$$\begin{aligned}\mathbf{E} = -\nabla\phi - \frac{\partial\mathbf{A}}{\partial t} = \mathbf{E}(t) \\ \mathbf{B} = \nabla \times \mathbf{A} = 0\end{aligned}$$

We also have $\nabla \cdot \mathbf{A} = 0$, since $\mathbf{A}$ does not depend on coordinates. Substituting into (8.4.1.3) we get:

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + \frac{ie\hbar}{m}\mathbf{A}\cdot\nabla + \frac{e^2}{2m}A^2\right)\psi = E\psi \tag{8.4.1.4}$$

The ∇ in the $\mathbf{A}\cdot\nabla$ term is a velocity, so this gauge is called a velocity gauge. We can apply a gauge transformation $\Lambda = -\mathbf{A}\cdot\mathbf{x}$ and we get:

$$\phi \to \phi - \frac{\partial\Lambda}{\partial t} = \phi + \frac{\partial\mathbf{A}}{\partial t}\cdot\mathbf{x} = \phi - \mathbf{E}\cdot\mathbf{x}$$
$$\mathbf{A} \to \mathbf{A} + \nabla\Lambda = \mathbf{A} - \mathbf{A} = 0$$

Substituting this into (8.4.1.4) (and using $V = e\phi$), we obtain:

$$\left(-\frac{\hbar^2}{2m}\nabla^2 - e\mathbf{E}\cdot\mathbf{x}\right)\psi = E\psi \tag{8.4.1.5}$$

Due to the $\mathbf{x}$ term in $\mathbf{E}\cdot\mathbf{x}$, this gauge is called the length gauge. Note that $e = -|e|$ is negative for electrons.

Nonrelativistic Limit in the Lagrangian

We use the identity $\frac{\partial}{\partial t}\left(e^{-imc^2t}f(t)\right) = e^{-imc^2t}(-imc^2 + \frac{\partial}{\partial t})f(t)$ to get:

$$L = c^2\partial^\mu\psi^*\partial_\mu\psi - m^2c^4\psi^*\psi = \frac{\partial}{\partial t}\psi^*\frac{\partial}{\partial t}\psi - c^2\partial^i\psi^*\partial_i\psi - m^2c^4\psi^*\psi =$$
$$= (imc^2 + \frac{\partial}{\partial t})\varphi^*(-imc^2 + \frac{\partial}{\partial t})\varphi - c^2\partial^i\varphi^*\partial_i\varphi - m^2c^4\varphi^*\varphi =$$
$$= 2mc^2\left[\frac{1}{2}i(\varphi^*\frac{\partial\varphi}{\partial t} - \varphi\frac{\partial\varphi^*}{\partial t}) - \frac{1}{2m}\partial^i\varphi^*\partial_i\varphi + \frac{1}{2mc^2}\frac{\partial\varphi^*}{\partial t}\frac{\partial\varphi}{\partial t}\right]$$

The constant factor $2mc^2$ in front of the Lagrangian is of course irrelevant, so we drop it and then we take the limit $c \to \infty$ (neglecting the last term) and we get

$$L = \frac{1}{2}i(\varphi^*\frac{\partial\varphi}{\partial t} - \varphi\frac{\partial\varphi^*}{\partial t}) - \frac{1}{2m}\partial^i\varphi^*\partial_i\varphi$$

After integration by parts we arrive at the Lagrangian for the Schrödinger equation:

$$L = i\varphi^*\frac{\partial\varphi}{\partial t} - \frac{1}{2m}\partial^i\varphi^*\partial_i\varphi$$

Klein-Gordon Equation

The Dirac equation implies the Klein-Gordon equation:

$$0 = (-i\hbar c\gamma^\mu D_\mu - mc^2)(i\hbar c\gamma^\nu D_\nu - mc^2)\psi = (\hbar^2c^2\gamma^\mu\gamma^\nu D_\mu D_\nu + m^2c^4)\psi =$$
$$= (\hbar^2c^2 g^{\mu\nu}D_\mu D_\nu + m^2c^4)\psi = (\hbar^2c^2 D^\mu D_\mu + m^2c^4)\psi$$

Note however, the ψ in the true Klein-Gordon equation is just a scalar, but here we get a 4-component spinor. Now:

$$D_\mu D_\nu = (\partial_\mu + ieA_\mu)(\partial_\nu + ieA_\nu) = \partial_\mu\partial_\nu + ie(A_\mu\partial_\nu + A_\nu\partial_\mu + (\partial_\mu A_\nu)) - e^2A_\mu A_\nu$$

$$[D_\mu, D_\nu] = D_\mu D_\nu - D_\nu D_\mu = ie(\partial_\mu A_\nu) - ie(\partial_\nu A_\mu)$$

We rewrite $D^\mu D_\mu$:

$$D^\mu D_\mu = g^{\mu\nu} D_\mu D_\nu = \partial^\mu \partial_\mu + ie((\partial^\mu A_\mu) + 2A^\mu \partial_\mu) - e^2 A^\mu A_\mu =$$

$$= \partial^\mu \partial_\mu + ie((\partial^0 A_0) + 2A^0 \partial_0 + (\partial^i A_i) + 2A^i \partial_i) - e^2 (A^0 A_0 + A^i A_i) =$$

$$= \partial^\mu \partial_\mu + i\frac{1}{c^2}\frac{\partial V}{\partial t} + 2i\frac{V}{c^2}\frac{\partial}{\partial t} + ie(\partial^i A_i) + 2ieA^i \partial_i - \frac{V^2}{c^2} - e^2 A^i A_i$$

The nonrelativistic limit can also be applied directly to the Klein-Gordon equation:

$$0 = (\hbar^2 c^2 D^\mu D_\mu + m^2 c^4)\psi =$$

$$= \left(\hbar^2 c^2 \partial^\mu \partial_\mu + i\frac{\partial V}{\partial t} + 2iV\frac{\partial}{\partial t} + i\hbar e c^2 (\partial^i A_i) + 2i\hbar e c^2 A^i \partial_i - V^2 - e^2 c^2 A^i A_i + m^2 c^4\right) e^{-\frac{i}{\hbar}mc^2 t}\varphi =$$

$$= \left(\hbar^2 \frac{\partial^2}{\partial t^2} - c^2\hbar^2\nabla^2 + 2iV\frac{\partial}{\partial t} + i\frac{\partial V}{\partial t} + i\hbar e c^2 (\partial^i A_i) + 2i\hbar e c^2 A^i \partial_i - V^2 - e^2 c^2 A^i A_i + m^2 c^4\right) e^{-\frac{i}{\hbar}mc^2 t}\varphi =$$

$$= e^{-\frac{i}{\hbar}mc^2 t}\left(\hbar^2(-\frac{i}{\hbar}mc^2 + \frac{\partial}{\partial t})^2 - \hbar^2 c^2\nabla^2 + 2iV(-\frac{i}{\hbar}mc^2 + \frac{\partial}{\partial t}) + i\frac{\partial V}{\partial t} + i\hbar e c^2(\partial^i A_i) + 2i\hbar e c^2 A^i\partial_i - V^2 + \right.$$

$$\left. - e^2 c^2 A^i A_i + m^2 c^4\right)\varphi =$$

$$= e^{-\frac{i}{\hbar}mc^2 t}\left(-2i\hbar mc^2\frac{\partial}{\partial t} + \hbar^2\frac{\partial^2}{\partial t^2} - c^2\hbar^2\nabla^2 + 2Vmc^2 + 2iV\frac{\partial}{\partial t} + i\frac{\partial V}{\partial t} + i\hbar e c^2(\partial^i A_i) + 2i\hbar e c^2 A^i\partial_i - V^2 + \right.$$

$$\left. - e^2 c^2 A^i A_i\right)\varphi =$$

$$= -2mc^2 e^{-\frac{i}{\hbar}mc^2 t}\left(i\hbar\frac{\partial}{\partial t} + \hbar^2\frac{\nabla^2}{2m} - V - \frac{1}{2mc^2}\frac{\partial^2}{\partial t^2} - \frac{i}{2mc^2}\frac{\partial V}{\partial t} + \frac{V^2}{2mc^2} - \frac{iV}{mc^2}\frac{\partial}{\partial t} + \right.$$

$$\left. - \frac{i\hbar e}{2m}\partial^i A_i - \frac{i\hbar e}{m}A^i\partial_i + \frac{e^2}{2m}A^i A_i\right)\varphi$$

Taking the limit $c \to \infty$ we again recover the Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}\varphi = \left(-\hbar^2\frac{\nabla^2}{2m} + V + \frac{i\hbar e}{2m}\partial^i A_i + \frac{i\hbar e}{m}A^i\partial_i - \frac{e^2}{2m}A^i A_i\right)\varphi\,,$$

we rewrite the right hand side a little bit:

$$i\hbar\frac{\partial}{\partial t}\varphi = \left(\frac{\hbar^2}{2m}(\partial^i\partial_i + \frac{i}{\hbar}e\partial^i A_i + 2\frac{i}{\hbar}eA^i\partial_i - \frac{e^2}{\hbar^2}A^i A_i) + V\right)\varphi\,,$$

$$i\hbar\frac{\partial}{\partial t}\varphi = \left(\frac{\hbar^2}{2m}(\partial^i + \frac{i}{\hbar}eA^i)(\partial_i + \frac{i}{\hbar}eA_i) + V\right)\varphi\,,$$

$$i\hbar\frac{\partial}{\partial t}\varphi = \left(\frac{1}{2m}\hbar^2 D^i D_i + V\right)\varphi\,,$$

Using (see the appendix for details):

$$\hbar^2 D^i D_i = -\hbar^2\delta_{ij}D^i D^j = -\hbar^2\left(\frac{i}{\hbar}(\mathbf{p} - e\mathbf{A})\right)^2 = (\mathbf{p} - e\mathbf{A})^2$$

we get the usual form of the Schrödinger equation for the vector potential:

$$i\hbar\frac{\partial}{\partial t}\varphi = \left(\frac{(\mathbf{p} - e\mathbf{A})^2}{2m} + V\right)\varphi\,.$$

A little easier derivation:

$$0 = (\hbar^2 c^2 D^\mu D_\nu + m^2 c^4)\psi =$$

$$= (\hbar^2 c^2 D^0 D_0 + \hbar^2 c^2 D^i D_i + m^2 c^4)\psi =$$

$$= 2mc^2 \left(\frac{\hbar^2}{2m} D^0 D_0 + \frac{\hbar^2}{2m} D^i D_i + \tfrac{1}{2} mc^2 \right) \psi =$$

$$= 2mc^2 \left(\frac{\hbar^2}{2m} \left(\partial^0 + \frac{i}{\hbar} e A^0 \right) \left(\partial_0 + \frac{i}{\hbar} e A_0 \right) + \tfrac{1}{2} mc^2 + \frac{\hbar^2}{2m} D^i D_i \right) e^{-\frac{i}{\hbar} mc^2 t} \varphi =$$

$$= 2mc^2 \left(\frac{\hbar^2}{2m} \left(\partial^0 + \frac{i}{\hbar} e A^0 \right) e^{-\frac{i}{\hbar} mc^2 t} \left(\partial_0 - \frac{i}{\hbar} mc + \frac{i}{\hbar} e A_0 \right) + \tfrac{1}{2} mc^2 + \frac{\hbar^2}{2m} D^i D_i \right) \varphi =$$

$$= 2mc^2 e^{-\frac{i}{\hbar} mc^2 t} \left(\frac{\hbar^2}{2m} \left(\partial^0 - \frac{i}{\hbar} mc + \frac{i}{\hbar} e A^0 \right) \left(\partial_0 - \frac{i}{\hbar} mc + \frac{i}{\hbar} e A_0 \right) + \tfrac{1}{2} mc^2 + \frac{\hbar^2}{2m} D^i D_i \right) \varphi =$$

$$= 2mc^2 e^{-\frac{i}{\hbar} mc^2 t} \left(\frac{\hbar^2}{2m} \partial^0 \partial_0 - \tfrac{1}{2} mc^2 - \frac{e^2 A^0 A_0}{2m} + c e A^0 + \frac{\hbar^2}{m} \frac{i}{\hbar} e (\partial^0 A^0 + A^0 \partial^0) - i\hbar c \partial_0 + \tfrac{1}{2} mc^2 + \frac{\hbar^2}{2m} D^i D_i \right) \varphi =$$

$$= 2mc^2 e^{-\frac{i}{\hbar} mc^2 t} \left(-i\hbar \frac{\partial}{\partial t} + \frac{\hbar^2}{2m} D^i D_i + c e A^0 + \frac{\hbar^2}{2mc^2} \frac{\partial^2}{\partial t^2} - \frac{e^2 \phi^2}{2mc^2} + \frac{ie\hbar}{mc^2} \left(\frac{\partial}{\partial t} \phi + \phi \frac{\partial}{\partial t} \right) \right) \varphi =$$

$$= 2mc^2 e^{-\frac{i}{\hbar} mc^2 t} \left(-i\hbar \frac{\partial}{\partial t} + \frac{(\mathbf{p} - e\mathbf{A})^2}{2m} + e\phi + \frac{\hbar^2}{2mc^2} \frac{\partial^2}{\partial t^2} - \frac{e^2 \phi^2}{2mc^2} + \frac{ie\hbar}{mc^2} \left(\frac{\partial}{\partial t} \phi + \phi \frac{\partial}{\partial t} \right) \right) \varphi$$

and letting $c \to \infty$ we get the Schrödinger equation:

$$i\hbar \frac{\partial}{\partial t} \varphi = \left(\frac{(\mathbf{p} - e\mathbf{A})^2}{2m} + e\phi \right) \varphi$$

8.4.2 Perturbation Theory

We want to solve the equation:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} |\psi(t)\rangle = H(t) |\psi(t)\rangle \tag{8.4.2.1}$$

with $H(t) = H^0 + H^1(t)$, where H^0 is time-independent part whose eigenvalue problem has been solved:

$$H^0 |n^0\rangle = E_n^0 |n^0\rangle$$

and $H^1(t)$ is a small time-dependent perturbation. $|n^0\rangle$ form a complete basis, so we can express $|\psi(t)\rangle$ in this basis:

$$|\psi(t)\rangle = \sum_n d_n(t) e^{-\frac{i}{\hbar} E_n^0 t} |n^0\rangle \tag{8.4.2.2}$$

Substituting this into (8.4.2.1), we get:

$$\sum_n \left(i\hbar \frac{\mathrm{d}}{\mathrm{d}t} d_n(t) + E_n^0 d_n(t) \right) e^{-\frac{i}{\hbar} E_n^0 t} |n^0\rangle = \sum_n \left(E_n^0 d_n(t) + H^1 d_n(t) \right) e^{-\frac{i}{\hbar} E_n^0 t} |n^0\rangle$$

so:

$$\sum_n i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \left(d_n(t) \right) e^{-\frac{i}{\hbar} E_n^0 t} |n^0\rangle = \sum_n d_n(t) e^{-\frac{i}{\hbar} E_n^0 t} H^1 |n^0\rangle$$

Choosing some particular state $|f^0\rangle$ of the H^0 Hamiltonian, we multiply the equation from the left by $\langle f^0| e^{\frac{i}{\hbar}E_f^0 t}$:

$$\sum_n i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\left(d_n(t)\right) e^{iw_{fn}t} \langle f^0|n^0\rangle = \sum_n d_n(t) e^{iw_{fn}t} \langle f^0|H^1|n^0\rangle$$

where $w_{fn} = \frac{E_f^0 - E_n^0}{\hbar}$. Using $\langle f^0|n^0\rangle = \delta_{fn}$:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} d_f(t) = \sum_n d_n(t) e^{iw_{fn}t} \langle f^0|H^1|n^0\rangle$$

we integrate from t_1 to t:

$$i\hbar\left(\left(d_f(t) - d_f(t_1)\right)\right) = \sum_n \int_{t_1}^{t} d_n(t') e^{iw_{fn}t'} \langle f^0|H^1(t')|n^0\rangle \,\mathrm{d}t'$$

Let the initial wavefunction at time t_1 be some particular state $|\psi(t_1)\rangle = |i^0\rangle$ of the unperturbed Hamiltonian, then $d_n(t_1) = \delta_{ni}$ and we get:

$$d_f(t) = \delta_{fi} - \frac{i}{\hbar} \sum_n \int_{t_1}^{t} d_n(t') e^{iw_{fn}t'} \langle f^0|H^1(t')|n^0\rangle \,\mathrm{d}t' \tag{8.4.2.3}$$

This is the equation that we will use for the perturbation theory.

In the zeroth order of the perturbation theory, we set $H^1(t) = 0$ and we get:

$$d_f(t) = \delta_{fi}$$

In the first order of the perturbation theory, we take the solution $d_n(t) = \delta_{ni}$ obtained in the zeroth order and substitute into the right hand side of (8.4.2.3):

$$d_f(t) = \delta_{fi} - \frac{i}{\hbar} \int_{t_1}^{t} e^{iw_{fi}t'} \langle f^0|H^1(t')|i^0\rangle \,\mathrm{d}t'$$

In the second order, we take the last solution, substitute into the right hand side of (8.4.2.3) again:

$$d_f(t) = \delta_{fi} + \left(-\frac{i}{\hbar}\right) \int_{t_1}^{t} e^{iw_{fi}t'} \langle f^0|H^1(t')|i^0\rangle \,\mathrm{d}t' +$$

$$+ \left(-\frac{i}{\hbar}\right)^2 \sum_n \int_{t_1}^{t} \mathrm{d}t'' \int_{t_1}^{t''} \mathrm{d}t' e^{iw_{fn}t''} \langle f^0|H^1(t'')|n^0\rangle \, e^{iw_{ni}t'} \langle n^0|H^1(t')|i^0\rangle$$

And so on for higher orders of the perturbation theory — more terms will arise on the right hand side of the last formula, so this is our main formula for calculating the $d_n(t)$ coefficients.

Time Independent Perturbation Theory

As a special case, if H^1 doesn't depend on time, the coefficients $d_n(t)$ simplify, so we calculate them in this section explicitly. Let's take

$$H(t) = H^0 + e^{t/\tau} H^1$$

so at the time $t_1 = -\infty$ the Hamiltonian $H(t) = H^0$ is unperturbed and we are interested in the time $t = 0$, when the Hamiltonian becomes $H(t) = H^0 + H^1$ (the coefficients $d_n(t)$ will still depend on the τ variable) and we do the limit $\tau \to \infty$ (this corresponds to smoothly applying the perturbation H^1 at the time negative infinity).

Let's calculate $d_f(0)$:

$$d_f(0) = \delta_{fi} + \left(-\frac{i}{\hbar}\right)\int_{-\infty}^{0} e^{iw_{fi}t'}e^{\frac{t}{\tau}}\mathrm{d}t'\,\langle f^0|H^1|i^0\rangle +$$

$$+\left(-\frac{i}{\hbar}\right)^2\sum_n\int_{-\infty}^{0}\mathrm{d}t''\int_{-\infty}^{t''}\mathrm{d}t' e^{iw_{fn}t''}e^{iw_{ni}t'}e^{\frac{t''}{\tau}}e^{\frac{t'}{\tau}}\,\langle f^0|H^1|n^0\rangle\,\langle n^0|H^1|i^0\rangle =$$

$$= \delta_{fi} + \left(-\frac{i}{\hbar}\right)\frac{1}{\frac{1}{\tau}+i\omega_{fi}}\,\langle f^0|H^1|i^0\rangle +$$

$$+\left(-\frac{i}{\hbar}\right)^2\sum_n\frac{1}{\frac{1}{\tau}+i\omega_{ni}}\frac{1}{\frac{2}{\tau}+i\omega_{fn}+i\omega_{ni}}\,\langle f^0|H^1|n^0\rangle\,\langle n^0|H^1|i^0\rangle$$

Taking the limit $\tau\to\infty$:

$$d_f(0) = \delta_{fi} + \left(-\frac{1}{\hbar}\right)\frac{1}{\omega_{fi}}\,\langle f^0|H^1|i^0\rangle +$$

$$+\left(-\frac{1}{\hbar}\right)^2\sum_n\frac{1}{\omega_{ni}}\frac{1}{\omega_{fn}+\omega_{ni}}\,\langle f^0|H^1|n^0\rangle\,\langle n^0|H^1|i^0\rangle =$$

$$= \delta_{fi} - \frac{\langle f^0|H^1|i^0\rangle}{E_f^0-E_i^0} +$$

$$+\sum_n\frac{\langle f^0|H^1|n^0\rangle\,\langle n^0|H^1|i^0\rangle}{(E_n^0-E_i^0)(E_f^0-E_i^0)}$$

Substituting this into (8.4.2.2) evaluated for $t=0$:

$$|\psi(0)\rangle = \sum_n d_n(0)\,|n^0\rangle =$$

$$= |i^0\rangle - \sum_n\frac{|n^0\rangle\,\langle n^0|H^1|i^0\rangle}{E_n^0-E_i^0} +$$

$$+\sum_{n,m}\frac{|n^0\rangle\,\langle n^0|H^1|m^0\rangle\,\langle m^0|H^1|i^0\rangle}{(E_m^0-E_i^0)(E_n^0-E_i^0)}$$

The sum $\sum_n$ is over all $n\neq i$, similarly for the other sum. Let's also calculate the energy:

$$E = \langle\psi(0)|H|\psi(0)\rangle = \langle\psi(0)|H^0+H^1|\psi(0)\rangle =$$

$$\left(\cdots - \sum_{n'\neq i}\frac{\langle i^0|H^1|n'^0\rangle\,\langle n'^0|}{E_{n'}^0-E_i^0} + \langle i^0|\right)(H^0+H^1)\left(|i^0\rangle - \sum_{n\neq i}\frac{|n^0\rangle\,\langle n^0|H^1|i^0\rangle}{E_n^0-E_i^0}+\cdots\right)$$

To evaluate this, we use the fact that $\langle i^0|H^0|i^0\rangle = E_i^0$ and $\langle i^0|H^0|n^0\rangle = E_i^0\delta_{ni}$:

$$E = E_i^0 + \langle i^0|H^1|i^0\rangle - \sum_{n\neq i}\frac{\langle i^0|H^1|n^0\rangle\,\langle n^0|H^1|i^0\rangle}{E_n^0-E_i^0}+\cdots =$$

$$= E_i^0 + \langle i^0|H^1|i^0\rangle - \sum_{n\neq i}\frac{|\,\langle n^0|H^1|i^0\rangle\,|^2}{E_n^0-E_i^0}+\cdots$$

Where we have neglected the higher order terms, so we can identify the corrections to the energy E coming from the particular orders of the perturbation theory:

$$E_i^0 = \langle i^0|H^0|i^0\rangle$$

$$E_i^1 = \langle i^0 | H^1 | i^0 \rangle$$
$$E_i^2 = -\sum_{n \neq i} \frac{|\langle n^0 | H^1 | i^0 \rangle|^2}{E_n^0 - E_i^0}$$

8.4.3 Scattering Theory

The incoming plane wave state is a solution of

$$H_0 |\mathbf{k}\rangle = E_k |\mathbf{k}\rangle$$

with $H_0 = \frac{p^2}{2m}$. E.g.

$$\langle \mathbf{r} | \mathbf{k} \rangle = e^{i\mathbf{r}\cdot\mathbf{k}}$$
$$E_k = \frac{\hbar^2 k^2}{2m}$$

We want to solve:

$$(H_0 + V) |\psi\rangle = E_k |\psi\rangle$$

The solution of this is:

$$|\psi\rangle = |\mathbf{k}\rangle + \frac{1}{E_k - H_0} V |\psi\rangle = |\mathbf{k}\rangle + GV |\psi\rangle$$

where

$$G = \frac{1}{E_k - H_0}$$

is the Green function for the Schrödinger equation. G is not unique, it contains both outgoing and ingoing waves. As shown below, one can distinguish between these two by adding a small $i\epsilon$ into the denominator, that moves the poles of the Green functions above and below the x-axis:

$$G_+ = \frac{1}{E_k - H_0 + i\epsilon}$$
$$G_- = \frac{1}{E_k - H_0 - i\epsilon}$$

Both G_+ and G_- are well-defined and unique. One can calculate both Green functions explicitly:

$$G_+(\mathbf{r}, \mathbf{r}') = \langle \mathbf{r} | G_+ | \mathbf{r}' \rangle = \langle \mathbf{r} | \frac{1}{E_k - H_0 + i\epsilon} | \mathbf{r}' \rangle =$$
$$= \int \frac{\mathrm{d}^3 k'}{(2\pi)^3} \frac{\langle \mathbf{r} | \mathbf{k}' \rangle \langle \mathbf{k}' | \mathbf{r}' \rangle}{E_k - E_{k'} + i\epsilon} = \int \frac{\mathrm{d}^3 k'}{(2\pi)^3} \frac{e^{i\mathbf{k}'\cdot(\mathbf{r}-\mathbf{r}')}}{E_k - E_{k'} + i\epsilon} = \frac{2m}{\hbar^2} \int \frac{\mathrm{d}^3 k'}{(2\pi)^3} \frac{e^{i\mathbf{k}'\cdot(\mathbf{r}-\mathbf{r}')}}{k^2 - k'^2 + i\epsilon} =$$
$$= \frac{4\pi m}{(2\pi)^3 \hbar^2 i |\mathbf{r} - \mathbf{r}'|} \int_{-\infty}^{\infty} \mathrm{d}^3 k' k' \frac{e^{ik'|\mathbf{r}-\mathbf{r}'|}}{k^2 - k'^2 + i\epsilon} = \frac{4\pi m}{(2\pi)^3 \hbar^2 i |\mathbf{r} - \mathbf{r}'|} (2\pi i) k \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{2k} =$$
$$= \frac{m e^{ik|\mathbf{r}-\mathbf{r}'|}}{2\pi\hbar^2 |\mathbf{r} - \mathbf{r}'|}$$

Similarly:

$$G_-(\mathbf{r}, \mathbf{r}') = \langle \mathbf{r} | G_- | \mathbf{r}' \rangle = \langle \mathbf{r} | \frac{1}{E_k - H_0 - i\epsilon} | \mathbf{r}' \rangle = \cdots = \frac{m e^{-ik|\mathbf{r}-\mathbf{r}'|}}{2\pi\hbar^2 |\mathbf{r} - \mathbf{r}'|}$$

Assuming $|\mathbf{r}'| \ll |\mathbf{r}|$, we can taylor expand $|\mathbf{r} - \mathbf{r}'|$:

$$|\mathbf{r} - \mathbf{r}'| = e^{-\mathbf{r}' \cdot \nabla} |\mathbf{r}| = \left(1 - \mathbf{r}' \cdot \nabla + (-\mathbf{r}' \cdot \nabla)^2 + O\left(r'^3\right)\right) |\mathbf{r}| = |\mathbf{r}| - \mathbf{r}' \cdot \nabla |\mathbf{r}| + O\left(r'^2\right) =$$
$$= r - \mathbf{r}' \cdot \hat{\mathbf{r}} + O\left(r'^2\right)$$

so:

$$e^{ik|\mathbf{r}-\mathbf{r}'|} \approx e^{ikr} e^{-ik\mathbf{r}' \cdot \hat{\mathbf{r}}}$$
$$|\mathbf{r} - \mathbf{r}'| \approx r$$

and simplify the result even further:

$$G_+(\mathbf{r}, \mathbf{r}') = \frac{m}{2\pi\hbar^2} \frac{e^{ikr}}{r} e^{-ik\mathbf{r}' \cdot \hat{\mathbf{r}}}$$
$$G_-(\mathbf{r}, \mathbf{r}') = \frac{m}{2\pi\hbar^2} \frac{e^{-ikr}}{r} e^{ik\mathbf{r}' \cdot \hat{\mathbf{r}}}$$

Let's get back to the solution of the Schrödinger equation:

$$|\psi\rangle = |\mathbf{k}\rangle + G_+ V |\psi\rangle$$

It contains the solution $|\psi\rangle$ on both sides of the equation, so we express it explicitly:

$$|\psi\rangle - G_+ V |\psi\rangle = |\mathbf{k}\rangle$$

$$|\psi\rangle = \frac{1}{1 - G_+ V} |\mathbf{k}\rangle$$

and multiply by V:

$$V |\psi\rangle = \frac{V}{1 - G_+ V} |\mathbf{k}\rangle = T |\mathbf{k}\rangle$$

where T is the transition matrix:

$$T = \frac{V}{1 - G_+ V} = V(1 + G_+ V + (G_+ V)^2 + \cdots) =$$

$$= V + V G_+ V + V G_+ V G_+ V + \cdots =$$

$$= V + V \frac{1}{E_k - H_0 + i\epsilon} V + V \frac{1}{E_k - H_0 + i\epsilon} V \frac{1}{E_k - H_0 + i\epsilon} V + \cdots$$

Then the final solution is:

$$|\psi\rangle = |\mathbf{k}\rangle + G_+ V |\psi\rangle = |\mathbf{k}\rangle + G_+ T |\mathbf{k}\rangle$$

and in a coordinate representation:

$$\psi(\mathbf{r}) = \langle \mathbf{r} | \psi \rangle = \langle \mathbf{r} | \mathbf{k} \rangle + \langle \mathbf{r} | G_+ T | \mathbf{k} \rangle = \langle \mathbf{r} | \mathbf{k} \rangle + \int \mathrm{d}^3 r' \, \langle \mathbf{r} | G_+ | \mathbf{r}' \rangle \, \langle \mathbf{r}' | T | \mathbf{k} \rangle =$$

$$= \langle \mathbf{r} | \mathbf{k} \rangle + \int \mathrm{d}^3 r' \mathrm{d}^3 k' \, \langle \mathbf{r} | G_+ | \mathbf{r}' \rangle \, \langle \mathbf{r}' | \mathbf{k}' \rangle \, \langle \mathbf{k}' | T | \mathbf{k} \rangle =$$

$$= e^{i\mathbf{k} \cdot \mathbf{r}} + \int \mathrm{d}^3 r' \mathrm{d}^3 k' G_+(\mathbf{r}, \mathbf{r}') e^{i\mathbf{k}' \cdot \mathbf{r}'} \, \langle \mathbf{k}' | T | \mathbf{k} \rangle$$

Plugging the representation of the Green function for $|\mathbf{r}'| \ll |\mathbf{r}|$ in:

$$\psi(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} + \frac{m}{2\pi\hbar^2}\frac{e^{ikr}}{r}\int \mathrm{d}^3r'\mathrm{d}^3k' e^{-ik\mathbf{r}'\cdot\hat{\mathbf{r}}} e^{i\mathbf{k}'\cdot\mathbf{r}'} \langle\mathbf{k}'|T|\mathbf{k}\rangle =$$
$$= e^{i\mathbf{k}\cdot\mathbf{r}} + \frac{m}{2\pi\hbar^2}\frac{e^{ikr}}{r}\int \mathrm{d}^3r'\mathrm{d}^3k' e^{i\mathbf{r}'\cdot(\mathbf{k}'-k\hat{\mathbf{r}})} \langle\mathbf{k}'|T|\mathbf{k}\rangle =$$
$$= e^{i\mathbf{k}\cdot\mathbf{r}} + \frac{m}{2\pi\hbar^2}\frac{e^{ikr}}{r}\int \mathrm{d}^3k' \delta(\mathbf{k}'-k\hat{\mathbf{r}}) \langle\mathbf{k}'|T|\mathbf{k}\rangle =$$
$$= e^{i\mathbf{k}\cdot\mathbf{r}} + \frac{m}{2\pi\hbar^2}\frac{e^{ikr}}{r} \langle k\hat{\mathbf{r}}|T|\mathbf{k}\rangle =$$
$$= e^{i\mathbf{k}\cdot\mathbf{r}} + f(\theta,\phi)\frac{e^{ikr}}{r}$$

where the scattering amplitude $f(\theta,\phi)$ is:

$$f(\theta,\phi) = \frac{m}{2\pi\hbar^2}\langle k\hat{\mathbf{r}}|T|\mathbf{k}\rangle = \frac{m}{2\pi\hbar^2}\langle\mathbf{k}'|T|\mathbf{k}\rangle$$

Where $\mathbf{k}' = k\hat{\mathbf{r}}$ is the final momentum.

The differential cross section $\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}$ is defined as the probability to observe the scattered particle in a given state per solid angle, e.g. the scattered flux per unit of solid angle per incident flux:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{|\mathbf{j}_i|}\frac{\mathrm{d}n}{\mathrm{d}\Omega} = \frac{r^2}{|\mathbf{j}_i|}\frac{\mathrm{d}n}{r^2\mathrm{d}\Omega} = \frac{r^2}{|\mathbf{j}_i|}\frac{\mathrm{d}n}{\mathrm{d}S} = \frac{r^2}{|\mathbf{j}_i|}\mathbf{j}_o\cdot\mathbf{n} = \frac{r^2}{|\mathbf{j}_i|}\mathbf{j}_o\cdot\hat{\mathbf{r}} =$$
$$= \frac{r^2}{\frac{\hbar k}{m}}\frac{\hbar k}{m}\left(\frac{1}{r^2}+\frac{i}{kr^3}\right)|f(\theta,\phi)|^2 = \left(1+\frac{i}{kr}\right)|f(\theta,\phi)|^2 \to |f(\theta,\phi)|^2$$

where we used $|\mathbf{j}_i| = \frac{\hbar k}{m}$ and

$$\mathbf{j}_o\cdot\hat{\mathbf{r}} = \frac{\hbar}{2mi}(\psi^*\nabla\psi - \psi\nabla\psi^*)\cdot\hat{\mathbf{r}} = \frac{\hbar}{2mi}\left(\psi^*\frac{\partial}{\partial r}\psi - \psi\frac{\partial}{\partial r}\psi^*\right) =$$
$$= \frac{\hbar}{2mi}\left(f^*(\theta,\phi)\frac{e^{-ikr}}{r}\frac{\partial}{\partial r}\left(f(\theta,\phi)\frac{e^{ikr}}{r}\right) - f(\theta,\phi)\frac{e^{ikr}}{r}\frac{\partial}{\partial r}\left(f^*(\theta,\phi)\frac{e^{-ikr}}{r}\right)\right) =$$
$$= \frac{\hbar k}{m}\left(\frac{1}{r^2}+\frac{i}{kr^3}\right)|f(\theta,\phi)|^2$$

Let's write the explicit formula for the transition matrix:

$$\langle\mathbf{k}'|T|\mathbf{k}\rangle = \int \mathrm{d}^3r\,\langle\mathbf{k}'|\mathbf{r}\rangle\langle\mathbf{r}|V|\mathbf{k}\rangle + \int \mathrm{d}^3r\mathrm{d}^3r'\,\langle\mathbf{k}'|\mathbf{r}\rangle\langle\mathbf{r}|VG_+|\mathbf{r}'\rangle\langle\mathbf{r}'|V|\mathbf{k}\rangle + \cdots =$$
$$= \int \mathrm{d}^3r e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}}V(\mathbf{r}) + \int \mathrm{d}^3r\mathrm{d}^3r' e^{-i\mathbf{k}'\cdot\mathbf{r}}V(\mathbf{r})\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}V(\mathbf{r}')e^{i\mathbf{k}\cdot\mathbf{r}'} + \cdots$$

Born Approximation

The Born approximation is just the first term:

$$\langle\mathbf{k}'|T|\mathbf{k}\rangle \approx \int \mathrm{d}^3r e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}}V(\mathbf{r}) = \int \mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\phi\, e^{iqr\cos\theta}V(r)r^2\sin\theta =$$
$$= 4\pi\int_0^\infty rV(r)\sin(qr)\,\mathrm{d}r$$

We can also write it as:

$$\langle \mathbf{k}'|T|\mathbf{k}\rangle \approx \int \mathrm{d}^3 r e^{-i\mathbf{q}\cdot\mathbf{r}} V(\mathbf{r}) = \tilde{V}(\mathbf{q})$$

where $\mathbf{q} = \mathbf{k}' - \mathbf{k}$. Note that for $|\mathbf{k}'| \approx |\mathbf{k}|$ we can write $|\mathbf{q}|$ using the angle θ between the vectors $\mathbf{k}'$ and $\mathbf{k}$:

$$\begin{aligned} |\mathbf{q}| = |\mathbf{k}' - \mathbf{k}| = \sqrt{k'^2 + k^2 - 2k'k\cos\theta} \approx \sqrt{k^2 + k^2 - 2k^2\cos\theta} = \\ = \sqrt{2k^2(1-\cos\theta)} = \sqrt{4k^2\sin^2\frac{\theta}{2}} = 2k\sin\left(\frac{\theta}{2}\right) \end{aligned}$$

Given the $\tilde{V}(\mathbf{q})$ we can then calculate the scattering potential $V(\mathbf{r})$ by the Fourier transform:

$$V(\mathbf{r}) = \int \frac{\mathrm{d}^3 q}{(2\pi)^3} \tilde{V}(\mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{r}}$$

Example 1:

$$\begin{aligned} \tilde{V}(\mathbf{q}) &= -\frac{g^2}{|\mathbf{q}|^2 + m_\phi^2} \\ V(\mathbf{r}) = \int \frac{\mathrm{d}^3 q}{(2\pi)^3} \frac{-g^2}{|\mathbf{q}|^2 + m_\phi^2} e^{i\mathbf{q}\cdot\mathbf{r}} &= \cdots = -\frac{g^2}{4\pi}\frac{1}{r} e^{-m_\phi r} \end{aligned}$$

Example 2:

$$\begin{aligned} \tilde{V}(\mathbf{q}) &= \frac{e^2}{|\mathbf{q}|^2} \\ V(\mathbf{r}) = \int \frac{\mathrm{d}^3 q}{(2\pi)^3} \frac{e^2}{|\mathbf{q}|^2} e^{i\mathbf{q}\cdot\mathbf{r}} &= \cdots = \frac{e^2}{4\pi r} \end{aligned}$$

Example 3 — Yukawa potential in Born approximation:

$$V(r) = -V_0 \frac{e^{-\alpha r}}{r}$$

$$\tilde{V}(\mathbf{q}) = -\frac{4\pi V_0}{|\mathbf{q}|^2 + \alpha^2}$$

$$f(\theta, \phi) = \frac{m}{2\pi\hbar^2} \langle \mathbf{k}' | T | \mathbf{k} \rangle = \frac{m}{2\pi\hbar^2} \tilde{V}(\mathbf{q}) = -\frac{m}{2\pi\hbar^2} \frac{4\pi V_0}{|\mathbf{q}|^2 + \alpha^2} = -\frac{2m}{\hbar^2} \frac{V_0}{|\mathbf{q}|^2 + \alpha^2}$$

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = |f(\theta, \phi)|^2 = \left(\frac{2mV_0}{\hbar^2}\right)^2 \frac{1}{\left(|\mathbf{q}|^2 + \alpha^2\right)^2} = \left(\frac{2mV_0}{\hbar^2}\right)^2 \frac{1}{\left(4k^2 \sin^2\left(\frac{\theta}{2}\right) + \alpha^2\right)^2}$$

$$\sigma = \int \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \mathrm{d}\Omega = \int \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \sin\theta \mathrm{d}\theta \mathrm{d}\phi =$$

$$= \left(\frac{2mV_0}{\hbar^2}\right)^2 \int \frac{1}{\left(4k^2 \sin^2\left(\frac{\theta}{2}\right) + \alpha^2\right)^2} \sin\theta \mathrm{d}\theta \mathrm{d}\phi =$$

$$= \left(\frac{2mV_0}{\hbar^2}\right)^2 2\pi \int_0^\pi \frac{\sin\theta \mathrm{d}\theta}{\left(4k^2 \sin^2\left(\frac{\theta}{2}\right) + \alpha^2\right)^2} =$$

$$= \left(\frac{2mV_0}{\hbar^2}\right)^2 2\pi \int_0^\pi \frac{\sin\theta \mathrm{d}\theta}{\left(2k^2(1 - \cos\theta) + \alpha^2\right)^2} =$$

$$= \left(\frac{2mV_0}{\hbar^2}\right)^2 2\pi \int_{-1}^1 \frac{\mathrm{d}y}{\left(2k^2(1 + y) + \alpha^2\right)^2} =$$

$$= \left(\frac{2mV_0}{\hbar^2}\right)^2 2\pi \int_{\alpha^2}^{4k^2 + \alpha^2} \frac{2k^2 \mathrm{d}z}{z^2} =$$

$$= \left(\frac{2mV_0}{\hbar^2}\right)^2 2\pi 2k^2 \left(\frac{1}{\alpha^2} - \frac{1}{4k^2 + \alpha^2}\right)$$

Example 4 — Coulomb potential in Born approximation:

$$\alpha \to 0$$

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left(\frac{2mV_0}{\hbar^2}\right)^2 \frac{1}{\left(4k^2 \sin^2\left(\frac{\theta}{2}\right)\right)^2} = \left(\frac{2mV_0}{4\hbar^2 k^2}\right)^2 \frac{1}{\sin^4\frac{\theta}{2}}$$

$$E = \frac{p^2}{2m} = \frac{\hbar^2 k^2}{2m}$$

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left(\frac{V_0}{4E}\right)^2 \frac{1}{\sin^4\frac{\theta}{2}}$$

$$V_0 \to \frac{ZZ'e^2}{4\pi\epsilon_0} = ZZ'\alpha\hbar c$$

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left(\frac{ZZ'\alpha\hbar c}{4E}\right)^2 \frac{1}{\sin^4\frac{\theta}{2}}$$

By setting $E = \frac{1}{2}mv_0^2$ we obtain the classical Rutherford cross-section formula.

8.5 Systematic Perturbation Theory in QM

We have

$$H = H_0 + e^{-\epsilon|t|} H_1$$

where the ground state of the noninteracting Hamiltonian H_0 is:

$$H_0 |0\rangle = E_0 |0\rangle$$

and the ground state of the interacting Hamiltonian H is:

$$H |\Omega\rangle = E |\Omega\rangle$$

Then:

$$\begin{aligned} H |\Omega\rangle = (H_0 + H_1) |\Omega\rangle = E |\Omega\rangle \\ \langle 0|H_0 + H_1|\Omega\rangle = E \langle 0|\Omega\rangle \\ E_0 \langle 0|\Omega\rangle + \langle 0|H_1|\Omega\rangle = E \langle 0|\Omega\rangle \\ E = E_0 + \frac{\langle 0|H_1|\Omega\rangle}{\langle 0|\Omega\rangle} \end{aligned}$$

We can also write

$$|\Omega\rangle = \lim_{\epsilon \to 0+} U_\epsilon(0, -\infty) |0\rangle$$

where

$$U_\epsilon(t, t_0) = T \exp\left(-\frac{i}{\hbar} \int_{t_0}^{t} \mathrm{d}t' e^{-\epsilon|t'|} H_1(t')\right)$$

Let's write several common expressions for the ground state energy:

$$\begin{aligned} \Delta E = E - E_0 = \frac{\langle 0|H_1|\Omega\rangle}{\langle 0|\Omega\rangle} = \frac{\langle 0|H_1 U(0,-\infty)|0\rangle}{\langle 0|U(0,-\infty)|0\rangle} = \\ = \lim_{t\to 0} \frac{\langle 0|H_1 U(t,-\infty)|0\rangle}{\langle 0|U(t,-\infty)|0\rangle} = \lim_{t\to 0} \frac{\langle 0|i\partial_t U(t,-\infty)|0\rangle}{\langle 0|U(t,-\infty)|0\rangle} = \lim_{t\to 0} \frac{i\partial_t \langle 0|U(t,-\infty)|0\rangle}{\langle 0|U(t,-\infty)|0\rangle} = \\ = \lim_{t\to 0} i\partial_t \log \langle 0|U(t,-\infty)|0\rangle \equiv \lim_{t\to\infty(1-i\epsilon)} i\frac{\mathrm{d}}{\mathrm{d}t} \log \langle 0|U(t,-\infty)|0\rangle \end{aligned}$$

The last expression incorporates the ϵ dependence of U_ϵ explicitly. The vacuum amplitude is sometimes denoted by $R(t)$:

$$R(t) = \langle 0|U(t,-\infty)|0\rangle$$

The two point (interacting) Green (or correlation) function is:

$$G(x,y) = \langle \Omega|T\phi(x)\phi(y)|\Omega\rangle = \frac{\langle 0|T\phi(x)\phi(y)U(\infty,-\infty)|0\rangle}{\langle 0|U(\infty,-\infty)|0\rangle}$$

The $\epsilon \to 0$ limit of U_ϵ is tacitly assumed to make this formula well defined (sometimes the other way $t \to \infty(1-i\epsilon)$ of writing the same limit is used). Another way of writing the formula above for the Green function in QM is:

$$G(\mathbf{k}_1, \mathbf{k}_2, t_2 - t_1) = i \langle \Omega|T c_{\mathbf{k}_2}(t_2) c^\dagger_{\mathbf{k}_1}(t_1)|\Omega\rangle = i\frac{\langle 0|T c_{\mathbf{k}_2}(t_2) c^\dagger_{\mathbf{k}_1}(t_1) U(\infty,-\infty)|0\rangle}{\langle 0|U(\infty,-\infty)|0\rangle}$$

Last type of similar expressions to consider is the scattering amplitude:

$$\langle f|U(\infty,-\infty)|i\rangle$$

where the initial state is let's say a boson+fermion and the final state a boson+antifermion:

$$\begin{aligned} |i\rangle = a^\dagger_{\mathbf{k}} b^{s\dagger}_{\mathbf{l}} |0\rangle \\ |f\rangle = a^\dagger_{\mathbf{p}} a^{r\dagger}_{\mathbf{q}} |0\rangle \end{aligned}$$

This is just an example, the $|i\rangle$ and $|f\rangle$ states can contain any number of (arbitrary) particles.

8.6 Appendix

8.6.1 Units and Dimensional Analysis

The evolution operator is dimensionless:

$$U(-\infty,\infty) = T\exp\left(\frac{i}{\hbar}\int_{-\infty}^{\infty}\mathrm{d}^4x\mathcal{L}(x)\right)$$

So:

$$\left[\int_{-\infty}^{\infty}\mathrm{d}^4x\mathcal{L}(x)\right] = [\hbar] = M^0$$

where M is an arbitrary mass scale. Length unit is M^{-1}, so then

$$[\mathcal{L}(x)] = M^4$$

For the particular forms of the Lagrangians above we get:

$$[m\bar{e}e] = [m^2 Z_\mu Z^\mu] = [m^2 H^2] = [i\bar{e}\gamma^\mu\partial_\mu e] = [\mathcal{L}] = M^4$$

so $[\bar{e}e] = M^3$, $[Z_\mu Z^\mu] = [H^2] = M^2$ and we get

$$[e] = [\bar{e}] = M^{\frac{3}{2}}$$

$$[Z_\mu] = [Z^\mu] = [H] = [\partial_\mu] = [\partial^\mu] = M^1$$

Example: what is the dimension of G_μ in $\mathcal{L} = -\frac{G_\mu}{\sqrt{2}}[\bar{\psi}_{\nu_\mu}\gamma^\mu(1-\gamma_5)\psi_\mu][\bar{\psi}_e\gamma^\mu(1-\gamma_5)\psi_{\nu_e}]$? Answer:

$$[\mathcal{L}] = [G_\mu\bar{\psi}\psi\bar{\psi}\psi]$$

$$M^4 = [G_\mu]M^{\frac{3}{2}}M^{\frac{3}{2}}M^{\frac{3}{2}}M^{\frac{3}{2}}$$

$$[G_\mu] = M^{-2}$$

In order to get the above units from the SI units, one has to do the following identification:

$$kg \to M^1$$

$$m \to M^{-1}$$

$$s \to M^{-1}$$

$$A \to M^1$$

The SI units of the above quantities are:

$$
\begin{aligned}
= \mathrm{V} = \frac{\mathrm{kg\,m^2}}{\mathrm{A\,s^3}} &= \mathrm{M} \\
[A_\mu] = \frac{[\phi]}{[c]} = \frac{\mathrm{V\,s}}{\mathrm{m}} = \frac{\mathrm{kg\,m}}{\mathrm{A\,s^2}} &= \mathrm{M} \\
[c] = \frac{\mathrm{m}}{\mathrm{s}} &= 1 \\
[e] = \mathrm{C} = \mathrm{A\,s} &= 1 \\
[\hbar] = \mathrm{J\,s} = \frac{\mathrm{m^2\,kg}}{\mathrm{s}} &= 1 \\
[\partial_\mu] = \frac{1}{\mathrm{m}} &= \mathrm{M} \\
[F_{\mu\nu}] = [\partial_\mu A_\nu] = \frac{\mathrm{kg}}{\mathrm{A\,s^2}} &= \mathrm{M^2} \\
[\mathcal{L}] = [F_{\mu\nu}]^2 = \frac{\mathrm{kg^2}}{\mathrm{A^2\,s^4}} &= \mathrm{M^4} \\
[\psi] = \frac{\mathrm{kg}^{\frac{1}{2}}}{\mathrm{A\,m\,s}} &= \mathrm{M}^{\frac{3}{2}}
\end{aligned}
$$

The SI units are useful for checking that the c, e and $\hbar$ constants are at correct places in the expression.

8.6.2 Atomic Units

Hartree atomic units are defined using the relations:

$$\hbar = m = e = 4\pi\epsilon_0 = 1$$

so for example for the Bohr radius we get:

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{me^2} = 1$$

for fine structure constant ($\alpha = 1/137.036...$) we get:

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} = \frac{1}{c}$$

from which we calculate the speed of light c in atomic units as:

$$c = \frac{1}{\alpha}$$

Energy is measured in Hartrees, one Hartree being

$$1\,\mathrm{Ha} = \frac{\hbar^2}{ma_0^2} = 1\,(\mathrm{a.u.}) = 27.211\,\mathrm{eV}$$

Hamiltonian and the corresponding spectrum of the Hydrogen atom:

$$
\begin{aligned}
H &= -\frac{\hbar^2}{2m}\nabla^2 - \frac{1}{4\pi\epsilon_0}\frac{e^2}{r} \\
E_n &= -\frac{\hbar^2}{ma_0^2}\frac{1}{2n^2}
\end{aligned}
$$

become in atomic units:

$$H = -\frac{1}{2}\nabla^2 - \frac{1}{r}$$
$$E_n = -\frac{1}{2n^2}$$

Poisson equation (Gauss's law)

$$\nabla^2\phi = -\frac{\rho}{\epsilon_0}$$

becomes:

$$\nabla^2\phi = -4\pi\rho$$

8.6.3 Tensors in Special Relativity and QFT

In general, the covariant and contravariant vectors and tensors work just like in special (and general) relativity. We use the metric $g_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$ (e.g. signature -2, but it's possible to also use the metric with signature +2). The four potential A^μ is given by:

$$A^\mu = \left(\frac{\phi}{c}, \mathbf{A}\right) = (A^0, A^1, A^2, A^3)$$

where ϕ is the electrostatic potential. Whenever we write $\mathbf{A}$, the components of it are given by the upper indices, e.g. $\mathbf{A} = (A^1, A^2, A^3)$. The components with lower indices can be calculated using the metric tensor, so it depends on the signature convention:

$$A_\mu = g_{\mu\nu}A^\nu = (A^0, -\mathbf{A}) = (A^0, -A^1, -A^2, -A^3)$$

In our case we got $A_0 = A^0$ and $A_i = -A^i$ (if we used the other signature convention, then the sign of A_0 would differ and A_i would stay the same). The length (squared) of the vector is:

$$A^2 = A_\mu A^\mu = \left(A^0\right)^2 - |\mathbf{A}|^2 = \left(A^0\right)^2 - \mathbf{A}^2$$

where $\mathbf{A}^2 \equiv |\mathbf{A}|^2 = (A^1)^2 + (A^2)^2 + (A^3)^2$.

The position 4-vector is (in any metric):

$$x^\mu = (ct, \mathbf{x})$$

Gradient is defined as (in any metric):

$$\partial_\mu = (\partial_0, \partial_1, \partial_2, \partial_3) = \frac{\partial}{\partial x^\mu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$$

the upper indices depend on the signature, e.g. for -2:

$$\partial^\mu = (\partial^0, \partial^1, \partial^2, \partial^3) = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z}\right)$$

and +2:

$$\partial^\mu = (\partial^0, \partial^1, \partial^2, \partial^3) = \left(-\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$$

The d'Alembert operator is:

$$\partial^2 \equiv \partial_\mu\partial^\mu$$

the 4-velocity is (in any metric):

$$v^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}t}{\mathrm{d}\tau}\frac{\mathrm{d}x^\mu}{\mathrm{d}t} = \gamma(c, \mathbf{v})$$

where τ is the proper time, $\gamma = \frac{\mathrm{d}t}{\mathrm{d}\tau} = \frac{1}{\sqrt{1-\frac{\mathbf{v}^2}{c^2}}}$ and $\mathbf{v} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}$ is the velocity in the coordinate time t. In the metric with signature +2:

$$v^2 = v_\mu v^\mu = g_{\mu\nu} v^\mu v^\nu = -\gamma^2 c^2 + \gamma^2 \mathbf{v}^2 = \frac{-c^2 + \mathbf{v}^2}{1 - \frac{\mathbf{v}^2}{c^2}} = -c^2$$

With signature -2 we get $v^2 = c^2$. The 4-momentum is (in any metric)

$$p^\mu = m v^\mu = m\gamma(c, \mathbf{v})$$

where m is the rest mass. The fluid-density 4-current is (in any metric):

$$j^\mu = \rho v^\mu = \rho\gamma(c, \mathbf{v})$$

where ρ is the fluid density at rest. For example the vanishing 4-divergence (the continuity equation) is written as (in any metric):

$$0 = \partial_\mu j^\mu = \frac{1}{c}\frac{\partial}{\partial t}(\rho\gamma c) + \nabla\cdot(\rho\gamma\mathbf{v}) = \frac{\partial}{\partial t}(\rho\gamma) + \nabla\cdot(\rho\mathbf{v}\gamma) = \frac{\partial}{\partial t}\left(\frac{\rho}{\sqrt{1-\frac{\mathbf{v}^2}{c^2}}}\right) + \nabla\cdot\left(\frac{\rho\mathbf{v}}{\sqrt{1-\frac{\mathbf{v}^2}{c^2}}}\right)$$

Momentum ($\mathbf{p} = -i\hbar\nabla$) and energy ($E = i\hbar\frac{\partial}{\partial t}$) is combined into 4-momentum as

$$p^\mu = \left(\frac{E}{c}, \mathbf{p}\right) = i\hbar\left(\frac{1}{c}\frac{\partial}{\partial t}, -\nabla\right) = i\hbar\left(\partial_0, -\partial_j\right) = i\hbar\left(\partial^0, \partial^j\right) = i\hbar\partial^\mu$$
$$p_\mu = g_{\mu\nu}p^\nu = i\hbar g_{\mu\nu}\partial^\nu = i\hbar\partial_\mu$$

For the signature $+2$ we get $p^\mu = -i\hbar\partial^\mu$ and $p_\mu = -i\hbar\partial_\mu$.

For p^2 we get (signature -2):

$$p^2 = p_\mu p^\mu = (p^0)^2 - \mathbf{p}^2 = (p_0)^2 - \mathbf{p}^2 = \frac{E^2}{c^2} - \mathbf{p}^2$$
$$p^2 = p_\mu p^\mu = m^2 v_\mu v^\mu = m^2 c^2$$

comparing those two we get the following useful relations (valid in any metric):

$$\frac{E^2}{c^2} - \mathbf{p}^2 = m^2c^2$$
$$E^2 = m^2c^4 + \mathbf{p}^2c^2$$
$$E = \sqrt{m^2c^4 + \mathbf{p}^2c^2} = mc^2\sqrt{1 + \frac{\mathbf{p}^2}{m^2c^2}} = mc^2\left(1 + \frac{\mathbf{p}^2}{2m^2c^2} + O\left(\frac{p^4}{m^4c^4}\right)\right) =$$
$$= mc^2 + \frac{\mathbf{p}^2}{2m} + O\left(\frac{p^4}{m^3c^2}\right)$$

the following relations are also useful:

$$p^2 = p_\mu p^\mu = -\hbar^2\partial_\mu\partial^\mu \equiv -\hbar^2\partial^2 = -\hbar^2\left(\partial_0\partial^0 + \partial_i\partial^i\right) = -\hbar^2\left(\partial_0\partial_0 - \partial_i\partial_i\right) =$$
$$= -\hbar^2\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right) = -\frac{\hbar^2}{c^2}\frac{\partial^2}{\partial t^2} + \hbar^2\nabla^2$$

For the signature $+2$ we get:

$$p^2 = p_\mu p^\mu = -\hbar^2 \partial_\mu \partial^\mu \equiv -\hbar^2 \partial^2 = -\hbar^2 \left(\partial_0 \partial^0 + \partial_i \partial^i\right) = -\hbar^2 \left(-\partial_0 \partial_0 + \partial_i \partial_i\right) = \\ = -\hbar^2 \left(-\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \nabla^2\right) = \frac{\hbar^2}{c^2}\frac{\partial^2}{\partial t^2} - \hbar^2 \nabla^2$$

So for example the Klein-Gordon equation:

$$\left(\frac{\hbar^2}{c^2}\frac{\partial^2}{\partial t^2} - \hbar^2 \nabla^2 + m^2 c^2\right) \psi = 0$$

can be for signature -2 written as:

$$(+\hbar^2 \partial^2 + m^2 c^2)\psi = (-p^2 + m^2 c^2)\psi = 0$$

and for $+2$ as:

$$(-\hbar^2 \partial^2 + m^2 c^2)\psi = (p^2 + m^2 c^2)\psi = 0$$

Note: for the signature +2, we would get $p^\mu = -i\hbar\partial^\mu$ and $p_\mu = -i\hbar\partial_\mu$.

For the minimal coupling $D_\mu = \partial_\mu + \frac{i}{\hbar} e A_\mu$ we get:

$$D^0 = \partial^0 + \frac{i}{\hbar} e A^0$$

$$D^j = \partial^j + \frac{i}{\hbar} e A^j = -\frac{i}{\hbar}(i\hbar\partial^j - e A^j) = -\frac{i}{\hbar}(\mathbf{p} - e\mathbf{A})$$

and for the lower indices:

$$D_0 = \partial_0 + \frac{i}{\hbar} e A_0$$

$$D_j = \partial_j + \frac{i}{\hbar} e A_j = -\frac{i}{\hbar}(i\hbar\partial_j - e A_j) = \frac{i}{\hbar}(i\hbar\partial^j - e A^j) = \frac{i}{\hbar}(\mathbf{p} - e\mathbf{A})$$

8.6.4 Adding Angular Momenta

Angular momenta are added using the Clebsch-Gordan coefficients (or equivalently $3j$ symbols):

$$|j_1 j_2 j_3 m_3\rangle = \sum_{m_1 m_2} (j_1 m_1 j_2 m_2 | j_3 m_3) \, |j_1 m_1 j_2 m_2\rangle = \\ = \sum_{m_1 m_2} (-1)^{j_1 - j_2 + m_3} \sqrt{2j_3 + 1} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{pmatrix} |j_1 m_1 j_2 m_2\rangle \tag{8.6.4.1}$$

Spin Orbit Coupling (Spin Spherical Harmonics)

This is just a special case of (8.6.4.1) for:

$$j_1 = l$$
$$m_1 = m$$
$$j_2 = \tfrac{1}{2}$$
$$m_2 = s$$

So the kets $|j_1 m_1 j_2 m_2\rangle$ can be written as:

$$|j_1 m_1 j_2 m_2\rangle = |l m \tfrac{1}{2} s\rangle = Y_{lm} \Phi_s$$

Where:

$$\Phi_{\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$
$$\Phi_{-\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Where s is a spin, $s = \pm\frac{1}{2}$. Then we get:

$$\begin{aligned}
|l \tfrac{1}{2} j_3 m_3\rangle &= \sum_{m=-l}^{l} \sum_{s=-\frac{1}{2}}^{\frac{1}{2}} (-1)^{l-\frac{1}{2}+m_3} \sqrt{2j_3+1} \begin{pmatrix} l & \frac{1}{2} & j_3 \\ m & s & -m_3 \end{pmatrix} |l m \tfrac{1}{2} s\rangle = \\
&= (-1)^{l-\frac{1}{2}+m_3} \sqrt{2j_3+1} \sum_{m=-l}^{l} \left(\begin{pmatrix} l & \frac{1}{2} & j_3 \\ m & -\frac{1}{2} & -m_3 \end{pmatrix} |l m \tfrac{1}{2} (-\tfrac{1}{2})\rangle + \right. \\
&\qquad \left. + \begin{pmatrix} l & \frac{1}{2} & j_3 \\ m & \frac{1}{2} & -m_3 \end{pmatrix} |l m \tfrac{1}{2} \tfrac{1}{2}\rangle \right) = \\
&= (-1)^{l-\frac{1}{2}+m_3} \sqrt{2j_3+1} \left(\begin{pmatrix} l & \frac{1}{2} & j_3 \\ m_3+\frac{1}{2} & -\frac{1}{2} & -m_3 \end{pmatrix} |l (m_3+\tfrac{1}{2}) \tfrac{1}{2} (-\tfrac{1}{2})\rangle + \right. \\
&\qquad \left. + \begin{pmatrix} l & \frac{1}{2} & j_3 \\ m_3-\frac{1}{2} & \frac{1}{2} & -m_3 \end{pmatrix} |l (m_3-\tfrac{1}{2}) \tfrac{1}{2} \tfrac{1}{2}\rangle \right) = \\
&= (-1)^{l-\frac{1}{2}+m_3} \sqrt{2j_3+1} \begin{pmatrix} \begin{pmatrix} l & \frac{1}{2} & j_3 \\ m_3-\frac{1}{2} & \frac{1}{2} & -m_3 \end{pmatrix} Y_{l,m_3-\frac{1}{2}} \\ \begin{pmatrix} l & \frac{1}{2} & j_3 \\ m_3+\frac{1}{2} & -\frac{1}{2} & -m_3 \end{pmatrix} Y_{l,m_3+\frac{1}{2}} \end{pmatrix}
\end{aligned}$$

These are called spin-angular functions or spin spherical harmonics. Using the triangle selection rule of the $3j$ symbols, we can see that there are only two options for j_3:

$$j_3 = l + \tfrac{1}{2}$$
$$j_3 = l - \tfrac{1}{2}$$

So we get for $j_3 = l + \frac{1}{2}$:

$$|(j_3 - \tfrac{1}{2})\tfrac{1}{2} j_3 m_3\rangle = (-1)^{j_3-\frac{1}{2}-\frac{1}{2}+m_3}\sqrt{2j_3+1}\begin{pmatrix}\begin{pmatrix} j_3-\frac{1}{2} & \frac{1}{2} & j_3 \\ m_3-\frac{1}{2} & \frac{1}{2} & -m_3\end{pmatrix} Y_{j_3-\frac{1}{2},m_3-\frac{1}{2}} \\ \begin{pmatrix} j_3-\frac{1}{2} & \frac{1}{2} & j_3 \\ m_3+\frac{1}{2} & -\frac{1}{2} & -m_3\end{pmatrix} Y_{j_3-\frac{1}{2},m_3+\frac{1}{2}}\end{pmatrix} =$$

$$= (-1)^{j_3+m_3-1}\sqrt{2j_3+1}\begin{pmatrix}(-1)^{j_3+m_3-1}\sqrt{\frac{j_3+m_3}{2j_3(2j_3+1)}}Y_{j_3-\frac{1}{2},m_3-\frac{1}{2}} \\ (-1)^{2j_3}(-1)^{j_3-m_3-1}\sqrt{\frac{j_3-m_3}{2j_3(2j_3+1)}}Y_{j_3-\frac{1}{2},m_3+\frac{1}{2}}\end{pmatrix} =$$

$$= \begin{pmatrix}\sqrt{\frac{j_3+m_3}{2j_3}}Y_{j_3-\frac{1}{2},m_3-\frac{1}{2}} \\ (-1)^{4j_3}\sqrt{\frac{j_3-m_3}{2j_3}}Y_{j_3-\frac{1}{2},m_3+\frac{1}{2}}\end{pmatrix} =$$

$$= \frac{1}{\sqrt{2j_3}}\begin{pmatrix}\sqrt{j_3+m_3}\,Y_{j_3-\frac{1}{2},m_3-\frac{1}{2}} \\ \sqrt{j_3-m_3}\,Y_{j_3-\frac{1}{2},m_3+\frac{1}{2}}\end{pmatrix} =$$

$$= \frac{1}{\sqrt{2l+1}}\begin{pmatrix}\sqrt{l+m_3+\frac{1}{2}}\,Y_{l,m_3-\frac{1}{2}} \\ \sqrt{l-m_3+\frac{1}{2}}\,Y_{l,m_3+\frac{1}{2}}\end{pmatrix}$$

The allowed values for m_3 are $m_3 = -l-\frac{1}{2}, -l+\frac{1}{2}, -l+\frac{1}{2}+1, \ldots, l-\frac{1}{2}, l+\frac{1}{2}$, total of $2l+2$ values. For the case $l = \pm(l+\frac{1}{2})$, the spherical harmonic is not defined ($m > l$) but its coefficient (the square root $\sqrt{l \pm m_3 + \frac{1}{2}}$) is zero, so the whole element is defined as zero.

For $j_3 = l - \frac{1}{2}$:

$$|(j_3 + \tfrac{1}{2})\tfrac{1}{2} j_3 m_3\rangle = (-1)^{j_3+\frac{1}{2}-\frac{1}{2}+m_3}\sqrt{2j_3+1}\begin{pmatrix}\begin{pmatrix} j_3+\frac{1}{2} & \frac{1}{2} & j_3 \\ m_3-\frac{1}{2} & \frac{1}{2} & -m_3\end{pmatrix} Y_{j_3+\frac{1}{2},m_3-\frac{1}{2}} \\ \begin{pmatrix} j_3+\frac{1}{2} & \frac{1}{2} & j_3 \\ m_3+\frac{1}{2} & -\frac{1}{2} & -m_3\end{pmatrix} Y_{j_3+\frac{1}{2},m_3+\frac{1}{2}}\end{pmatrix} =$$

$$= (-1)^{j_3+m_3}\sqrt{2j_3+1}\begin{pmatrix}(-1)^{2j_3+1}(-1)^{j_3-m_3}\sqrt{\frac{j_3-m_3+1}{(2j_3+1)(2j_3+2)}}Y_{j_3+\frac{1}{2},m_3-\frac{1}{2}} \\ (-1)^{j_3+m_3}\sqrt{\frac{j_3+m_3+1}{(2j_3+1)(2j_3+2)}}Y_{j_3+\frac{1}{2},m_3+\frac{1}{2}}\end{pmatrix} =$$

$$= \begin{pmatrix}(-1)^{4j_3+1}\sqrt{\frac{j_3-m_3+1}{2j_3+2}}Y_{j_3+\frac{1}{2},m_3-\frac{1}{2}} \\ \sqrt{\frac{j_3+m_3+1}{2j_3+2}}Y_{j_3+\frac{1}{2},m_3+\frac{1}{2}}\end{pmatrix} =$$

$$= \frac{1}{\sqrt{2j_3+2}}\begin{pmatrix}-\sqrt{j_3-m_3+1}\,Y_{j_3+\frac{1}{2},m_3-\frac{1}{2}} \\ \sqrt{j_3+m_3+1}\,Y_{j_3+\frac{1}{2},m_3+\frac{1}{2}}\end{pmatrix} =$$

$$= \frac{1}{\sqrt{2l+1}}\begin{pmatrix}-\sqrt{l-m_3+\frac{1}{2}}\,Y_{l,m_3-\frac{1}{2}} \\ \sqrt{l+m_3+\frac{1}{2}}\,Y_{l,m_3+\frac{1}{2}}\end{pmatrix}$$

The allowed values for m_3 are $m_3 = -l+\frac{1}{2}, -l+\frac{1}{2}+1, \ldots, l-\frac{1}{2}$, total of $2l$ values (in particular, the values $m_3 = \pm(l+\frac{1}{2})$ are not allowed).

The last formula is the spin spherical harmonics given in terms of l, m_3, the second last formula is in terms of j_3, m_3 (both are used). The spin spherical harmonics is usually denoted by $\chi_\kappa^{m_3}$ or $y_l^{j_3,m_3}$. See the next section for the definition of κ.

Kappa

In order to define the state, one needs to specify both j_3 and l (distinguishng the two cases $j_3 = l \pm \frac{1}{2}$). This can be unified into just one integer κ, where $-\hbar\kappa$ is defined as the eigenvalue of the operator:

$$K = \boldsymbol{\sigma} \cdot \mathbf{L} + \hbar = \left(\frac{2}{\hbar^2} \mathbf{S} \cdot \mathbf{L} + 1 \right) \hbar = \left(\frac{1}{\hbar^2} \left(\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2 \right) + 1 \right) \hbar$$

Then:

$$\begin{aligned} K\psi = \left(\frac{1}{\hbar^2} \left(\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2 \right) + 1 \right) \hbar\psi = \\ = \left(j_3(j_3+1) - l(l+1) - s(s+1) + 1 \right) \hbar\psi = \\ = \left(j_3(j_3+1) - l(l+1) + \frac{1}{4} \right) \hbar\psi = \\ = -\kappa\hbar\psi \end{aligned}$$

from which

$$\begin{aligned} \kappa = -j_3(j_3+1) + l(l+1) - \frac{1}{4} = \\ = \begin{cases} -j_3(j_3+1) + (j_3 - \frac{1}{2})(j_3 - \frac{1}{2} + 1) - \frac{1}{4}; & \text{for } j_3 = l + \frac{1}{2} \\ -j_3(j_3+1) + (j_3 + \frac{1}{2})(j_3 + \frac{1}{2} + 1) - \frac{1}{4}; & \text{for } j_3 = l - \frac{1}{2} \end{cases} = \\ = \begin{cases} -(j_3 + \frac{1}{2}); & \text{for } j_3 = l + \frac{1}{2} \\ +(j_3 + \frac{1}{2}); & \text{for } j_3 = l - \frac{1}{2} \end{cases} = \\ = \begin{cases} -l - 1; & \text{for } j_3 = l + \frac{1}{2} \\ l; & \text{for } j_3 = l - \frac{1}{2} \end{cases} \end{aligned}$$

The opposite relation is:

$$l = \begin{cases} -\kappa - 1; & \text{for } \kappa < 0, \text{ equivalently } j_3 = l + \frac{1}{2} \\ \kappa; & \text{for } \kappa > 0, \text{ equivalently } j_3 = l - \frac{1}{2} \end{cases}$$

Code:

```
>>> from sympy import var, S
>>> var("j l")
(j, l)
>>> k = -j*(j+1) + l*(l+1) - S(1)/4
>>> k.subs(l, j-S(1)/2).expand()
-j - 1/2
>>> k.subs(l, j+S(1)/2).expand()
j + 1/2
```

Some useful relations with κ that follow from the above for both cases $j_3 = l \pm \frac{1}{2}$:

$$\begin{aligned} l(l+1) = \kappa(\kappa+1) \\ l = |\kappa + \tfrac{1}{2}| - \tfrac{1}{2} \\ j_3 = |\kappa| - \tfrac{1}{2} \end{aligned}$$

In order to enumerate all possibilities, one needs to count all integers except zero: $\kappa = -1, 1, -2, 2, -3, 3, \ldots$:

κ	l	j_3	$j_3 - l$	degeneracy	label
-1	0	0.5	0.5	2	$s_{1/2}$
1	1	0.5	-0.5	2	$p_{1/2}$
-2	1	1.5	0.5	4	$p_{3/2}$
2	2	1.5	-0.5	4	$d_{3/2}$
-3	2	2.5	0.5	6	$d_{5/2}$
3	3	2.5	-0.5	6	$f_{5/2}$
-4	3	3.5	0.5	8	$f_{7/2}$
4	4	3.5	-0.5	8	$g_{7/2}$
-5	4	4.5	0.5	10	$g_{9/2}$
5	5	4.5	-0.5	10	$h_{9/2}$
-6	5	5.5	0.5	12	$h_{11/2}$
6	6	5.5	-0.5	12	$i_{11/2}$
$\ldots$					

The degeneracy of the individual states for each κ is equal to $2j_3 + 1 = 2|\kappa|$ (which is equal to $2l + 2$ for $j_3 = l + \frac{1}{2}$ and $2l$ for $j_3 = l - \frac{1}{2}$, see the previous section), that is 2, 4, 6, 8 for $j_3 = 0.5, 1.5, 2.5, 3.5$ (or equivalently $\kappa = \pm 1, \pm 2, \pm 3, \pm 4$) respectively. All states together with the given l have total degeneracy $2l + 2 + 2l = 2(2l + 1)$, that is 2, 6, 10, 14 for $l = 0, 1, 2, 3$ respectively.

The states are labeled by a letter corresponding to $l = 0, 1, 2, 3, \ldots$ (s, p, d, f, g, h, i, j, k, l, m, n, o, q, r, t, u, v, w, x, y, z, a, b, c, e, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z, A, B, C, D, E) with a subscript equal to the total angular momentum $j_3 = n/2$ with $n = 1, 3, 5, 7, \ldots$.

8.7 Examples

8.7.1 Two Particles in Harmonic Potential

It is a 1D, two-body problem with an interacting Hamiltonian

$$H(x_1, x_2) = -\tfrac{1}{2}\frac{\partial^2}{\partial x_1^2} - \tfrac{1}{2}\frac{\partial^2}{\partial x_2^2} + \frac{1}{|x_1 - x_2|} + \tfrac{1}{2}\omega^2 x_1^2 + \tfrac{1}{2}\omega^2 x_2^2$$

and it can be solved analytically. The Schrödinger equation is

$$\left(-\tfrac{1}{2}\frac{\partial^2}{\partial x_1^2} - \tfrac{1}{2}\frac{\partial^2}{\partial x_2^2} + \frac{1}{|x_1 - x_2|} + \tfrac{1}{2}\omega^2 x_1^2 + \tfrac{1}{2}\omega^2 x_2^2\right)\Psi(x_1, x_2) = E\Psi(x_1, x_2)$$

we use the substitution:

$$u = \frac{1}{\sqrt{2}}(x_1 - x_2)$$
$$v = \frac{1}{\sqrt{2}}(x_1 + x_2)$$

then

$$\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} = \frac{\partial^2}{\partial u^2} + \frac{\partial^2}{\partial v^2}$$
$$|x_1 - x_2| = \sqrt{2}|u|$$
$$x_1^2 + x_2^2 = u^2 + v^2$$

and

$$\left(-\tfrac{1}{2}\frac{\partial^2}{\partial u^2}-\tfrac{1}{2}\frac{\partial^2}{\partial v^2}+\frac{1}{\sqrt{2}|u|}+\tfrac{1}{2}\omega^2u^2+\tfrac{1}{2}\omega^2v^2\right)\Psi(u,v)=E\Psi(u,v)$$

Note also the symmetry of the Hamiltonian $H(x_1,x_2)=H(x_2,x_1)$ which after substitution is equivalent to $H(u,v)=H(-u,v)$. Now we can separate the equation:

$$\begin{aligned}\Psi(u,v)&=f(u)g(v)\\ \left(-\tfrac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}u^2}+\frac{1}{\sqrt{2}|u|}+\tfrac{1}{2}\omega^2u^2\right)f_k(u)&=\epsilon_kf_k(u)\\ \left(-\tfrac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}v^2}+\tfrac{1}{2}\omega^2v^2\right)g_l(v)&=\epsilon_lg_l(v)\\ E_{kl}&=\epsilon_k+\epsilon_l\end{aligned}$$

the solution of the second equation is:

$$\begin{aligned}g_l(v)&=\frac{1}{\sqrt{2^ll!}}\left(\frac{\omega}{\pi}\right)^{\frac{1}{4}}e^{-\frac{\omega v^2}{2}}H_l(\sqrt{\omega}v)\\ \epsilon_l&=\omega(l+\tfrac{1}{2})\quad\text{for } l=0,1,2,\ldots\end{aligned}$$

where $H_n(x)$ are the Hermite polynomials:

$$H_n(x)=(-1)^ne^{x^2}\frac{\mathrm{d}^n}{\mathrm{d}x^n}e^{-x^2}$$

The solution to the first equation can be approximated around the minimum of the potential, which occurs at point $u=u_0$ (since the potential is symmetric with respect to u, we only treat the branch $u>0$):

$$\begin{aligned}V(u)=\frac{1}{\sqrt{2}|u|}+\tfrac{1}{2}\omega^2u^2&=\left(2^{-\frac{1}{3}}+2^{-\frac{4}{3}}\right)\omega^{\frac{2}{3}}+\frac{3}{2}\omega^2(u-u_0)^2+O\left((u-u_0)^3\right)\\ u_0&=2^{-\frac{1}{6}}\omega^{-\frac{2}{3}}\end{aligned}$$

So the first few states can be approximated by the harmonic oscillator solution with frequency $\sqrt{3}\omega$:

$$\begin{aligned}f_k(u)&=\frac{1}{\sqrt{2^kk!}}\left(\frac{\sqrt{3}\omega}{\pi}\right)^{\frac{1}{4}}e^{-\frac{\sqrt{3}\omega(u-u_0)^2}{2}}H_k(3^{\frac{1}{4}}\sqrt{\omega}(u-u_0))\\ \epsilon_k&=\left(2^{-\frac{1}{3}}+2^{-\frac{4}{3}}\right)\omega^{\frac{2}{3}}+\sqrt{3}\omega(k+\tfrac{1}{2})\quad\text{for } k=0,1,2,\ldots\end{aligned}$$

The final solution is then:

$$\begin{aligned}\Psi_{kl}(u,v)&=f_k(u)g_l(v)=\\ &=\frac{1}{\sqrt{2^kk!}}\left(\frac{\sqrt{3}\omega}{\pi}\right)^{\frac{1}{4}}e^{-\frac{\sqrt{3}\omega(u-u_0)^2}{2}}H_k(3^{\frac{1}{4}}\sqrt{\omega}(u-u_0))\frac{1}{\sqrt{2^ll!}}\left(\frac{\omega}{\pi}\right)^{\frac{1}{4}}e^{-\frac{\omega v^2}{2}}H_l(\sqrt{\omega}v)\\ E_{kl}&=\epsilon_k+\epsilon_l=\left(2^{-\frac{1}{3}}+2^{-\frac{4}{3}}\right)\omega^{\frac{2}{3}}+\sqrt{3}\omega(k+\tfrac{1}{2})+\omega(l+\tfrac{1}{2})\end{aligned}$$

8.7.2 Quantum Harmonic Oscillator

The quantum harmonic oscillator for one particle in 1D is:

$$i\hbar\frac{\partial}{\partial t}\psi(x,t)=-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x,t)+V(x)\psi(x,t)$$

$$V(x) = \frac{1}{2}m\omega^2 x^2$$

This is a partial differential equation for the time evolution of the wave function $\psi(x,t)$, but one method to solve it is the eigenvalues expansion:

$$\psi(x,t) = \sum_E c_E \psi_E(x) e^{-\frac{i}{\hbar}Et}$$

where the sum goes over the whole spectrum (for continuous spectrum the sum turns into an integral), the c_E coefficients are determined from the initial condition and $\psi_E(x)$ satisfies the one dimensional one particle time independent Schrödinger equation:

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}\psi_E(x) + V(x)\psi_E(x) = E\psi_E(x)$$

and this is just an ODE and thus can be solved with Hermes1D. There can be many types of boundary conditions for this equation, depending on the physical problem, but in our case we simply have $\lim_{x\to\pm\infty}\psi_E(x) = 0$ and the normalization condition $\int_{-\infty}^{\infty}|\psi_E(x)|^2\mathrm{d}x = 1$.

We can set $m = \hbar = 1$ and from now on we'll just write $\psi(x)$ instead of $\psi_E(x)$:

$$-\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}\psi(x) + V(x)\psi(x) = E\psi(x)$$

and we will solve it on the interval (a,b) with the boundary condition $\psi(a) = \psi(b) = 0$. The weak formulation is

$$\int_a^b \frac{1}{2}\frac{\mathrm{d}\psi(x)}{\mathrm{d}x}\frac{\mathrm{d}v(x)}{\mathrm{d}x} + V(x)\psi(x)v(x)\,\mathrm{d}x - \left[\frac{\mathrm{d}\psi(x)}{\mathrm{d}x}v(x)\right]_b^a = E\int_a^b \psi(x)v(x)\,\mathrm{d}x$$

but due to the boundary condition $v(a) = v(b) = 0$ so $[\psi'(x)v(x)]_b^a = 0$ and we get

$$\int_a^b \frac{1}{2}\frac{\mathrm{d}\psi(x)}{\mathrm{d}x}\frac{\mathrm{d}v(x)}{\mathrm{d}x} + V(x)\psi(x)v(x)\,\mathrm{d}x = E\int_a^b \psi(x)v(x)\,\mathrm{d}x$$

And the finite element formulation is then $\psi(x) = \sum_j y_j\phi_j(x)$ and $v = \phi_i(x)$:

$$\left(\int_a^b \frac{1}{2}\phi_i'(x)\phi_j'(x) + V(x)\phi_i(x)\phi_j(x)\,\mathrm{d}x\right)y_j = E\int_a^b \phi_i(x)\phi_j(x)\,\mathrm{d}x\; y_j$$

which is a generalized eigenvalue problem:

$$A_{ij}y_j = EB_{ij}y_j$$

with

$$A_{ij} = \int_a^b \frac{1}{2}\phi_i'(x)\phi_j'(x) + V(x)\phi_i(x)\phi_j(x)\,\mathrm{d}x$$

$$B_{ij} = \int_a^b \phi_i(x)\phi_j(x)\,\mathrm{d}x$$

8.7.3 Radial Schrödinger Equation

Another important example is the three dimensional one particle time independent Schrödinger equation for a spherically symmetric potential:

$$-\frac{1}{2}\nabla^2\psi(\mathbf{x}) + V(r)\psi(\mathbf{x}) = E\psi(\mathbf{x})$$

The way to solve it is to separate the equation into radial and angular parts by writing the Laplace operator in spherical coordinates as:

$$\nabla^2 f = \frac{\partial^2 f}{\partial \rho^2} + \frac{2}{\rho}\frac{\partial^2 f}{\partial \rho^2} - \frac{L^2}{\rho^2}$$

$$L^2 = -\frac{\partial^2 f}{\partial \theta^2} - \frac{1}{\sin^2\theta}\frac{\partial^2 f}{\partial \phi^2} - \frac{1}{\tan\theta}\frac{\partial f}{\partial \theta}$$

Substituting $\psi(\mathbf{x}) = R(\rho)Y(\theta,\phi)$ into the Schrödinger equation yields:

$$-\frac{1}{2}\nabla^2(RY) + VRY = ERY$$

$$-\frac{1}{2}R''Y - \frac{1}{\rho}R'Y + \frac{L^2 RY}{2\rho^2} + VRY = ERY$$

Using the fact that $L^2 Y = l(l+1)Y$ we can cancel Y and we get the radial Schrödinger equation:

$$-\frac{1}{2}R'' - \frac{1}{\rho}R' + \frac{l(l+1)R}{2\rho^2} + VR = ER$$

The solution is then:

$$\psi(\mathbf{x}) = \sum_{nlm} c_{nlm} R_{nl}(r) Y_{lm}\left(\frac{\mathbf{x}}{r}\right)$$

where $R_{nl}(r)$ satisfies the radial Schrödinger equation (from now on we just write $R(r)$):

$$-\frac{1}{2}R''(r) - \frac{1}{r}R'(r) + \left(V + \frac{l(l+1)}{2r^2}\right)R(r) = ER(r)$$

Again there are many types of boundary conditions, but the most common case is $\lim_{r\to\infty} R(r) = 0$ and $R(0) = 1$ or $R(0) = 0$. One solves this equation on the interval $(0, a)$ for large enough a.

The procedure is similar to the previous example, only we need to remember that we always have to use covariant integration (in the previous example the covariant integration was the same as the coordinate integration), in this case $r^2 \sin\theta \mathrm{d}r\mathrm{d}\theta\mathrm{d}\phi$, so the weak formulation is:

$$\int \left(-\frac{1}{2}R''(r) - \frac{1}{r}R'(r) + \left(V + \frac{l(l+1)}{2r^2}\right)R(r)\right) v(r) r^2 \sin\theta \mathrm{d}r\mathrm{d}\theta\mathrm{d}\phi =$$

$$= \int ER(r)v(r)r^2 \sin\theta \mathrm{d}r\mathrm{d}\theta\mathrm{d}\phi$$

Integrating over the angles gives 4π which we cancel out at both sides and we get:

$$\int_0^a \left(-\frac{1}{2}R''(r) - \frac{1}{r}R'(r) + \left(V + \frac{l(l+1)}{2r^2}\right)R(r)\right) v(r) r^2 \mathrm{d}r =$$

$$= E\int_0^a R(r)v(r)r^2 \mathrm{d}r$$

We apply per partes to the first two terms on the left hand side:

$$\int_0^a \left(-\frac{1}{2}R''(r) - \frac{1}{r}R'(r)\right) v(r) r^2 \mathrm{d}r = \int_0^a -\frac{1}{2r^2}\left(r^2 R'(r)\right)' v(r) r^2 \mathrm{d}r =$$

$$= \int_0^a -\frac{1}{2}\left(r^2 R'(r)\right)' v(r) \mathrm{d}r = \int_0^a \frac{1}{2} r^2 R'(r) v'(r) \mathrm{d}r - \frac{1}{2}[r^2 R'(r) v(r)]_0^a =$$

$$= \int_0^a \frac{1}{2} R'(r) v'(r) r^2 \mathrm{d}r - \frac{1}{2} a^2 R'(a) v(a)$$

We used the fact that $\lim_{r\to 0} r^2 R'(r) = 0$. If we also prescribe the boundary condition $R'(a) = 0$, then the boundary term vanishes completely. The weak formulation is then:

$$\int_0^a \frac{1}{2} R'(r) v'(r) r^2 + \left(V + \frac{l(l+1)}{2r^2} \right) R(r) v(r) r^2 \, \mathrm{d}r = E \int_0^a R(r) v(r) r^2 \, \mathrm{d}r$$

or

$$\int_0^a \frac{1}{2} R'(r) v'(r) r^2 + V(r) R(r) v(r) r^2 + \frac{l(l+1)}{2} R(r) v(r) \, \mathrm{d}r = E \int_0^a R(r) v(r) r^2 \, \mathrm{d}r$$

Another approach

Another (equivalent) approach is to write a weak formulation for the 3D problem in cartesian coordinates:

$$\int_\Omega \frac{1}{2} \nabla \psi(\mathbf{x}) \nabla v(\mathbf{x}) + V(r) \psi(\mathbf{x}) v(\mathbf{x}) \, \mathrm{d}^3 x = E \int_\Omega \psi(\mathbf{x}) v(\mathbf{x}) \, \mathrm{d}^3 x$$

and only then transform to spherical coordinates:

$$\int_0^{2\pi} \mathrm{d}\varphi \int_0^\pi \mathrm{d}\theta \int_0^a \mathrm{d}r \left(\frac{1}{2} \nabla \psi(\mathbf{x}) \nabla v(\mathbf{x}) + V(r) \psi(\mathbf{x}) v(\mathbf{x}) \right) r^2 \sin\theta =$$

$$= E \int_0^{2\pi} \mathrm{d}\varphi \int_0^\pi \mathrm{d}\theta \int_0^a \mathrm{d}r \, \psi(\mathbf{x}) v(\mathbf{x}) r^2 \sin\theta$$

The 3d eigenvectors $\psi(\mathbf{x})$ however are not spherically symmetric. Nevertheless we can still proceed by choosing our basis as

$$v_{ilm}(\mathbf{x}) = \phi_{il}(r) Y_{lm}(\theta, \varphi)$$

and seek our solution as

$$\psi(\mathbf{x}) = \sum_{jlm} y_{jlm} \phi_{jl}(r) Y_{lm}(\theta, \varphi)$$

Using the properties of spherical harmonics and the gradient:

$$\int Y_{lm} Y_{l'm'} \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\varphi = \delta_{ll'} \delta_{mm'}$$

$$\int r^2 \nabla Y_{lm} \nabla Y_{l'm'} \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\varphi = l(l+1) \delta_{ll'} \delta_{mm'}$$

$$\nabla f = \frac{\partial f}{\partial r} \hat{\boldsymbol{r}} + \frac{1}{r} \frac{\partial f}{\partial \theta} \hat{\boldsymbol{\theta}} + \frac{1}{r \sin\theta} \frac{\partial f}{\partial \phi} \hat{\boldsymbol{\phi}}$$

the weak formulation becomes:

$$\left(\int_0^a \frac{1}{2} r^2 \phi'_{il}(r) \phi'_{jl}(r) + \frac{1}{2} X + \frac{l(l+1)}{2} \phi_{il}(r) \phi_{jl}(r) + r^2 V(r) \phi_{il}(r) \phi_{jl}(r) \, \mathrm{d}r \right) y_{jlm} =$$

$$= E \int_0^a r^2 \phi_{il}(r) \phi_{jl}(r) \, \mathrm{d}r \; y_{jlm}$$

where both l and m indices are given by the indices of the particular base function v_{ilm}. The X term is (schematically):

$$X = \int r^2 \sin\theta(r) Y_{lm}(\theta, \varphi) (\phi_{il} \nabla \phi_{jl} + \nabla \phi_{il} \phi_{jl}) \nabla Y_{lm}$$

There is an interesting identity:

$$\int r \hat{\mathbf{r}} Y_{lm} \nabla Y_{l'm'} \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\varphi = 0$$

But it cannot be applied, because we have one more r in the expression. Nevertheless the term is probably zero, as can be seen when we compare the weak formulation to the one we got directly from the radial equation.

How Not To Derive The Weak Formulation

If we forgot that we have to integrate covariantly, this section is devoted to what happens if we integrate using the coordinate integration. We would get:

$$\int_0^a \frac{1}{2}R'(x)v'(x) - \frac{1}{r}R'(x)v(x) + \left(V + \frac{l(l+1)}{2r^2}\right)R(x)v(x)\,\mathrm{d}x = E\int_0^a R(x)v(x)\,\mathrm{d}x$$

Notice the matrix on the left hand side is not symmetric. There is another way of writing the weak formulation by applying per-partes to the $R'(r)v(r)$ term:

$$-\int_0^a \frac{1}{r}R'(x)v(x)\mathrm{d}x =$$

$$= \int_0^a \frac{1}{r}R(x)v'(x)\mathrm{d}x - \int_0^a \frac{1}{r^2}R(x)v(x)\mathrm{d}x - \left[\frac{1}{r}R'(x)v'(x)\right]_0^a + \left[\frac{1}{r^2}R'(x)v(x)\right]_0^a$$

We can use $v(a) = 0$ and $R'(a) = 0$ to simplify a bit:

$$-\int_0^a \frac{1}{r}R'(x)v(x)\mathrm{d}x =$$

$$= \int_0^a \frac{1}{r}R(x)v'(x)\mathrm{d}x - \int_0^a \frac{1}{r^2}R(x)v(x)\mathrm{d}x + \lim_{r\to 0}\left(\frac{R'(x)v'(x)}{r} - \frac{R'(x)v(x)}{r^2}\right)$$

Since $R(x) \sim r^l$ near $r = 0$, we can see that for $l \geq 3$ the limits on the right hand side are zero, but for $l = 0, 1, 2$ they are not zero and need to be taken into account. Let's assume $l \geq 3$ for now, then our weak formulation looks like:

$$\int_0^a \frac{1}{2}R'(x)v'(x) + \frac{1}{r}R(x)v'(x) + \left(V + \frac{l(l+1)}{2r^2} - \frac{1}{r^2}\right)R(x)v(x)\,\mathrm{d}x = E\int_0^a R(x)v(x)\,\mathrm{d}x$$

or

$$\int_0^a \frac{1}{2}R'(x)v'(x) + \frac{1}{r}R(x)v'(x) + \left(V + \frac{(l-2)(l+1)}{2r^2}\right)R(x)v(x)\,\mathrm{d}x = E\int_0^a R(x)v(x)\,\mathrm{d}x$$

The left hand side is also not symmetric, however we can now take an average of our both weak formulations to get a symmetric weak formulation:

$$\int_0^a \frac{1}{2}R'(x)v'(x) + \frac{R(x)v'(x) - R'(x)v(x)}{2r} + \left(V + \frac{l(l+1)-1}{2r^2}\right)R(x)v(x)\,\mathrm{d}x =$$

$$= E\int_0^a R(x)v(x)\,\mathrm{d}x$$

Keep in mind, that this symmetric version is only correct for $l \geq 3$. For $l < 3$ we need to use our first nonsymmetric version.

As you can see, this is something very different to what we got in the previous section. First there were lots of technical difficulties and second the final result is wrong, since it doesn't correspond to the 3D Schrödinger equation.

Scattering in radial potential

If $V = 0$, the radial equation is:

$$-\frac{1}{2}R''_{El}(r) - \frac{1}{r}R'_{El}(r) + \frac{l(l+1)}{2r^2}R_{El}(r) = ER_{El}(r)$$

The general solution is a linear combination of the spherical Bessel functions $j_l(kr)$ and $n_l(kr)$:

$$R_{El}(r) = A_l j_l(kr) + B_l n_l(kr)$$

where $k = \sqrt{2E}$ and $E > 0$ is a continuous spectrum. The asymptotic expansion for $r \to \infty$ is:

$$j_l(kr) \to \frac{1}{kr} \sin\left(kr - \frac{l\pi}{2}\right)$$
$$n_l(kr) \to \frac{1}{kr} \cos\left(kr - \frac{l\pi}{2}\right)$$

so we get for large r:

$$R_{El}(r) = A_l j_l(kr) + B_l n_l(kr) \to$$
$$\to A_l \frac{1}{kr} \sin\left(kr - \frac{l\pi}{2}\right) + B_l \frac{1}{kr} \cos\left(kr - \frac{l\pi}{2}\right) =$$
$$= \sqrt{A_l^2 + B_l^2} \frac{1}{kr} \sin\left(kr - \frac{l\pi}{2} + \operatorname{atan2}(B_l, A_l)\right) = C_l \frac{1}{kr} \sin\left(kr - \frac{l\pi}{2} + \delta_l\right)$$

where

$$\begin{aligned} \delta_l &= \operatorname{atan2}(B_l, A_l) \\ C_l &= \sqrt{A_l^2 + B_l^2} \end{aligned} \tag{8.7.3.1}$$

The C_l and δ_l are physical variables, so we express A_l and B_l using them:

$$\begin{aligned} A_l &= C_l \cos\delta_l \\ B_l &= C_l \sin\delta_l \end{aligned} \tag{8.7.3.2}$$

and write the exact solution R_{El} as:

$$R_{El}(r) = C_l(\cos\delta_l\, j_l(kr) + \sin\delta_l\, n_l(kr)) \tag{8.7.3.3}$$

We can then compare this to $\phi \approx e^{ikz} + f(\theta,\phi)\frac{e^{ikr}}{r}$, by expanding $e^{ikz} = e^{ikr\cos\theta} = \sum(2l+1)i^l j_l(kr)P_l(\cos\theta)$:

$$C_l = \frac{e^{i\delta_l}}{k}$$
$$f(\theta,\phi) = \frac{1}{2ik}\sum(2l+1)(e^{2i\delta_l} - 1)P_l(\cos\theta)$$

Since $\sigma(\theta) = |f(\theta)|^2$ and integrating over ω we get the total cross section:

$$\sigma = \frac{4\pi}{k}\sum(2l+1)\sin^2\delta_l$$

In order to find the phase shifts δ_l, we solve the radial equation for the full potential

$$-\frac{1}{2}R_{nl}''(r) - \frac{1}{r}R_{nl}'(r) + \left(V + \frac{l(l+1)}{2r^2}\right)R_{nl}(r) = ER_{nl}(r)$$

and then fit it to the above asymptotic solution for V=0. We require that the value and the slope must be continuous, so we use (8.7.3.3) and R_{nl} must satisfy the following two equations (for the value and the derivative) at the point $r = a$:

$$\begin{aligned} R_{El}(a) = Aj_l(ka) + Bn_l(ka)) &= C_l(\cos\delta_l\, j_l(ka) + \sin\delta_l\, n_l(ka)) \\ R_{El}'(a) = kAj_l'(ka) + kBn_l'(ka)) &= C_l k(\cos\delta_l\, j_l'(ka) + \sin\delta_l\, n_l'(ka)) \end{aligned} \tag{8.7.3.4}$$

This is a set of two equations for two unknowns C_l and δ_l. The solution is:

$$\begin{aligned} D &= k(j_l(ka)n_l'(ka) - j_l'(ka)n_l(ka)) \\ A &= \frac{R_{El}(a)kn_l'(ka) - R_{El}'(a)n_l(ka)}{D} \\ B &= -\frac{R_{El}(a)kj_l'(ka) - R_{El}'(a)j_l(ka)}{D} \end{aligned} \tag{8.7.3.5}$$

And one can calculate C_l and δ_l from (8.7.3.1). Code:

```
>>> from sympy import var, solve
>>> var("R Rp j jp n np A B k")
(R, Rp, j, jp, n, np, A, B, k)
>>> eq1 = R - A*j - B*n
>>> eq2 = Rp - k*A*jp - k*B*np
>>> solve([eq1, eq2], [A, B])
{A: (R*k*np - Rp*n)/(k*(j*np - jp*n)), B: (-R*jp*k + Rp*j)/(k*(j*np - jp*n))}
```

Another approach to calculate δ_l is to take the logarithmic derivative ($(\log|u|)' = \frac{u'}{u}$) at the point $r = a$:

$$\gamma_l \equiv \frac{\mathrm{d}}{\mathrm{d}r}\log|R_l(r)|\Big|_{r=a} = \frac{R'_{El}(a)}{R_{El}(a)} = \frac{C_l k(\cos\delta_l\, j'_l(ka) + \sin\delta_l\, n'_l(ka))}{C_l(\cos\delta_l\, j_l(ka) + \sin\delta_l\, n_l(ka))} =$$

$$= k\frac{j'_l(ka) + \tan\delta_l\, n'_l(ka)}{j_l(ka) + \tan\delta_l\, n_l(ka)} \tag{8.7.3.6}$$

and solving for δ_l we get:

$$\tan\delta_l = -\frac{kj'_l(ka) - \gamma_l j_l(ka)}{kn'_l(ka) - \gamma_l n_l(ka)} = -\frac{-kj_{l+1}(ka) + kl\frac{j_l(ka)}{ka} - \gamma_l j_l(ka)}{-kn_{l+1}(ka) + kl\frac{n_l(ka)}{ka} - \gamma_l n_l(ka)} =$$

$$= -\frac{kaj_{l+1}(ka) - lj_l(ka) + aj_l(ka)\gamma_l}{kan_{l+1}(ka) - ln_l(ka) + an_l(ka)\gamma_l} \tag{8.7.3.7}$$

where we used the following relations:

$$j'_l(z) = -j_{l+1}(z) + l\frac{j_l(z)}{z}$$
$$n'_l(z) = -n_{l+1}(z) + l\frac{n_l(z)}{z}$$

The disadvantage of (8.7.3.7) is that we only know $\tan\delta_l$, while in (8.7.3.5) we know δ_l directly using the $\operatorname{atan}2$ function.

Now we can use these δ_l in the formula for the total cross section. We can define a reduced phase-shift η_l by

$$\delta_l = (n - l - 1)\pi + \eta_l$$

where $n - l - 1$ is the number of radial nodes and $0 \le \eta_l \le \pi$.

The problem can now be formulated in two ways. Either to solve the radial equation for a potential with finite reach and then "measure" those phase shifts in the solution. Or by prescribing those phase shifts and we now need to calculate the solutions (e.g. the energies) from the radial equation.

8.8 Radial Schrödinger and Dirac Equations

8.8.1 Variational Formulation of the Schrödinger equation

Lagrangian is:

$$\mathcal{L}(\psi) = \tfrac{1}{2}(\nabla\psi)^2 + V(x)\psi^2(x)$$

Subject to the normalization constrain:

$$N[\psi] = \int |\psi(x)|^2 \mathrm{d}^3x - 1 = 0$$

The action is:

$$S[\psi] = \int \mathcal{L} \mathrm{d}^3 x$$

Variating it (subject to the normalization condition) we get:

$$\begin{aligned} 0 = \delta(S - \epsilon N) = \delta \int \tfrac{1}{2}(\nabla\psi)^2 + V(x)\psi^2(x)\mathrm{d}^3x - \epsilon\left(\int |\psi(x)|^2 \mathrm{d}^3x - 1\right) = \\ = \int (\nabla\psi)\cdot(\nabla\delta\psi) + 2V\psi\delta\psi - 2\epsilon\psi\delta\psi \mathrm{d}^3x \\ = 2\int \left(-\tfrac{1}{2}\nabla^2\psi + V\psi - \epsilon\psi\right)\delta\psi \mathrm{d}^3x + \int (\mathbf{n}\cdot\nabla\psi)\delta\psi \,\mathrm{d}^2x \end{aligned}$$

Which gives the Schrödinger equation assuming the surface integral vanishes.

Note: to apply the variation δ correctly, one uses the definition:

$$\delta F[\psi] \equiv \frac{\mathrm{d}}{\mathrm{d}\epsilon} F[\psi + \epsilon\delta\psi]\Big|_{\epsilon=0}$$

Weak Formulation

The weak formulation is obtained from the above by substituting $\delta\psi \to v$ (the test function) so we get:

$$\int \tfrac{1}{2}(\nabla\psi)\cdot(\nabla v) + V\psi v - \epsilon\psi v \,\mathrm{d}^3x$$

8.8.2 Radial Schrödinger equation

There are two ways to obtain the radial Schrödinger equation. Either from the Lagrangian, or from the equation itself.

From the Equation

$$-\frac{1}{2}\nabla^2\psi(\mathbf{x}) + V(r)\psi(\mathbf{x}) = E\psi(\mathbf{x})$$

The way to solve it is to separate the equation into radial and angular parts by writing the Laplace operator in spherical coordinates as:

$$\nabla^2 f = \frac{\partial^2 f}{\partial\rho^2} + \frac{2}{\rho}\frac{\partial f}{\partial\rho} - \frac{L^2 f}{\rho^2}$$

$$L^2 = -\frac{\partial^2}{\partial\theta^2} - \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} - \frac{1}{\tan\theta}\frac{\partial}{\partial\theta}$$

Substituting $\psi(\mathbf{x}) = R(\rho)Y(\theta,\phi)$ into the Schrödinger equation yields:

$$-\frac{1}{2}\nabla^2(RY) + VRY = ERY$$

$$-\frac{1}{2}R''Y - \frac{1}{\rho}R'Y + \frac{L^2 RY}{2\rho^2} + VRY = ERY$$

Using the fact that $L^2 Y = l(l+1)Y$ we can cancel Y and we get the radial Schrödinger equation:

$$-\frac{1}{2}R'' - \frac{1}{\rho}R' + \frac{l(l+1)R}{2\rho^2} + VR = ER$$

Normalization:

$$1 = \int |\psi|^2 \mathrm{d}^3x = \int R^2|Y|^2\mathrm{d}^3x = \int R^2|Y|^2\rho^2\mathrm{d}\Omega\mathrm{d}\rho = \int R^2\rho^2\mathrm{d}\rho\int |Y|^2\mathrm{d}\Omega = \int R^2\rho^2\mathrm{d}\rho$$

From the Lagrangian

We need to convert the Lagrangian to spherical coordinates. In order to easily make sure we do things covariantly, we start from the action (which is a scalar):

$$\begin{aligned}
S[\psi] &= \int \tfrac{1}{2}(\nabla\psi)^2 + V(x)\psi^2(x)\,\mathrm{d}^3x = \\
&= \int (\tfrac{1}{2}(\nabla(RY))^2 + V(RY)^2)\rho^2\mathrm{d}\rho\mathrm{d}\Omega = \\
&= \int (\tfrac{1}{2}(R'^2Y^2 + R^2(\nabla Y)^2 + 2RR'(\hat{\rho}Y)\cdot\nabla Y) + V(RY)^2)\rho^2\mathrm{d}\rho\mathrm{d}\Omega = \\
&= \int \left(\tfrac{1}{2}\left(R'^2 + R^2\frac{l(l+1)}{\rho^2}\right) + VR^2\right)\rho^2\mathrm{d}\rho = \\
&= \int \tfrac{1}{2}\rho^2R'^2 + (\rho^2V + \tfrac{1}{2}l(l+1))R^2\,\mathrm{d}\rho\,,
\end{aligned}$$

where we used the following properties of spherical harmonics:

$$\begin{aligned}
\int Y^2\mathrm{d}\Omega &= 1 \\
\int \rho^2(\nabla Y)^2\mathrm{d}\Omega &= l(l+1) \\
(Y\hat{\rho})\cdot(\rho\nabla Y) &= 0
\end{aligned}$$

We now minimize the action (subject to the normalization $\int \rho^2R^2\mathrm{d}\rho = 1$) to obtain the radial equation:

$$\begin{aligned}
0 = \delta(S - \epsilon N) &= \delta\int \tfrac{1}{2}\rho^2R'^2 + (\rho^2V + \tfrac{1}{2}l(l+1))R^2 - \epsilon\rho^2R^2\,\mathrm{d}\rho = \\
&= 2\int \tfrac{1}{2}\rho^2R'(\delta R)' + (\rho^2V + \tfrac{1}{2}l(l+1))R\delta R - \epsilon\rho^2R\delta R\,\mathrm{d}\rho = \\
&= 2\int \left((-\tfrac{1}{2}\rho^2R')' + (\rho^2V + \tfrac{1}{2}l(l+1))R - \epsilon\rho^2R\right)\delta R\,\mathrm{d}\rho + [\rho^2R'\delta R]_b^a
\end{aligned}$$

So the radial equation is:

$$(-\tfrac{1}{2}\rho^2R')' + (\rho^2V + \tfrac{1}{2}l(l+1))R = \epsilon\rho^2R \tag{8.8.2.1}$$

In agreement with the previous result.

Solving for u=rR

We can also make the substitution $u = rR$ and solve for u:

$$\begin{aligned}
R &= \frac{u}{r} \\
R' &= \frac{u'}{r} - \frac{u}{r^2}
\end{aligned}$$

and we substitute this to (8.8.2.1):

$$\begin{aligned}
-\tfrac{1}{2}\left(r^2\left(\frac{u'}{r} - \frac{u}{r^2}\right)\right)' + \left(V + \frac{l(l+1)}{2r^2}\right)ru &= \epsilon ru \\
-\tfrac{1}{2}ru'' + \left(V + \frac{l(l+1)}{2r^2}\right)ru &= \epsilon ru \\
-\tfrac{1}{2}u'' + \left(V + \frac{l(l+1)}{2r^2}\right)u &= \epsilon u
\end{aligned}$$

Perturbative Correction to Energy

We introduce P and Q by $P(r) = u(r)$ and $Q(r) = P'(r) = u'(r)$. The radial Schrödinger equation is then:

$$\begin{aligned} P'(r) &= Q(r) \\ Q'(r) &= -2\left(E - V(r) - \frac{l(l+1)}{2r^2}\right) P(r) \end{aligned}$$

Let P_1 and Q_1 represent the radial wave function and its derivative at E_1 and P_2, Q_2 at E_2, so the following holds:

$$\begin{aligned} Q_1'(r) &= -2\left(E_1 - V(r) - \frac{l(l+1)}{2r^2}\right) P_1(r) \\ Q_2'(r) &= -2\left(E_2 - V(r) - \frac{l(l+1)}{2r^2}\right) P_2(r) \end{aligned}$$

Now we evaluate $(Q_2P_1 - P_2Q_1)'$ using the relations above:

$$(Q_2P_1 - P_2Q_1)' = Q_2'P_1 + Q_2P_1' - P_2'Q_1 - P_2Q_1' = Q_2'P_1 + Q_2Q_1' - Q_2'Q_1 - P_2Q_1' = Q_2'P_1 - P_2Q_1' = 2(E_1 - E_2)P_1P_2$$

We integrate the last formula on the intervals $(0, a_c)$ and (a_c, ∞):

$$\begin{aligned} {}_0^{a_c} &= 2(E_1 - E_2)\int_0^{a_c} P_1(r)P_2(r)\,\mathrm{d}r \\ [Q_2P_1 - P_2Q_1]_{a_c}^{\infty} &= 2(E_1 - E_2)\int_{a_c}^{\infty} P_1(r)P_2(r)\,\mathrm{d}r \end{aligned}$$

On the interval $(0, a_c)$ we know the exact solution corresponding to the energies E_1 and E_2 by integrating outwards (the solution will eventually diverge for large r except for the eigenvalues, but we only need it up to a_c) and we know that $P_1(0) = P_2(0) = 0$, so we get:

$$Q_2(a_c^-)P_1(a_c^-) - P_2(a_c^-)Q_1(a_c^-) = 2(E_1 - E_2)\int_0^{a_c} P_1(r)P_2(r)\,\mathrm{d}r$$

where a_c^- means that we need the values at a_c when integrating the equation from the left (the value will generally be different when integrating the equation from the right, unless the energy is an eigenvalue). Similarly on the other interval where $P_1(\infty) = P_2(\infty) = 0$:

$$-(Q_2(a_c^+)P_1(a_c^+) - P_2(a_c^+)Q_1(a_c^+)) = 2(E_1 - E_2)\int_{a_c}^{\infty} P_1(r)P_2(r)\,\mathrm{d}r$$

Taking the sum of the last two expressions:

$$2(E_1 - E_2)\int_0^{\infty} P_1(r)P_2(r)\,\mathrm{d}r = Q_2(a_c^-)P_1(a_c^-) - P_2(a_c^-)Q_1(a_c^-) - (Q_2(a_c^+)P_1(a_c^+) - P_2(a_c^+)Q_1(a_c^+))$$

Now we use the fact that $P_1(a_c^-) = P_1(a_c^+)$ and $P_2(a_c^-) = P_2(a_c^+)$, because we match the two solutions from the left and right, so that the function is continuous (it's derivative will have a jump though):

$$2(E_1 - E_2)\int_0^{\infty} P_1(r)P_2(r)\,\mathrm{d}r = P_1(a_c)(Q_2(a_c^-) - Q_2(a_c^+)) - P_2(a_c)(Q_1(a_c^-) - Q_1(a_c^+))$$

By requiring, that the energy E_2 is an eigenvalue, it follows that there is no jump in the derivative, so we set $Q_2(a_c^-) = Q_2(a_c^+)$ and we get:

$$2(E_1 - E_2)\int_0^{\infty} P_1(r)P_2(r)\,\mathrm{d}r = -P_2(a_c)(Q_1(a_c^-) - Q_1(a_c^+))$$

that gives us an exact formula for the eigenvalue E_2:

$$E_2 = E_1 + \frac{P_2(a_c)(Q_1(a_c^-) - Q_1(a_c^+))}{2\int_0^\infty P_1(r)P_2(r)\,\mathrm{d}r}$$

We approximate the value of $P_2(a_c)$ by $P_1(a_c)$ as well as the integral $\int_0^\infty P_1(r)P_2(r)\,\mathrm{d}r$ by $\int_0^\infty P_1^2(r)\,\mathrm{d}r$ and we get an approximation for the eigenenergy:

$$E_2 \approx E_1 + \frac{P_1(a_c)(Q_1(a_c^-) - Q_1(a_c^+))}{2\int_0^\infty P_1^2(r)\,\mathrm{d}r}$$

We use this approximation iteratively until the convergence is achieved (the discontinuity in $Q(r)$ at $r = a_c$ is small enough, or equivalently, the correction to the energy is small enough).

For Dirac equation, one obtains a similar formula:

$$E_2 \approx E_1 + c\frac{P_1(a_c)(Q_1(a_c^-) - Q_1(a_c^+))}{\int_0^\infty P_1^2(r) + Q_1^2(r)\,\mathrm{d}r}$$

So it is just the previous formula multiplied by $2c$ and the normalization is calculated using both P and Q (as usual for the Dirac equation).

Weak Formulation

The weak formulation is obtained from the action above by substituting $\delta R \to v$ (the test function) so we get:

$$\int \tfrac{1}{2}\rho^2 R'v' + (\rho^2 V + \tfrac{1}{2}l(l+1))Rv\,\mathrm{d}\rho = \epsilon\int \rho^2 Rv\,\mathrm{d}\rho$$

We can also start from the equation itself, multiply by a test function v:

$$(-\tfrac{1}{2}\rho^2 R')'v + (\rho^2 V + \tfrac{1}{2}l(l+1))Rv = \epsilon\rho^2 Rv$$

We integrate it. Normally we need to be using $\rho^2\mathrm{d}\rho$ in order to integrate covariantly, but the above equation was already multiplied by ρ^2 (i.e. strictly speaking, it is not coordinate independent anymore), so we only integrate by $\mathrm{d}\rho$:

$$\int (-\tfrac{1}{2}\rho^2 R')'v + (\rho^2 V + \tfrac{1}{2}l(l+1))Rv\mathrm{d}\rho = \epsilon\int \rho^2 Rv\mathrm{d}\rho$$

After integration by parts:

$$\int \tfrac{1}{2}\rho^2 R'v' + (\rho^2 V + \tfrac{1}{2}l(l+1))Rv\mathrm{d}\rho - \tfrac{1}{2}[\rho^2 R'v]_0^a = \epsilon\int \rho^2 Rv\mathrm{d}\rho$$

Where a is the end of the domain (the origin is at 0). The boundary term is zero at the origin, so we get:

$$\int \tfrac{1}{2}\rho^2 R'v' + (\rho^2 V + \tfrac{1}{2}l(l+1))Rv\mathrm{d}\rho + \tfrac{1}{2}\rho^2 R'(a)v(a) = \epsilon\int \rho^2 Rv\mathrm{d}\rho$$

We usually want to have the boundary term $\frac{1}{2}\rho^2 R'(a)v(a)$ equal to zero. This is equivalent to either letting $R'(a) = 0$ (we prescribe the zero derivative of the radial wave function at a) or we set $v(a) = 0$ (which corresponds to zero Dirichlet condition for R, i.e. setting $R(a) = 0$).

Weak Formulation for u

$$\int \tfrac{1}{2}u'v' + \left(V + \frac{l(l+1)}{2\rho^2}\right)uv\,\mathrm{d}\rho - \tfrac{1}{2}\left[u'v\right]_0^R = \epsilon\int uv\,\mathrm{d}\rho$$

We prescribe $u(0) = u(R) = 0$, so we get:

$$\int \tfrac{1}{2}u'v' + \left(V + \frac{l(l+1)}{2\rho^2}\right)uv\,\mathrm{d}\rho = \epsilon\int uv\,\mathrm{d}\rho$$

Dirac Notation

We can also write all the formulas using the Dirac notation:

$$\mathbb{1} = \int \mathrm{d}\rho\rho^2 \left|\rho\right\rangle \left\langle\rho\right|$$
$$\left\langle\rho|\rho'\right\rangle = \frac{\delta(\rho-\rho')}{\rho^2}$$
$$\left\langle\rho|R\right\rangle = R(\rho)$$
$$\left\langle\rho|\hat{H}|R\right\rangle = \frac{1}{\rho^2}(-\tfrac{1}{2}\rho^2 R')' + (V + \tfrac{1}{2}\frac{l(l+1)}{\rho^2})R$$
$$\hat{H}\left|R\right\rangle = E\left|R\right\rangle$$

Then normalization is:

$$\left\langle R|R\right\rangle = \int \mathrm{d}\rho\rho^2 \left\langle R|\rho\right\rangle \left\langle\rho|R\right\rangle = \int \mathrm{d}\rho\rho^2 R^2(\rho)$$

The operator $\hat{H}$ can be written as:

$$\left\langle\rho|\hat{H}|\rho'\right\rangle = \left\langle\rho|\rho'\right\rangle \left(-\tfrac{1}{2}\frac{1}{\rho^2}\frac{\mathrm{d}}{\mathrm{d}\rho}\left(\rho^2\frac{\mathrm{d}}{\mathrm{d}\rho}\right) + (V + \tfrac{1}{2}\frac{l(l+1)}{\rho^2})\right)$$

so to recover the above formula, we do:

$$\left\langle\rho|\hat{H}|R\right\rangle = \int \mathrm{d}\rho'\rho'^2 \left\langle\rho|\hat{H}|\rho'\right\rangle \left\langle\rho'|R\right\rangle =$$
$$= \int \mathrm{d}\rho'\rho'^2 \frac{\delta(\rho-\rho')}{\rho^2}\left(-\tfrac{1}{2}\frac{1}{\rho^2}\frac{\mathrm{d}}{\mathrm{d}\rho}\left(\rho^2\frac{\mathrm{d}}{\mathrm{d}\rho}\right) + (V + \tfrac{1}{2}\frac{l(l+1)}{\rho^2})\right) R(\rho') = \frac{1}{\rho^2}(-\tfrac{1}{2}\rho^2 R')' + (V + \tfrac{1}{2}\frac{l(l+1)}{\rho^2})R$$

Operator $\hat{H}$ is symmetric, because:

$$\int f\frac{1}{\rho^2}(\rho^2 g')'\rho^2 \mathrm{d}\rho = \int \frac{1}{\rho^2}(\rho^2 f')' g\rho^2 \mathrm{d}\rho$$

The weak formulation is:

$$\left\langle v|H|R\right\rangle = E\left\langle v|R\right\rangle$$
$$\int \mathrm{d}\rho\rho^2 \left\langle v|\rho\right\rangle \left\langle\rho|H|R\right\rangle = E\int \mathrm{d}\rho\rho^2 \left\langle v|\rho\right\rangle \left\langle\rho|R\right\rangle$$
$$\int \mathrm{d}\rho\rho^2 v(\rho)\left(\frac{1}{\rho^2}(-\tfrac{1}{2}\rho^2 R')' + (V + \tfrac{1}{2}\frac{l(l+1)}{\rho^2})R\right) = E\int \mathrm{d}\rho\rho^2 v(\rho)R(\rho)$$

and we obtain the FE formulation by expanding $\left|R\right\rangle = \sum_j R_j \left|j\right\rangle$ (note that the basis $\left|j\right\rangle$ is not orthogonal, so in particular $\sum_j \left|j\right\rangle \left\langle j\right| \neq 1$):

$$\sum_j \left\langle i|H|j\right\rangle R_j = E\sum_j \left\langle i|j\right\rangle R_j$$

This is a generalized eigenvalue problem. In the special case of an orthonormal basis, $\left\langle i|j\right\rangle = \delta_{ij}$ (which FE is not), we get:

$$\sum_j \left\langle i|H|j\right\rangle R_j = R_i$$
$$R_i = \left\langle i|R\right\rangle$$

Which is an eigenvalue problem.

8.8.3 Variational Formulation of the Dirac equation

The QED Lagrangian density is

$$\mathcal{L} = \bar{\psi}(i\hbar c\gamma^\mu D_\mu - mc^2)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

where:

$$D_\mu = \partial_\mu + \frac{i}{\hbar}eA_\mu$$
$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

We will treat the fields as classical fields, so we get the classical wave Dirac equation, after plugging this Lagrangian into the Euler-Lagrange equation of motion:

$$(i\hbar c\gamma^\mu D_\mu - mc^2)\psi = 0$$
$$\partial_\nu F^{\nu\mu} = -ec\bar{\psi}\gamma^\mu\psi$$

Notice that the Lagrangian happens to be zero for the solution of Dirac equation (e.g. the extremum of the action). This has nothing to do with the variational principle itself, it's just a coincindence.

In this section we are only interested in the Dirac equation, so we write the Lagrangian as:

$$\begin{aligned}
\mathcal{L} = \bar{\psi}(i\hbar c\gamma^\mu D_\mu - mc^2)\psi = \\
= \psi^\dagger\gamma^0(i\hbar c\gamma^\mu D_\mu - mc^2)\psi = \\
= \psi^\dagger\gamma^0(i\hbar c\gamma^0(\partial_0 + \frac{i}{\hbar}eA_0) + ic\gamma^i(\partial_i + \frac{i}{\hbar}eA_i) - mc^2)\psi = \\
= \psi^\dagger(i\hbar c\partial_0 + i\hbar c\gamma^0\gamma^i\partial_i - \gamma^0 mc^2 - ceA_0 - ce\gamma^0\gamma^i A_i)\psi = \\
= \psi^\dagger(i\hbar\frac{\partial}{\partial t} + c\alpha^i p_i - \beta mc^2 - ceA_0 - ce\alpha^i A_i)\psi = \\
= -\psi^\dagger(-i\hbar\frac{\partial}{\partial t} + c\alpha^i(-p_i + eA_i) + \beta mc^2 + ceA_0)\psi = \\
= -\psi^\dagger(-i\hbar\frac{\partial}{\partial t} + c\boldsymbol{\alpha}\cdot(\mathbf{p} - e\mathbf{A}) + \beta mc^2 + V)\psi
\end{aligned}$$

where we introduced the potential by $V = ceA_0$. We also could have done the same manipulation to the dirac equation itself and we would get the same expression:

$$(-i\hbar\frac{\partial}{\partial t} + c\boldsymbol{\alpha}\cdot(\mathbf{p} - e\mathbf{A}) + \beta mc^2 + V)\psi = 0$$

The corresponding eigenvalue problem is:

$$(c\boldsymbol{\alpha}\cdot(\mathbf{p} - e\mathbf{A}) + \beta mc^2 + V)\psi = W\psi$$

8.8.4 Radial Dirac equation

As for the Schrödinger equation, there are two ways to obtain the radial Dirac equation. Either from the Lagrangian, or from the equation itself.

From the Equation

The manipulations are well known, one starts by writing the Dirac spinors using the spin angular functions and radial components P and Q:

$$\psi = \begin{pmatrix} \frac{P}{\rho}\chi_\kappa^{j_3} \\ i\frac{Q}{\rho}\chi_{-\kappa}^{j_3} \end{pmatrix}$$
$$\psi^\dagger = \begin{pmatrix} \frac{P}{\rho}\chi_\kappa^{j_3} & -i\frac{Q}{\rho}\chi_{-\kappa}^{j_3} \end{pmatrix}$$

and putting this into the Dirac equation one obtains:

$$\begin{pmatrix} \left(-\hbar c\left(\frac{\mathrm{d}}{\mathrm{d}\rho} - \frac{\kappa}{\rho}\right)Q + (V + mc^2 - W)P\right) & 0 \\ 0 & \left(\hbar c\left(\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho}\right)P + (V - mc^2 - W)Q\right) \end{pmatrix} \begin{pmatrix} \frac{1}{\rho}\chi_\kappa^{j_3} \\ i\frac{1}{\rho}\chi_{-\kappa}^{j_3} \end{pmatrix} = 0$$

So one obtains the following radial equations:

$$-\hbar c\left(\frac{\mathrm{d}}{\mathrm{d}\rho} - \frac{\kappa}{\rho}\right)Q + (V + mc^2 - W)P = 0$$
$$\hbar c\left(\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho}\right)P + (V - mc^2 - W)Q = 0$$

From the Lagrangian

We can reuse the calculations from the previous sections, because the Lagrangian happens to be zero for the solution of the Dirac equation:

$$\mathcal{L} = \bar{\psi}(i\hbar c\gamma^\mu D_\mu - mc^2)\psi =$$
$$= -\psi^\dagger(-i\hbar\frac{\partial}{\partial t} + c\boldsymbol{\alpha}\cdot(\mathbf{p} - e\mathbf{A}) + \beta mc^2 + V)\psi =$$
$$= \begin{pmatrix} \frac{P}{\rho}\chi_\kappa^{j_3} & -i\frac{Q}{\rho}\chi_{-\kappa}^{j_3} \end{pmatrix} \begin{pmatrix} \left(-\hbar c\left(\frac{\mathrm{d}}{\mathrm{d}\rho} - \frac{\kappa}{\rho}\right)Q + (V + mc^2)P\right) & 0 \\ 0 & \left(\hbar c\left(\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho}\right)P + (V - mc^2)Q\right) \end{pmatrix} \begin{pmatrix} \frac{1}{\rho}\chi_\kappa^{j_3} \\ i\frac{1}{\rho}\chi_{-\kappa}^{j_3} \end{pmatrix} =$$
$$= \frac{1}{\rho^2}P\left(-\hbar c\left(\frac{\mathrm{d}}{\mathrm{d}\rho} - \frac{\kappa}{\rho}\right)Q + (V + mc^2)P\right)\chi_\kappa^{j_3}\chi_\kappa^{j_3} + \frac{1}{\rho^2}Q\left(\hbar c\left(\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho}\right)P + (V - mc^2)Q\right)\chi_{-\kappa}^{j_3}\chi_{-\kappa}^{j_3}$$

We can now write the action:

$$S = \int \mathcal{L}\,\rho^2\,\mathrm{d}\rho\mathrm{d}\Omega$$

the spin angular functions integrate to 1:

$$\int \chi_\kappa^{j_3}\chi_\kappa^{j_3}\mathrm{d}\Omega = 1$$
$$\int \chi_{-\kappa}^{j_3}\chi_{-\kappa}^{j_3}\mathrm{d}\Omega = 1$$

the ρ^2 cancels out and we get:

$$S[P,Q] = \int P\left(-\hbar c\left(\frac{\mathrm{d}}{\mathrm{d}\rho} - \frac{\kappa}{\rho}\right)Q + (V + mc^2)P\right) + Q\left(\hbar c\left(\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho}\right)P + (V - mc^2)Q\right)\,\mathrm{d}\rho =$$
$$= \int -\hbar c(PQ' - QP') + \hbar c\frac{2\kappa}{\rho}PQ + V(P^2 + Q^2) + mc^2(P^2 - Q^2)\mathrm{d}\rho$$

the normalization condition is:

$$N = \int P^2 + Q^2 \mathrm{d}\rho - 1 = 0$$

and we can variate the action, we also shift the energy $W = \epsilon + mc^2$:

$$0 = \delta(S - WN) = \delta(S - \epsilon N - mc^2 N)$$

which effectively adds $-mc^2(P^2 + Q^2)$ into the Lagrangian, which changes the term $mc^2(P^2 - Q^2)$ into $-2mc^2Q^2$. We can now variate the (constrained) action:

$$0 = \delta \int -\hbar c(PQ' - QP') + \hbar c\frac{2\kappa}{\rho}PQ + V(P^2 + Q^2) - 2mc^2Q^2 \mathrm{d}\rho =$$

$$= 2\int \left(-\hbar c((\delta P)Q' - P'\delta Q) + \hbar c\frac{\kappa}{\rho}((\delta P)Q + P\delta Q)) + (P\delta P + Q\delta Q)V - 2mc^2 Q\delta Q - \epsilon(P\delta P + Q\delta Q)\right)\mathrm{d}\rho$$

$$+[P\delta Q - Q\delta P]_0^R =$$

$$= 2\int \delta P\left(-\hbar cQ' + \hbar c\frac{\kappa}{\rho}Q + PV - \epsilon P\right) + \delta Q\left(\hbar cP' + \hbar c\frac{\kappa}{\rho}P + QV - 2mc^2Q - \epsilon Q\right)\mathrm{d}\rho + [P\delta Q - Q\delta P]_0^R =$$

which gives the two radial equations:

$$-\hbar cQ' + \hbar c\frac{\kappa}{\rho}Q + PV = \epsilon P$$

$$\hbar cP' + \hbar c\frac{\kappa}{\rho}P + QV - 2mc^2Q = \epsilon Q$$

Weak Formulation

The weak formulation can be obtained by substituting $\delta P \to v_1$ and $\delta Q \to v_2$ into the action above (and separating the integrals) and omitting the the boundary term:

$$\int -\hbar cQ'v_1 + \hbar c\frac{\kappa}{\rho}Qv_1 + PVv_1 \mathrm{d}\rho = \epsilon \int Pv_1 \mathrm{d}\rho$$

$$\int \hbar cP'v_2 + \hbar c\frac{\kappa}{\rho}Pv_2 + QVv_2 - 2mc^2Qv_2 \mathrm{d}\rho = \epsilon \int Qv_2 \mathrm{d}\rho$$

We can also start from the radial equations themselves to get the same result. If we start from the equations themselves (which is the most elementary approach), there are no boundary terms (because we didn't integrate by parts). We can separate the integrals according to the matrix elements that they contribute to:

$$\int PVv_1 \mathrm{d}\rho + \int -\hbar cQ'v_1 + \hbar c\frac{\kappa}{\rho}Qv_1 \mathrm{d}\rho = \epsilon \int Pv_1 \mathrm{d}\rho$$

$$\int \hbar cP'v_2 + \hbar c\frac{\kappa}{\rho}Pv_2 + \int (V - 2mc^2)Qv_2 \mathrm{d}\rho = \epsilon \int Qv_2 \mathrm{d}\rho$$

To show that this problem generates a symmetric matrix, it is helpful to write the radial equations in the following form:

$$\hat{H}\left|P, Q\right\rangle = \epsilon \left|P, Q\right\rangle$$

where:

$$\left|P, Q\right\rangle = \begin{pmatrix} P(\rho) \\ Q(\rho) \end{pmatrix}$$

$$\hat{H} = \begin{pmatrix} V(\rho) & \hbar c\left(-\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho}\right) \\ \hbar c\left(\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho}\right) & V(\rho) - 2mc^2 \end{pmatrix}$$

the operator $\hat{H}$ is Hermitean ($\hat{H}^\dagger = \hat{H}$), because $\left(-\frac{\mathrm{d}}{\mathrm{d}\rho}\right)^\dagger = \frac{\mathrm{d}}{\mathrm{d}\rho}$:

$$\int f \frac{\mathrm{d}}{\mathrm{d}\rho} g \mathrm{d}\rho = \int \left(-\frac{\mathrm{d}}{\mathrm{d}\rho}\right) f g \mathrm{d}\rho$$

and all the other quantities are just scalars.

Stricly speaking, the exact Dirac notation (that is coordinate/representation independent) would be the following (notice the missing ρ^2 in the completeness relation, which is different to the radial Schrödinger equation):

$$\begin{aligned}
\hat{H}\left|P,Q\right\rangle &= \epsilon \left|P,Q\right\rangle \\
\mathbb{1} &= \int \mathrm{d}\rho \left|\rho\right\rangle \left\langle\rho\right| \\
\left\langle\rho|\rho'\right\rangle &= \delta(\rho - \rho') \\
\int \left\langle\rho|\hat{H}|\rho'\right\rangle \left\langle\rho'|P,Q\right\rangle \mathrm{d}\rho' &= \epsilon \left\langle\rho|P,Q\right\rangle \\
\left\langle\rho|P,Q\right\rangle &= \begin{pmatrix} P(\rho) \\ Q(\rho) \end{pmatrix} \\
\left\langle\rho|\hat{H}|\rho'\right\rangle = \delta(\rho - \rho') &\begin{pmatrix} V(\rho) & \hbar c \left(-\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho}\right) \\ \hbar c \left(\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho}\right) & V(\rho) - 2mc^2 \end{pmatrix}
\end{aligned}$$

The normalization is:

$$\left\langle P,Q|P,Q\right\rangle = \int \mathrm{d}\rho \left\langle P,Q|\rho\right\rangle \left\langle\rho|P,Q\right\rangle = \int \mathrm{d}\rho (P^2 + Q^2) = 1$$

The weak formulation is:

$$\left\langle v|\hat{H}|P,Q\right\rangle = \epsilon \left\langle v|P,Q\right\rangle$$

where the test function $|v\rangle$ is one of:

$$|v\rangle = \begin{cases} |v_1\rangle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ |v_2\rangle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{cases}$$

The FE formulation is then obtained by expanding $|P,Q\rangle = \sum_k q_k |k\rangle$:

$$\sum_l \left\langle k|\hat{H}|l\right\rangle q_l = \epsilon \sum_l \left\langle k|l\right\rangle q_l$$

The basis $|k\rangle$ can be for example the FE basis, some spline basis set, or gaussians. The basis has actually $2n$ base functions and it enumerates each equation like this:

$$|k\rangle = \begin{cases} |i\rangle \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \text{for } i = k < n \\ |i\rangle \begin{pmatrix} 0 \\ 1 \end{pmatrix} & \text{for } i = k >= n \end{cases}$$

So at the end of the day, the $\langle k|\hat{H}|l\rangle$ matrix looks like this:

$$\left\langle k|\hat{H}|l\right\rangle = \begin{pmatrix} \langle i|V(r)|j\rangle & \hbar c \left\langle i\right| - \frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho}\left|j\right\rangle \\ \hbar c \left\langle i\right| \frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho}\left|j\right\rangle & \langle i|V(r) - 2mc^2|j\rangle \end{pmatrix}$$

The matrix is $2n \times 2n$, composed of those 4 matrices $n \times n$. The $\langle k|l\rangle$ matrix looks like this:

$$\langle k|l\rangle = \begin{pmatrix} \langle i|j\rangle & 0 \\ 0 & \langle i|j\rangle \end{pmatrix}$$

We can also write the matrix elements explicitly. Let $|i\rangle = B_i(\rho)$, then:

$$\begin{aligned} \langle i|j\rangle &= \int B_i B_j \,\mathrm{d}\rho \\ \langle i|V|j\rangle &= \int B_i V B_j \,\mathrm{d}\rho \\ \langle i|V - 2mc^2|j\rangle &= \int B_i (V - 2mc^2) B_j \,\mathrm{d}\rho \\ \hbar c \,\langle i|\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho}|j\rangle &= \hbar c \int B_i B_j' + B_i \frac{\kappa}{\rho} B_j \,\mathrm{d}\rho \\ \hbar c \,\langle i| - \frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho}|j\rangle &= \hbar c \int -B_i B_j' + B_i \frac{\kappa}{\rho} B_j \,\mathrm{d}\rho \end{aligned}$$

8.8.5 Other Forms of Dirac Equations

The radial Dirac equations are:

$$\begin{aligned} \hat{H}\,|P,Q\rangle &= \epsilon\,|P,Q\rangle \\ |P,Q\rangle &= \begin{pmatrix} P(\rho) \\ Q(\rho) \end{pmatrix} \\ \hat{H} &= \begin{pmatrix} V(\rho) & \hbar c\left(-\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho}\right) \\ \hbar c\left(\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho}\right) & V(\rho) - 2mc^2 \end{pmatrix} \end{aligned}$$

After substitution $S = f(r)P$ and $T = f(r)Q$, we get:

$$\begin{aligned} \hat{H}\,|P,Q\rangle &= \epsilon\,|P,Q\rangle \\ \hat{H}\frac{1}{f}\,|S,T\rangle &= \epsilon\frac{1}{f}\,|S,T\rangle \\ f\hat{H}\frac{1}{f}\,|S,T\rangle &= \epsilon\,|S,T\rangle \end{aligned}$$

where:

$$\begin{aligned} f\hat{H}\frac{1}{f} &= f\begin{pmatrix} V(\rho) & \hbar c\left(-\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho}\right) \\ \hbar c\left(\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho}\right) & V(\rho) - 2mc^2 \end{pmatrix}\frac{1}{f} = \\ &= \begin{pmatrix} V(\rho) & \hbar c\left(-f\frac{\mathrm{d}}{\mathrm{d}\rho}\frac{1}{f} + \frac{\kappa}{\rho}\right) \\ \hbar c\left(f\frac{\mathrm{d}}{\mathrm{d}\rho}\frac{1}{f} + \frac{\kappa}{\rho}\right) & V(\rho) - 2mc^2 \end{pmatrix} \end{aligned}$$

and after using:

$$f\frac{\mathrm{d}}{\mathrm{d}\rho}\frac{1}{f} = \frac{\mathrm{d}}{\mathrm{d}\rho} - \frac{f'}{f}$$

we get:

$$f\hat{H}\frac{1}{f} = \begin{pmatrix} V(\rho) & \hbar c\left(-\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho} + \frac{f'}{f}\right) \\ \hbar c\left(\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho} - \frac{f'}{f}\right) & V(\rho) - 2mc^2 \end{pmatrix}$$

Example I

In order to obtain equations for g and f, related to P and Q by:

$$P = \rho g$$
$$Q = \rho f$$

so $f(r) = \frac{1}{\rho}$ and

$$\frac{f'}{f} = -\frac{1}{\rho}$$

and we get the radial Dirac equation for g and f:

$$f\hat{H}\frac{1}{f} = \begin{pmatrix} V(\rho) & \hbar c\left(-\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa-1}{\rho}\right) \\ \hbar c\left(\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa+1}{\rho}\right) & V(\rho) - 2mc^2 \end{pmatrix}$$

Example II

For $f(r) = \rho$ we get

$$\frac{f'}{f} = \frac{1}{\rho}$$

and

$$f\hat{H}\frac{1}{f} = \begin{pmatrix} V(\rho) & \hbar c\left(-\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa+1}{\rho}\right) \\ \hbar c\left(\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa-1}{\rho}\right) & V(\rho) - 2mc^2 \end{pmatrix}$$

Example III

For $f(r) = \rho^n$ we get

$$\frac{f'}{f} = \frac{n}{\rho}$$

and

$$f\hat{H}\frac{1}{f} = \begin{pmatrix} V(\rho) & \hbar c\left(-\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa+n}{\rho}\right) \\ \hbar c\left(\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa-n}{\rho}\right) & V(\rho) - 2mc^2 \end{pmatrix}$$

Example I is just a special case for $n = -1$, Example II for $n = 1$.

Example IV

For $f(r) = \tanh\rho$ we get

$$\frac{f'}{f} = \frac{1}{\sinh\rho\cosh\rho}$$

and

$$f\hat{H}\frac{1}{f} = \begin{pmatrix} V(\rho) & \hbar c\left(-\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho} + \frac{1}{\sinh\rho\cosh\rho}\right) \\ \hbar c\left(\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho} - \frac{1}{\sinh\rho\cosh\rho}\right) & V(\rho) - 2mc^2 \end{pmatrix}$$

Example V

For $f(r) = \tanh^n k\rho$ we get

$$\frac{f'}{f} = \frac{nk}{\sinh k\rho \cosh k\rho}$$

and

$$f\hat{H}\frac{1}{f} = \begin{pmatrix} V(\rho) & \hbar c\left(-\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho} + \frac{nk}{\sinh k\rho \cosh k\rho}\right) \\ \hbar c\left(\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{\kappa}{\rho} - \frac{nk}{\sinh k\rho \cosh k\rho}\right) & V(\rho) - 2mc^2 \end{pmatrix}$$

Example IV is just a special case for $n = 1$, $k = 1$.

8.8.6 Asymptotic

Schrödinger

The radial Schrödinger equation is:

$$P''(r) + 2\left(E - V(r) - \frac{l(l+1)}{2r^2}\right)P(r) = 0$$
$$Q(r) = P'(r)$$

For $r \to \infty$, assuming $V(r) \to 0$ we get:

$$P''(r) + 2EP(r) = 0$$

And the asymptotic is:

$$P(r) = e^{-\sqrt{-2E}r}$$
$$Q(r) = P'(r) = -\sqrt{-2E}e^{-\sqrt{-2E}r}$$

For $r \to 0$ and assuming that $V(r)$ can be neglected compared to the $\frac{l(l+1)}{2r^2}$ term (for example $V(r) = -Z/r + O(1)$ is ok) we get:

$$P''(r) - \frac{l(l+1)}{r^2}P(r) = 0$$

And the asymptotic is:

$$P(r) = r^{l+1}$$
$$Q(r) = (l+1)r^l$$

From the derivation this is valid for $l > 0$, but it turns out to be valid also for $l = 0$, because for $V(r) = -Z/r + O(1)$ the equations become:

$$P''(r) + \frac{2Z}{r}P(r) = 0$$

the asymptotic of which is:

$$P(r) = r(1 - Zr)$$
$$P'(r) = 1 - 2Zr$$
$$P''(r) = -2Z$$

Which in the first order is just the above asymptotic for $l = 0$:

$$P(r) = r$$
$$Q(r) = 1$$

Note that Z can be both positive and negative.

Dirac

The Dirac equation is:

$$H = \begin{pmatrix} V(r) + c^2 & c\left(-\frac{\mathrm{d}}{\mathrm{d}r} + \frac{\kappa}{r}\right) \\ c\left(\frac{\mathrm{d}}{\mathrm{d}r} + \frac{\kappa}{r}\right) & V(r) - c^2 \end{pmatrix}$$
$$H \begin{pmatrix} P(r) \\ Q(r) \end{pmatrix} = W \begin{pmatrix} P(r) \\ Q(r) \end{pmatrix}$$

Where the relativistic energy $W = E + c^2$. In terms of the nonrelativistic energy it becomes:

$$H_{nonrel} = \begin{pmatrix} V(r) & c\left(-\frac{\mathrm{d}}{\mathrm{d}r} + \frac{\kappa}{r}\right) \\ c\left(\frac{\mathrm{d}}{\mathrm{d}r} + \frac{\kappa}{r}\right) & V(r) - 2c^2 \end{pmatrix}$$
$$H_{nonrel} \begin{pmatrix} P(r) \\ Q(r) \end{pmatrix} = E \begin{pmatrix} P(r) \\ Q(r) \end{pmatrix}$$

For $r \to \infty$, assuming $V(r) \to 0$ we get:

$$H = \begin{pmatrix} c^2 & -c\frac{\mathrm{d}}{\mathrm{d}r} \\ c\frac{\mathrm{d}}{\mathrm{d}r} & -c^2 \end{pmatrix}$$

and in terms of $P(r)$ and $Q(r)$:

$$c^2 P - cQ' = WP$$
$$cP' - c^2 Q = WQ$$

let's put the derivatives on the left hand side:

$$cP' = (W + c^2)Q$$
$$cQ' = -(W - c^2)P$$

write a second order equation:

$$c^2 P'' = (W + c^2)cQ' = -(W + c^2)(W - c^2)P = -(W^2 - c^4)P$$

and finally we get:

$$P'' + \frac{W^2 - c^4}{c^2} P = 0$$
$$Q = \frac{c}{W + c^2} P'$$

The asymptotic is:

$$P(r) = e^{-\sqrt{\frac{c^4 - W^2}{c^2}} r}$$
$$Q(r) = \frac{c}{W + c^2} \left(-\sqrt{\frac{c^4 - W^2}{c^2}}\right) e^{-\sqrt{\frac{c^4 - W^2}{c^2}} r} = -\sqrt{\frac{c^2 - W}{c^2 + W}} e^{-\sqrt{\frac{c^4 - W^2}{c^2}} r}$$

We can also write it in terms of E:

$$P(r) = e^{-\sqrt{-2E - \left(\frac{E}{c}\right)^2} r}$$
$$Q(r) = -\sqrt{-\frac{E}{E + 2c^2}} P(r)$$

For $r \to 0$ we write the full equations:

$$(V + c^2)P - cQ' + c\frac{\kappa}{r}Q = WP$$
$$cP' + c\frac{\kappa}{r}P + (V - c^2)Q = WQ$$

Then we assume $P(r) = r^\beta$ and use the second equation to express $Q(r)$:

$$Q(r) = \frac{cP' + c\frac{\kappa}{r}P}{W - V + c^2} = \frac{c\beta r^{\beta-1} + c\frac{\kappa}{r}r^\beta}{W - V + c^2} = r^{\beta-1}\frac{c(\beta + \kappa)}{W - V + c^2}$$

We can always write any potential as $V(r) = -\frac{Z(r)}{r}$ and we get:

$$Q(r) = r^{\beta-1}\frac{c(\beta + \kappa)}{W + \frac{Z(r)}{r} + c^2} = r^\beta \frac{c(\beta + \kappa)}{Z(r) + (W + c^2)r}$$

If $Z(r) \to Z$ as $r \to 0$ then the term $(W + c^2)r$ goes to zero and we get:

$$Q(r) = r^\beta \frac{c(\beta + \kappa)}{Z}$$

If $Z(r) \to Z_1 r$, then we get:

$$Q(r) = r^\beta \frac{c(\beta + \kappa)}{Z_1 r + (W + c^2)r} = r^{\beta-1}\frac{c(\beta + \kappa)}{Z_1 + W + c^2}$$

If $Z(r) \sim r^3$ (harmonic oscillator) or $Z(r) \sim r^2$, then the $Z(r)$ term goes to zero and we get:

$$Q(r) = r^{\beta-1}\frac{c(\beta + \kappa)}{W + c^2}$$

In order to determine the constant β for $Z \neq 0$, we write the fraction $\frac{Q(r)}{P(r)}$ in two ways:

$$\frac{Q(r)}{P(r)} = \frac{c(\beta + \kappa)}{Z} = \frac{Z}{c(\kappa - \beta)}$$

The second equation follows from first assuming $Q(r) = r^\beta$ and using the first Dirac equation to express $P(r) = r^\beta \frac{c(\kappa-\beta)}{Z}$. Now we can express β (we can assume $\beta > 0$):

$$\frac{Z^2}{c^2} = \kappa^2 - \beta^2$$
$$\beta = \sqrt{\kappa^2 - \left(\frac{Z}{c}\right)^2}$$

8.9 Density Functional Theory (DFT)

8.9.1 Many Body Schrödinger Equation

We use (Hartree) atomic units in this whole section about DFT. We use the Born-Oppenheimer approximation, which says that the nuclei of the treated atoms are seen as fixed. A stationary electronic state (for N electrons) is then described by a wave function $\Psi(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r}_N)$ fulfilling the many-body Schrödinger equation

$$\hat{H}\,|\Psi\rangle = (\hat{T} + \hat{U} + \hat{V})\,|\Psi\rangle = E\,|\Psi\rangle$$

where

$$\hat{T} = \sum_{i}^{N} -\frac{1}{2}\nabla_i^2$$

is the kinetic term,

$$\hat{U} = \sum_{i<j} U(\mathbf{r_i}, \mathbf{r_j}) = \frac{1}{2}\sum_{i,j} U(\mathbf{r_i}, \mathbf{r_j})$$
$$U(\mathbf{r_i}, \mathbf{r_j}) = U(\mathbf{r_j}, \mathbf{r_i}) = \frac{1}{|\mathbf{r_i} - \mathbf{r_j}|}$$

is the electron-electron interaction term and

$$\hat{V} = \sum_{i}^{N} v(\mathbf{r_i})$$
$$v(\mathbf{r_i}) = \sum_{k} -\frac{Z_k}{|\mathbf{r_i} - \mathbf{R_k}|}$$

is the interaction term between electrons and nuclei, where R_k are positions of nuclei and Z_k the number of nucleons in each nucleus (we are using atomic units). So for one atomic calculation with the atom nucleus in the origin, we have just $v(\mathbf{r_i}) = -\frac{Z}{|\mathbf{r_i}|}$.

$|\Psi|^2 = \Psi^*\Psi$ gives the probability density of measuring the first electron at the position $\mathbf{r_1}$, the second at $\mathbf{r_2}$, dots and the Nth electron at the position $\mathbf{r}_N$. The normalization is such that $\int |\Phi|^2 \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 \ldots \mathrm{d}^3 r_N = 1$. The Ψ is antisymmetric, i.e. $\Psi(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r}_N) = -\Psi(\mathbf{r_2}, \mathbf{r_1}, \cdots, \mathbf{r}_N) = -\Psi(\mathbf{r_1}, \mathbf{r}_N, \cdots, \mathbf{r_2})$ etc.

Integrating $|\Psi|^2$ over the first $N-1$ electrons is the probability density that the N-th electron is at the position $\mathbf{r}_N$. Thus the probability density $n(\mathbf{r})$ that any of the N electrons (i.e the first, or the second, or the third, dots, or the N-th) is at the position $\mathbf{r}$ is called the particle (or number) density and is therefore given by:

$$\begin{aligned}
n(\mathbf{r}) = &\int \Psi^*(\mathbf{r}, \mathbf{r}_2, \cdots, \mathbf{r}_N)\Psi(\mathbf{r}, \mathbf{r}_2, \cdots, \mathbf{r}_N)\,\mathrm{d}^3 r_2\,\mathrm{d}^3 r_3 \cdots \mathrm{d}^3 r_N + \\
&+ \int \Psi^*(\mathbf{r}_1, \mathbf{r}, \cdots, \mathbf{r}_N)\Psi(\mathbf{r}_1, \mathbf{r}, \cdots, \mathbf{r}_N)\,\mathrm{d}^3 r_1\,\mathrm{d}^3 r_3 \cdots \mathrm{d}^3 r_N + \cdots \\
&+ \int \Psi^*(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r})\Psi(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r})\,\mathrm{d}^3 r_1\,\mathrm{d}^3 r_2\,\mathrm{d}^3 r_3 \cdots \mathrm{d}^3 r_{N-1} = \\
&= \int (\delta(\mathbf{r} - \mathbf{r}_1) + \delta(\mathbf{r} - \mathbf{r}_2) + \cdots + \delta(\mathbf{r} - \mathbf{r}_N)) \\
&\Psi^*(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)\Psi(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)\,\mathrm{d}^3 r_1\,\mathrm{d}^3 r_2\,\mathrm{d}^3 r_3 \cdots \mathrm{d}^3 r_N = \\
= \sum_{i=1}^{N} &\int \langle\Psi|\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N\rangle\,\delta(\mathbf{r} - \mathbf{r}_i)\,\langle\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N|\Psi\rangle\;\mathrm{d}^3 r_1\,\mathrm{d}^3 r_2\,\mathrm{d}^3 r_3 \cdots \mathrm{d}^3 r_N = \\
= N &\int \langle\Psi|\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N\rangle\,\delta(\mathbf{r} - \mathbf{r}_1)\,\langle\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N|\Psi\rangle\;\mathrm{d}^3 r_1\,\mathrm{d}^3 r_2\,\mathrm{d}^3 r_3 \cdots \mathrm{d}^3 r_N = \\
&= N \int \Psi^*(\mathbf{r}, \mathbf{r}_2, \cdots, \mathbf{r}_N)\Psi(\mathbf{r}, \mathbf{r}_2, \cdots, \mathbf{r}_N)\,\mathrm{d}^3 r_2\,\mathrm{d}^3 r_3 \cdots \mathrm{d}^3 r_N
\end{aligned} \tag{8.9.1.1}$$

Thus $\int_\Omega n(\mathbf{r})\,\mathrm{d}^3 r$ gives the number of particles in the region of integration Ω. Obviously $\int n(\mathbf{r})\,\mathrm{d}^3 r = N$.

Note that the number density $n(\mathbf{r})$ and potential $V(\mathbf{r})$ in the Schroedinger equation is related to the electron charge density $\rho(\mathbf{r})$ and electrostatic potential energy $\phi(\mathbf{r})$ by:

$$\rho(\mathbf{r}) = qn(\mathbf{r})$$
$$q\phi(\mathbf{r}) = V(\mathbf{r})$$

where q is the particle elementary charge, which for electrons is $q = -e = -1$ in atomic units. The amount of electronic charge in the region Ω is given by:

$$Q = \int_\Omega \rho(\mathbf{r})\,\mathrm{d}^3r = q\int_\Omega n(\mathbf{r})\,\mathrm{d}^3r = -\int_\Omega n(\mathbf{r})\,\mathrm{d}^3r$$

The energy of the system is given by

$$E = \langle\Psi|\hat{H}|\Psi\rangle = \langle\Psi|\hat{T}|\Psi\rangle + \langle\Psi|\hat{U}|\Psi\rangle + \langle\Psi|\hat{V}|\Psi\rangle = T + U + V \tag{8.9.1.2}$$

where

$$T = \langle\Psi|\hat{T}|\Psi\rangle = \sum_i^N \int \Psi^*(\mathbf{r_1},\mathbf{r_2},\cdots,\mathbf{r_N})(-\tfrac{1}{2}\nabla_i^2)\Psi(\mathbf{r_1},\mathbf{r_2},\cdots,\mathbf{r_N})\,\mathrm{d}^3r_1\,\mathrm{d}^3r_2\cdots\mathrm{d}^3r_N$$

$$U = \langle\Psi|\hat{U}|\Psi\rangle$$

$$\begin{aligned} V = \langle\Psi|\hat{V}|\Psi\rangle &= \sum_i^N \int \Psi^*(\mathbf{r_1},\mathbf{r_2},\cdots,\mathbf{r_N})v(\mathbf{r_i})\Psi(\mathbf{r_1},\mathbf{r_2},\cdots,\mathbf{r_N})\,\mathrm{d}^3r_1\,\mathrm{d}^3r_2\cdots\mathrm{d}^3r_N = \\ &= \sum_i^N \int \Psi^*(\mathbf{r_1},\mathbf{r_2},\cdots,\mathbf{r_N})v(\mathbf{r_1})\Psi(\mathbf{r_1},\mathbf{r_2},\cdots,\mathbf{r_N})\,\mathrm{d}^3r_1\,\mathrm{d}^3r_2\cdots\mathrm{d}^3r_N = \\ &= N\int \Psi^*(\mathbf{r_1},\mathbf{r_2},\cdots,\mathbf{r_N})v(\mathbf{r_1})\Psi(\mathbf{r_1},\mathbf{r_2},\cdots,\mathbf{r_N})\,\mathrm{d}^3r_1\,\mathrm{d}^3r_2\cdots\mathrm{d}^3r_N = \\ &= \int v(\mathbf{r})n(\mathbf{r})\mathrm{d}^3r = V[n] \end{aligned} \tag{8.9.1.3}$$

It needs to be stressed, that E generally is *not* a functional of n alone, only the $V[n]$ is. In the next section we show however, that if the $|\Psi\rangle$ is a ground state (of any system), then E becomes a functional of n.

8.9.2 The Hohenberg-Kohn Theorem

The Schrödinger equation gives the map

$$C : V \to \Psi$$

where Ψ is the ground state. C is bijective (one-to-one correspondence), because to every V we can compute the corresponding Ψ from Schrödinger equation and two different V and V' (differing by more than a constant) give two different Ψ, because if V and V' gave the same Ψ, then by substracting

$$\hat{H}\,|\Psi\rangle = E_{gs}\,|\Psi\rangle$$

from

$$\hat{H}'\,|\Psi\rangle = (\hat{H} - \hat{V} + \hat{V}')\,|\Psi\rangle = E'_{gs}\,|\Psi\rangle$$

we would get $V - V' = E - E'$, which is a contradiction with the assumption that V and V' differ by more than a constant.

Similarly, from the ground state wavefunction Ψ we can compute the charge density n giving rise to the map

$$D : \Psi \to n$$

which is also bijective, because to every Ψ we can compute n from (8.9.1.1) and two different Ψ and Ψ' give two different n and n', because different Ψ and Ψ' give

$$E_{gs} = \langle\Psi|\hat{H}|\Psi\rangle < \langle\Psi'|\hat{H}|\Psi'\rangle = \langle\Psi'|\hat{H}' + \hat{V} - \hat{V}'|\Psi'\rangle = E'_{gs} + \int n'(\mathbf{r})(v(\mathbf{r}) - v'(\mathbf{r}))\,\mathrm{d}^3r$$

$$E'_{gs} = \langle\Psi'|\hat{H}'|\Psi'\rangle < \langle\Psi|\hat{H}'|\Psi\rangle = \langle\Psi|\hat{H} + \hat{V}' - \hat{V}|\Psi\rangle = E_{gs} + \int n(\mathbf{r})(v'(\mathbf{r}) - v(\mathbf{r}))\,\mathrm{d}^3r$$

adding these two inequalities together gives

$$0 < \int n'(\mathbf{r})(v(\mathbf{r}) - v'(\mathbf{r}))\,\mathrm{d}^3r + \int n(\mathbf{r})(v'(\mathbf{r}) - v(\mathbf{r}))\,\mathrm{d}^3r = \int (n(\mathbf{r}) - n'(\mathbf{r}))(v'(\mathbf{r}) - v(\mathbf{r}))\,\mathrm{d}^3r$$

which for $n = n'$ gives $0 < 0$, which is nonsense, so $n \neq n'$.

So we have proved that for a given ground state density $n_0(\mathbf{r})$ (generated by a potential $\hat{V}_0$) it is possible to calculate the corresponding ground state wavefunction $\Psi_0(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_N})$, in other words, Ψ_0 is a unique functional of n_0:

$$\Psi_0 = \Psi_0[n_0]$$

so the ground state energy E_0 is also a functional of n_0

$$E_0 = \langle \Psi_0[n_0] | \hat{T} + \hat{U} + \hat{V}_0 | \Psi_0[n_0] \rangle = E[n_0]$$

We define an energy functional

$$E_{v_0}[n] = \langle \Psi[n] | \hat{T} + \hat{U} + \hat{V}_0 | \Psi[n] \rangle = \langle \Psi[n] | \hat{T} + \hat{U} | \Psi[n] \rangle + \int v_0(\mathbf{r}) n(\mathbf{r}) \mathrm{d}^3r \tag{8.9.2.1}$$

where $|\Psi[n]\rangle$ is any ground state wavefunction (generated by an arbitrary potential), that is, n is a ground state density belonging to an arbitrary system. E_0 which is generated by the potential V_0 can then be expressed as

$$E_0 = E_{v_0}[n_0]$$

and for $n \neq n_0$ we have (from the Ritz principle)

$$E_0 < E_{v_0}[n]$$

and one has to minimize the functional $E_{v_0}[n]$:

$$E_0 = \min_n E_{v_0}[n] \tag{8.9.2.2}$$

The term

$$\langle \Psi[n] | \hat{T} + \hat{U} | \Psi[n] \rangle \equiv F[n]$$

in (8.9.2.1) is universal in the sense that it doesn't depend on $\hat{V}_0$. It can be proven *[DFT]*, that $F[n]$ is a functional of n for degenerated ground states too, so (8.9.2.2) stays true as well.

The ground state densities in (8.9.2.1) and (8.9.2.2) are called *pure-state v-representable* because they are the densities of (possible degenerate) ground state of the Hamiltonian with some local potential $v(\mathbf{r})$. One may ask a question if all possible functions are v-representable (this is called the v-representability problem). The question is relevant, because we need to know which functions to take into account in the minimization process (8.9.2.2). Even though not every function is v-representable *[DFT]*, every density defined on a grid (finite of infinite) which is strictly positive, normalized and consistent with the Pauli principle is ensemble v-representable. Ensemble v-representation is just a simple generalization of the above, for details see *[DFT]*.

The functional $E_{v_0}[n]$ in (8.9.2.2) depends on the particle number N, so in order to get n, we need to solve the variational formulation

$$\frac{\delta}{\delta n}\left(E_v[n] - \mu(N) \int n(\mathbf{r}) \mathrm{d}^{\mathbf{3}}\mathbf{r} \right) = 0$$

so

$$\frac{\delta E_v[n]}{\delta n} = \mu(N) \tag{8.9.2.3}$$

Let the $n_N(\mathbf{r})$ be the solution of (8.9.2.3) with a particle number N and the energy E_N:

$$E_N = E_v[n_N]$$

The Lagrangian multiplier μ is the exact chemical potential of the system

$$\mu(N) = \frac{\partial E_N}{\partial N}$$

becuase

$$E_{N+\epsilon} - E_N = E_v[n_{N+\epsilon}] - E_v[n_N] = \int \frac{\delta E_v}{\delta n}(n_{N+\epsilon} - n_N)\mathrm{d}^3r =$$
$$= \int \mu(N)(n_{N+\epsilon} - n_N)\mathrm{d}^3r = \mu(N)(N + \epsilon - N) = \mu(N)\epsilon$$

so

$$\mu(N) = \frac{E_{N+\epsilon} - E_N}{\epsilon} \longrightarrow \frac{\partial E_N}{\partial N}$$

8.9.3 The Kohn-Sham Equations

Consider an auxiliary system of N noninteracting electrons (noninteracting gas):

$$\hat{H}_s = \hat{T} + \hat{V}_s$$

the Schrödinger equation then becomes:

$$(-\tfrac{1}{2}\nabla^2 + v_s(\mathbf{r}))\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r})$$
$$n_s(\mathbf{r}) = \sum_i^N |\psi_i(\mathbf{r})|^2$$

and the total energy is:

$$E_s[n] = T_s[\{\psi_i[n]\}] + V_s[n]$$

where

$$T_s[n] = \langle\Psi[n]|\hat{T}|\Psi[n]\rangle = \sum_i \langle\psi_i| - \tfrac{1}{2}\nabla^2|\psi_i\rangle$$
$$V_s[n] = \langle\Psi[n]|\hat{V}|\Psi[n]\rangle = \int v_s(\mathbf{r})n(\mathbf{r})\mathrm{d}^3r$$

So:

$$E_s[n] = \sum_i \langle\psi_i| - \tfrac{1}{2}\nabla^2|\psi_i\rangle + \int v_s(\mathbf{r})n(\mathbf{r})\mathrm{d}^3r =$$
$$= \sum_i \int \psi_i^* \left(-\tfrac{1}{2}\nabla^2\right)\psi_i \,\mathrm{d}^3r + \int v_s(\mathbf{r}) \sum_i \psi_i^*\psi_i \,\mathrm{d}^3r =$$
$$= \sum_i \int \psi_i^* \left(-\tfrac{1}{2}\nabla^2 + v_s(\mathbf{r})\right)\psi_i \,\mathrm{d}^3r =$$
$$= \sum_i \epsilon_i \int \psi_i^*\psi_i \,\mathrm{d}^3r =$$
$$= \sum_i \epsilon_i$$

The total energy is the sum of eigenvalues (energies of the individual independent particles) as expected. From the last equation it follows:

$$T_s[n] = \sum_i \langle\psi_i| - \tfrac{1}{2}\nabla^2|\psi_i\rangle = \sum_i \epsilon_i - \int v_s(\mathbf{r})n(\mathbf{r})\mathrm{d}^3r$$

In other words, the kinetic energy of the noninteracting particles is equal to the sum of eigenvalues minus the potential energy coming from the total effective potential v_s used to construct the single particle orbitals ψ_i.

From (8.9.2.3) we get

$$\mu = \frac{\delta E_s[n]}{\delta n(\mathbf{r})} = \frac{\delta T_s[n]}{\delta n(\mathbf{r})} + \frac{\delta V_s[n]}{\delta n(\mathbf{r})} = \frac{\delta T_s[n]}{\delta n(\mathbf{r})} + v_s(\mathbf{r}) \tag{8.9.3.1}$$

Solution to this equation gives the density n_s.

Now we want to express the energy in (8.9.1.2) using T_s and E_H for convenience, where E_H is the classical electrostatic interaction energy of the charge distribution $\rho(\mathbf{r})$, defined using following relations - we start with a Poisson equation in atomic units

$$\nabla^2\phi_H(\mathbf{r}) = -4\pi\rho(\mathbf{r})$$

and substitute $\rho(\mathbf{r}) = qn(\mathbf{r})$, $V_H(\mathbf{r}) = q\phi_H(\mathbf{r})$ and we use the fact that $q^2 = 1$ in atomic units:

$$\nabla^2 V_H(\mathbf{r}) = -4\pi q^2 n(\mathbf{r}) = -4\pi n(\mathbf{r})$$

or equivalently by expressing V_H using the Green function:

$$V_H(\mathbf{r}) = -\frac{1}{4\pi}\int \frac{-4\pi n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}^3r' = \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}^3r' \tag{8.9.3.2}$$

and finally E_H is related to V_H using:

$$V_H(\mathbf{r}) = \frac{\delta E_H}{\delta n(\mathbf{r})}$$

so we get:

$$E_H[n] = \tfrac{1}{2}\int\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}^3r\mathrm{d}^3r'$$

Using the rules for functional differentiation, we can check that:

$$V_H(\mathbf{r}) = \frac{\delta E_H}{\delta n(\mathbf{r})} = \frac{\delta}{\delta n(\mathbf{r})}\tfrac{1}{2}\int\int \frac{n(\mathbf{r}')n(\mathbf{r}'')}{|\mathbf{r}'-\mathbf{r}''|}\mathrm{d}^3r'\mathrm{d}^3r'' = \\ = \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}^3r'$$

Using the above relations, we can see that

$$E_H[n] = \tfrac{1}{2}\int V_H(\mathbf{r})n(\mathbf{r})\mathrm{d}^3r$$

So from (8.9.2.1) we get

$$E[n] = (T+U)[n] + V[n] = T_s[n] + E_H[n] + (T - T_s + U - E_H)[n] + V[n] = \\ = T_s[n] + E_H[n] + E_{xc}[n] + V[n] \tag{8.9.3.3}$$

The rest of the energy is denoted by $E_{xc} = U - E_H + T - T_s$ and it is called is the exchange and correlation energy functional. From (8.9.2.3)

$$\mu = \frac{\delta E[n]}{\delta n(\mathbf{r})} = \frac{\delta T_s[n]}{\delta n(\mathbf{r})} + \frac{\delta E_H[n]}{\delta n(\mathbf{r})} + \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} + \frac{\delta V[n]}{\delta n(\mathbf{r})}$$

From (8.9.3.2) we have

$$\frac{\delta E_H}{\delta n(\mathbf{r})} = V_H(\mathbf{r})$$

from (8.9.1.3) we get

$$\frac{\delta V[n]}{\delta n(\mathbf{r})} = v(\mathbf{r})$$

we define

$$\frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} = V_{xc}(\mathbf{r}) \tag{8.9.3.4}$$

so we arrive at

$$\mu = \frac{\delta E[n]}{\delta n(\mathbf{r})} = \frac{\delta T_s[n]}{\delta n(\mathbf{r})} + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) + v(\mathbf{r}) \tag{8.9.3.5}$$

Solution to this equation gives the density n. Comparing (8.9.3.5) to (8.9.3.1) we see that if we choose

$$v_s \equiv V_H + V_{xc} + v \tag{8.9.3.6}$$

then $n_s(\mathbf{r}) \equiv n(\mathbf{r})$. So we solve the Kohn-Sham equations of this auxiliary non-interacting system

$$(-\tfrac{1}{2}\nabla^2 + v_s(\mathbf{r}))\psi_i(\mathbf{r}) \equiv (-\tfrac{1}{2}\nabla^2 + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) + v(\mathbf{r}))\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}) \tag{8.9.3.7}$$

which yield the orbitals ψ_i that reproduce the density $n(\mathbf{r})$ of the original interacting system

$$n(\mathbf{r}) \equiv n_s(\mathbf{r}) = \sum_i^N |\psi_i(\mathbf{r})|^2 \tag{8.9.3.8}$$

The sum is taken over the lowest N energies. Some of the ψ_i can be degenerated, but it doesn't matter - the index i counts every eigenfunction including all the degenerated. In plain words, the trick is in realizing, that the ground state energy can be found by minimizing the energy functional (8.9.2.1) and in rewriting this functional into the form (8.9.3.3), which shows that the interacting system can be treated as a noninteracting one with a special potential.

8.9.4 The XC Term

The exchange and correlation functional

$$E_{xc}[n] = (T + U)[n] - E_H[n] - T_S[n]$$

can always be written in the form

$$E_{xc}[n] = \int n(\mathbf{r}')\epsilon_{xc}(\mathbf{r}'; n)\mathrm{d}^3r'$$

where the $\epsilon_{xc}(\mathbf{r}'; n)$ is called the XC energy density. The XC potential is defined as:

$$V_{xc}(\mathbf{r}; n) = \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} = \epsilon_{xc}(\mathbf{r}; n) + \int n(\mathbf{r}')\frac{\delta \epsilon_{xc}(\mathbf{r}'; n)}{\delta n(\mathbf{r})}\mathrm{d}^3r'$$

8.9.5 Total Energy

We already derived all the necessary things above, so we just summarize it here. The total energy is given by:

$$E[n] = (T+U)[n] + V[n] = T_s[n] + E_H[n] + (T - T_s + U - E_H)[n] + V[n] = \\ = T_s[n] + E_H[n] + E_{xc}[n] + V[n]$$

where

$$T_s[n] = \sum_i \epsilon_i - \int v_s(\mathbf{r}) n(\mathbf{r}) \mathrm{d}^3 r$$
$$E_H[n] = \tfrac{1}{2} \int V_H(\mathbf{r}) n(\mathbf{r}) \mathrm{d}^3 r$$
$$E_{xc}[n] = \int \epsilon_{xc}(\mathbf{r}; n) n(\mathbf{r}) \mathrm{d}^3 r$$
$$V[n] = \int v(\mathbf{r}) n(\mathbf{r}) \mathrm{d}^3 r$$

This is the correct, quadratically convergent expression for the total energy. We use the whole input potential $V_{in} \equiv v_s$ and its associated eigenvalues ϵ_i to calculate the kinetic energy $T_s[n]$, this follows from the derivation of the expression for $T_s[n]$. Then we use the calculated charge density to express $E_H[n]$, $E_{xc}[n]$ and $V[n]$.

If one is not careful about the potential associated with the eigenvalues, i.e., confusing V_{in} with V_{out}, one gets a slowly converging formula for the total energy. By expanding v_s using (8.9.3.6):

$$\int v_s n(\mathbf{r}) \mathrm{d}^3 r = \int (V_H + V_{xc} + v) n(\mathbf{r}) \mathrm{d}^3 r = 2\tfrac{1}{2} \int V_H n(\mathbf{r}) \mathrm{d}^3 r + \int V_{xc} n(\mathbf{r}) \mathrm{d}^3 r + \int v n(\mathbf{r}) \mathrm{d}^3 r = \\ = 2E_H[n] + \int V_{xc} n(\mathbf{r}) \mathrm{d}^3 r + V[n]$$

So T_s is equal to:

$$T_s[n] = \sum_i \epsilon_i - \int v_s(\mathbf{r}) n(\mathbf{r}) \mathrm{d}^3 r = \\ = \sum_i \epsilon_i - 2E_H[n] - \int V_{xc} n(\mathbf{r}) \mathrm{d}^3 r - V[n]$$

And then the slowly converging form of total energy is:

$$E[n] = T_s[n] + E_H[n] + E_{xc}[n] + V[n] = \sum_i \epsilon_i - 2E_H[n] - \int V_{xc} n(\mathbf{r}) \mathrm{d}^3 r - V[n] + E_H[n] + E_{xc}[n] + V[n] = \\ = \sum_i \epsilon_i - E_H[n] + E_{xc}[n] - \int V_{xc}(\mathbf{r}; n) n(\mathbf{r}) \mathrm{d}^3 r$$

The reason it is slowly converging is because the new formula for kinetic energy is mixing V_{in} with V_{out}, so it is not as precise (see above) and converges much slower with SCF iterations. Once self-consistency has been achieved (i.e. $V_{in} = V_{out}$), the two expressions for total energy are equivalent.

8.9.6 XC Approximations

All the expressions above are exact (no approximation has been made so far). Unfortunately, no one knows $\epsilon_{xc}(\mathbf{r}'; n)$ exactly (yet). As such, various approximations for it exist.

LDA

The most simple approximation is the *local density approximation* (LDA), for which the xc energy density ϵ_{xc} at $\mathbf{r}$ is taken as that of a homogeneous electron gas (the nuclei are replaced by a uniform positively charged background, density $n = \text{const}$) with the same local density:

$$\epsilon_{xc}(\mathbf{r}; n) \approx \epsilon_{xc}^{LD}(n(\mathbf{r}))$$

The xc potential V_{xc} defined by (8.9.3.4) is then

$$V_{xc}(\mathbf{r}; n) = \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} = \epsilon_{xc}(\mathbf{r}; n) + \int n(\mathbf{r}') \frac{\delta \epsilon_{xc}(\mathbf{r}'; n)}{\delta n(\mathbf{r})} \mathrm{d}^3 r'$$

which in the LDA becomes

$$V_{xc}(\mathbf{r}; n) = \epsilon_{xc}^{LD}(n) + n \frac{\mathrm{d}\epsilon_{xc}^{LD}(n)}{\mathrm{d}n} = \frac{\mathrm{d}}{\mathrm{d}n} \left(n \epsilon_{xc}^{LD}(n) \right) = V_{xc}^{LD}(n) \tag{8.9.6.1}$$

The xc energy density ϵ_{xc}^{LD} of the homogeneous gas can be computed exactly:

$$\epsilon_{xc}^{LD}(n) = \epsilon_{x}^{LD}(n) + \epsilon_{c}^{LD}(n)$$

where the ϵ_x^{LD} is the electron gas exchange term given by

$$\epsilon_{x}^{LD}(n) = -\frac{3}{4\pi} (3\pi^2 n)^{\frac{1}{3}}$$

the rest of ϵ_{xc}^{LD} is hidden in $\epsilon_c^{LD}(n)$ for which there doesn't exist an analytic formula, but the correlation energies are known exactly from quantum Monte Carlo (QMC) calculations by Ceperley and Alder *[pickett]*. The energies were fitted by Vosko, Wilkes and Nussair (VWN) with $\epsilon_c^{LD}(n)$ and they got accurate results with errors less than $0.05\,\mathrm{mRy}$ in ϵ_c^{LD}, which means that $\epsilon_c^{LD}(n)$ is virtually known exactly. VWN result:

$$\epsilon_{c}^{LD}(n) \approx \frac{A}{2} \left\{ \ln\left(\frac{y^2}{Y(y)}\right) + \frac{2b}{Q} \arctan\left(\frac{Q}{2y+b}\right) + \right.$$
$$\left. - \frac{by_0}{Y(y_0)} \left[\ln\left(\frac{(y-y_0)^2}{Y(y)}\right) + \frac{2(b+2y_0)}{Q} \arctan\left(\frac{Q}{2y+b}\right) \right] \right\}$$

where $y = \sqrt{r_s}$, $Y(y) = y^2 + by + c$, $Q = \sqrt{4c - b^2}$, $y_0 = -0.10498$, $b = 3.72744$, $c = 12.9352$ (note that the value of c is wrong in *[pickett]*), $A = 0.0621814$ and r_s is the electron gas parameter, which gives the mean distance between electrons (in atomic units):

$$r_s = \left(\frac{3}{4\pi n}\right)^{\frac{1}{3}}$$

The xc potential is then computed from (8.9.6.1):

$$V_{xc}^{LD} = V_{x}^{LD} + V_{c}^{LD}$$
$$V_{x}^{LD} = -\frac{1}{\pi} (3\pi^2 n)^{\frac{1}{3}} = \frac{4}{3} \epsilon_{x}^{LD}$$
$$V_{c}^{LD} = \frac{A}{2} \left\{ \ln\left(\frac{y^2}{Y(y)}\right) + \frac{2b}{Q} \arctan\left(\frac{Q}{2y+b}\right) + \right.$$
$$\left. - \frac{by_0}{Y(y_0)} \left[\ln\left(\frac{(y-y_0)^2}{Y(y)}\right) + \frac{2(b+2y_0)}{Q} \arctan\left(\frac{Q}{2y+b}\right) \right] \right\} +$$
$$- \frac{A}{6} \frac{c(y-y_0) - by_0 y}{(y-y_0)Y(y)}$$

Some people also use Perdew and Zunger formulas, but they give essentially the same results. The LDA, although very simple, is surprisingly successful. More sophisticated approximations exist, for example the generalized gradient approximation (GGA), which sometimes gives better results than the LDA, but is not perfect either. Other options include orbital-dependent (implicit) density functionals or a linear response type functionals, but this topic is still evolving. The conclusion is, that the LDA is a good approximation to start with, and only when we are not satisfied, we will have to try some more accurate and modern approximation.

RLDA

Relativistic corrections to the energy-density functional (RLDA) were proposed by MacDonald and Vosko:

$$\epsilon_x^{RLD}(n) = \epsilon_x^{LD}(n)R$$
$$R = 1 - \frac{3}{2}\left(\frac{\beta\mu - \ln(\beta+\mu)}{\beta^2}\right)^2 = 1 - \frac{3}{2}A^2$$

where

$$\mu = \sqrt{1+\beta^2}$$
$$\beta = \frac{(3\pi^2 n)^{\frac{1}{3}}}{c} = -\frac{4\pi}{3c}\epsilon_x^{LD}$$
$$A = \frac{\beta\mu - \log(\beta+\mu)}{\beta^2}$$

We now calculate V_x^{RLD}:

$$\begin{aligned} V_x^{RLD} &= \epsilon_x^{LD}R + n\frac{\mathrm{d}\epsilon_x^{LD}R}{\mathrm{d}n} = \\ &= \epsilon_x^{LD}R + n\frac{\mathrm{d}\epsilon_x^{LD}}{\mathrm{d}n}R + n\epsilon_x^{LD}\frac{\mathrm{d}R}{\mathrm{d}n} = \\ &= \epsilon_x^{LD}R + n\frac{\mathrm{d}\epsilon_x^{LD}}{\mathrm{d}n}R + n\epsilon_x^{LD}\frac{\mathrm{d}R}{\mathrm{d}\beta}\frac{\mathrm{d}\beta}{\mathrm{d}n} \end{aligned} \tag{8.9.6.2}$$

where the derivative $\frac{\mathrm{d}\beta}{\mathrm{d}n}$ can be evaluated as follows:

$$\frac{\mathrm{d}\beta}{\mathrm{d}n} = \frac{\mathrm{d}}{\mathrm{d}n}\frac{(3\pi^2 n)^{\frac{1}{3}}}{c} = \frac{1}{3n}\frac{(3\pi^2 n)^{\frac{1}{3}}}{c} = \frac{\beta}{3n}$$

And $\frac{\mathrm{d}\epsilon_x^{LD}}{\mathrm{d}n}$ in exactly the same manner:

$$\frac{\mathrm{d}\epsilon_x^{LD}}{\mathrm{d}n} = ... = \frac{\epsilon_x^{LD}}{3n}$$

So we can write

$$\begin{aligned} V_x^{RLD} &= \epsilon_x^{LD}R + n\frac{\mathrm{d}\epsilon_x^{LD}}{\mathrm{d}n}R + n\epsilon_x^{LD}\frac{\mathrm{d}R}{\mathrm{d}\beta}\frac{\mathrm{d}\beta}{\mathrm{d}n} = \\ &= \epsilon_x^{LD}R + n\frac{\epsilon_x^{LD}}{3n}R + n\epsilon_x^{LD}\frac{\mathrm{d}R}{\mathrm{d}\beta}\frac{\beta}{3n} = \\ &= \frac{4}{3}\epsilon_x^{LD}R + \frac{1}{3}\epsilon_x^{LD}\frac{\mathrm{d}R}{\mathrm{d}\beta}\beta = \\ &= \frac{4}{3}\epsilon_x^{LD}\left(R + \frac{1}{4}\beta\frac{\mathrm{d}R}{\mathrm{d}\beta}\right) = \\ &= V_x^{LD}\left(R + \frac{1}{4}\beta\frac{\mathrm{d}R}{\mathrm{d}\beta}\right) \end{aligned}$$

where

$$\begin{aligned} \frac{\mathrm{d}R}{\mathrm{d}\beta} &= -\frac{3}{2}2AA' = -3AA' = \\ &= -6A\left(\frac{1}{\mu} - \frac{A}{\beta}\right) \end{aligned}$$

where we used the derivative of $A(\beta)$, which after a tedious, but straightforward differentiation is:

$$A'(\beta) = \cdots = 2\left(\frac{1}{\mu} - \frac{A}{\beta}\right)$$

Plugging this back in, we get:

$$\begin{aligned}
V_x^{RLD} &= V_x^{LD}\left(R + \frac{1}{4}\beta\frac{\mathrm{d}R}{\mathrm{d}\beta}\right) = \\
&= V_x^{LD}\left(1 - \frac{3}{2}A^2 + \frac{1}{4}\beta(-6A)\left(\frac{1}{\mu} - \frac{A}{\beta}\right)\right) = \\
&= V_x^{LD}\left(1 - \frac{3}{2}A^2 + \frac{6}{4}A^2 - \frac{6}{4}\beta\frac{A}{\mu}\right) = \\
&= V_x^{LD}\left(1 - \frac{3}{2}\frac{\beta}{\mu}A\right) = \\
&= V_x^{LD}\left(1 - \frac{3}{2}\frac{\beta}{\mu}\left(\frac{\beta\mu - \log(\beta+\mu)}{\beta^2}\right)\right) = \\
&= V_x^{LD}\left(1 - \frac{3}{2}\left(\frac{\beta\mu - \log(\beta+\mu)}{\beta\mu}\right)\right) = \\
&= V_x^{LD}\left(\frac{3\log(\beta+\mu)}{2\beta\mu} - \tfrac{1}{2}\right)
\end{aligned}$$

For $c \to \infty$ we get $\beta \to 0$, $R \to 1$ and $V_x^{RLD} \to \frac{4}{3}\epsilon_x^{LD} = V_x^{LD}$ as expected, because

$$\lim_{\beta\to 0}\frac{\beta\sqrt{1+\beta^2} - \ln(\beta + \sqrt{1+\beta^2})}{\beta^2} = 0$$

Code:

```
>>> from sympy import limit, var, sqrt, log
>>> var("beta")
beta
>>> limit((beta*sqrt(1+beta**2) - log(beta+sqrt(1+beta**2)))/beta**2, beta, 0)
0
```

8.9.7 Radial DFT Problem

Kohn-Sham Equations

For spherically symmetric potentials, we write all eigenfunctions as:

$$\psi_{nlm} = R_{nl}Y_{lm}$$

and we need to solve the following Kohn-Sham equations:

$$-\frac{1}{2}R_{nl}'' - \frac{1}{r}R_{nl}' + \left(V + \frac{l(l+1)R}{2r^2}\right)R_{nl} = \epsilon_{nl}R_{nl}$$

With normalization:

$$\int R_{nl}^2\, r^2\, \mathrm{d}r = 1$$

For Schroedinger equation, the charge density is calculated by adding all "(n, l, m)" states together, counting each one twice (for spin up and spin down):

$$n(\mathbf{r}) = \sum_{nlm} 2|\psi_{nlm}|^2 = \sum_{nlm} R_{nl}^2 2|Y_{lm}|^2 = \sum_{nl} R_{nl}^2 2 \sum_m |Y_{lm}|^2 = \frac{1}{4\pi} \sum_{nl} f_{nl} R_{nl}^2$$

where we have introduced the occupation numbers f_{nl} by

$$f_{nl} = 4\pi\, 2 \sum_m |Y_{lm}|^2$$

Normalization of the charge density is:

$$\begin{aligned} Z = \int n(\mathbf{r}) \mathrm{d}^3 x = \int n(r)\, r^2 \mathrm{d}\Omega \mathrm{d}r = 4\pi \int n(r)\, r^2 \mathrm{d}r = \\ = 4\pi \int \frac{1}{4\pi} \sum_{nl} f_{nl} R_{nl}^2\, r^2 \mathrm{d}\Omega \mathrm{d}r = \\ = \sum_{nl} f_{nl} \int R_{nl}^2\, r^2 \mathrm{d}r = \\ = \sum_{nl} f_{nl} \end{aligned}$$

So we can see, that it must hold:

$$\sum_{nl} f_{nl} = Z$$

where Z is the atomic number (number of electrons), and as such, f_{nl} are indeed the occupation numbers (integers). The factor 4π is explicitly factored out, as it cancels with the spherical harmonics: assuming all m states are occupied, this can be simplified to:

$$f_{nl} = 4\pi\, 2 \sum_m |Y_{lm}|^2 = 4\pi\, 2 \frac{2l+1}{4\pi} = 2(2l+1)$$

We can also use this machinery to prescribe "chemical occupation numbers", that don't necessarily correspond to the DFT ground state. For example for U atom we get:

$$\begin{aligned} f_{1l} &= 2(2l+1) \\ f_{2l} &= 2(2l+1) \\ f_{3l} &= 2(2l+1) \\ f_{4l} &= 2(2l+1) \\ f_{5l} = 2(2l+1) \qquad & l \leq 2 \\ f_{53} &= 3 \\ f_{60} &= 2 \\ f_{61} &= 6 \\ f_{62} &= 1 \\ f_{70} &= 2 \end{aligned}$$

By summing all these f_{nl}, we get 92 as expected:

$$\begin{aligned} \sum_{nl} f_{nl} = 2 + (2+6) + (2+6+10) + (2+6+10+14) + (2+6+10) + \\ +3 + 2 + 6 + 1 + 2 = 92 \end{aligned}$$

But this isn't the DFT ground state, because some KS energies are skipped, for example there is only one state for $n = 6$, $l = 2$, but there are nine more states with the same energy — instead two more states are occupied in $n = 7$, $l = 0$, but those have higher energy. So this corresponds to excited DFT state, strictly speaking not physically valid in the DFT formalism, but in practice this approach is often used. One can also prescribe fractional occupation numbers (in the Dirac case).

Poisson Equation

Poisson equation becomes:

$$V_H''(r) + \frac{2}{r}V_H'(r) = -4\pi n(r)$$

Total Energy

The total energy is given by:

$$E[n] = T_s[n] + E_H[n] + E_{xc}[n] + V[n]$$

where

$$\begin{aligned} T_s[n] &= \sum_{nl} f_{nl}\epsilon_{nl} - \int (V_H(r) + V_{xc}(r) + v(r))_{in} n(r)\mathrm{d}^3r = \\ &= \sum_{nl} f_{nl}\epsilon_{nl} - \int \left(V_H(r) + V_{xc}(r) - \frac{Z}{r}\right)_{in} n(r)\mathrm{d}^3r \\ E_H[n] &= \tfrac{1}{2}\int V_H(r)n(r)\mathrm{d}^3r \\ E_{xc}[n] &= \int \epsilon_{xc}(r;n)n(r)\mathrm{d}^3r \\ V[n] &= \int v(r)n(r)\mathrm{d}^3r = -\int \frac{Z}{r}n(r)\mathrm{d}^3r \end{aligned}$$

doing the integrals a bit we get:

$$\begin{aligned} T_s[n] &= \sum_{nl} f_{nl}\epsilon_{nl} - 4\pi\int \left(V_H(r) + V_{xc}(r) - \frac{Z}{r}\right)_{in} n(r)r^2\,\mathrm{d}r \\ E_H[n] &= 2\pi\int V_H(r)n(r)r^2\,\mathrm{d}r \\ E_{xc}[n] &= 4\pi\int \epsilon_{xc}(r;n)n(r)r^2\,\mathrm{d}r \\ V[n] &= -4\pi\int \frac{Z}{r}n(r)r^2\,\mathrm{d}r = -4\pi Z\int n(r)r\,\mathrm{d}r \end{aligned}$$

We can also express everything using the charge density $\rho(r) = -n(r)$:

$$\begin{aligned} T_s[n] &= \sum_{nl} f_{nl}\epsilon_{nl} + 4\pi\int \left(V_H(r) + V_{xc}(r) - \frac{Z}{r}\right)_{in} \rho(r)r^2\,\mathrm{d}r \\ E_H[n] &= -2\pi\int V_H(r)\rho(r)r^2\,\mathrm{d}r \\ E_{xc}[n] &= -4\pi\int \epsilon_{xc}(r;n)\rho(r)r^2\,\mathrm{d}r \\ V[n] &= 4\pi\int \frac{Z}{r}\rho(r)r^2\,\mathrm{d}r = 4\pi Z\int \rho(r)r\,\mathrm{d}r \end{aligned}$$

8.9.8 DFT As a Nonlinear Problem

The task is to find such a charge density n, so that all the equations below hold (e.g. are self-consistent):

$$\begin{aligned} V &= -\frac{Z}{r} + V_H + V_{xc} \\ \left(-\nabla^2 + V\right)\phi_m &= \epsilon_m \phi_m, \qquad m = 1, 2, \ldots, 4 \\ n &= \sum_{m=1}^{4} \phi_m^2 \\ V_{xc} &= f(n) \\ \nabla^2 V_H &= -4\pi n \end{aligned}$$

This is a standard nonlinear problem, except that the Jacobian is dense, as shown below.

Reformulation

Let's write everything in terms of $\phi_m(x)$ explicitly:

$$\begin{aligned} n(x) &= \sum_{m=1}^{4} \phi_m^2(x) \\ V_{xc}(x) &= f(n(x)) = f\left(\sum_{m=1}^{4} \phi_m^2(x)\right) \\ V_H(x) &= \int_\Omega \frac{n(x')}{|x' - x|}\mathrm{d}x' = \int_\Omega \frac{\sum_{m=1}^{4} \phi_m^2(x')}{|x' - x|}\mathrm{d}x' \\ V(x) &= -\frac{Z}{r} + V_H(x) + V_{xc}(x) = \\ &= -\frac{Z}{r} + \int_\Omega \frac{\sum_{m=1}^{4} \phi_m^2(x')}{|x' - x|}\mathrm{d}x' + f\left(\sum_{m=1}^{4} \phi_m^2(x)\right) \end{aligned}$$

Now we can write everything as just one (nonlinear) equation:

$$\left(-\nabla^2 - \frac{Z}{r} + \int_\Omega \frac{\sum_{m=1}^{4} \phi_m^2(x')}{|x' - x|}\mathrm{d}x' + f\left(\sum_{m=1}^{4} \phi_m^2(x)\right)\right)\phi_n = \epsilon_n \phi_n, \qquad n = 1, 2, \ldots, 4$$

FE Discretization

The correspondig discrete problem has the form

$$\begin{aligned} \int_\Omega \nabla\phi_n(x) \cdot \nabla v_i(x) + \left[-\frac{Z}{r} + \int_\Omega \frac{\sum_{m=1}^{4} \phi_m^2(x')}{|x' - x|}\mathrm{d}x' + f\left(\sum_{m=1}^{4} \phi_m^2(x)\right)\right]\phi_n(x)v_i(x)\mathrm{d}x = \\ = \int_\Omega \epsilon_n \phi_n(x) v_i(x)\mathrm{d}x, \qquad n = 1, 2, \ldots, 4; \quad i = 1, 2, \ldots, N \end{aligned}$$

where

$$\phi_n = \phi_n(\mathbf{Y}^{(n)}) = \sum_{j=1}^{N} y_j^{(n)} v_j(x)$$

Here $\mathbf{Y}^{(n)} = (y_1^{(n)}, y_2^{(n)}, \dots, y_N^{(n)})^T$ is the vector of unknown coefficients for the n-th wavefunction $\phi_n(x)$. Our equation can then be written in the compact form

$$\mathbf{F}(\mathbf{Y}^{(n)}) = \mathbf{0}, \qquad n = 1, 2, \dots, 4$$

where $\mathbf{F} = (F_1, F_2, \dots, F_N)^T$ with

$$F_i(\mathbf{Y}^{(n)}) = \int_\Omega \nabla\phi_n(x) \cdot \nabla v_i(x) + \left[-\frac{Z}{r} + \int_\Omega \frac{\sum_{m=1}^4 \phi_m^2(x')}{|x' - x|} \mathrm{d}x' + f\left(\sum_{m=1}^4 \phi_m^2(x) \right) \right] \phi_n(x) v_i(x) \mathrm{d}x - \\ - \int_\Omega \epsilon_n \phi_n(x) v_i(x) \mathrm{d}x$$

Jacobian

The Jacobi matrix has the elements:

$$J_{ik} = \frac{\partial F_i}{\partial y_k^{(s)}}$$

The only possible dense term is:

$$\frac{\partial}{\partial y_k^{(s)}} \int_\Omega \int_\Omega \frac{\sum_{m=1}^4 \phi_m^2(x')}{|x' - x|} \mathrm{d}x' \, \phi_n(x) v_i(x) \mathrm{d}x = \\ = \frac{\partial}{\partial y_k^{(s)}} \int_\Omega \int_\Omega \frac{\sum_{m=1}^4 \left(\sum_{j=1}^N y_j^{(m)} v_j(x') \right)^2}{|x' - x|} \mathrm{d}x' \left(\sum_{j=1}^N y_j^{(n)} v_j(x) \right) v_i(x) \mathrm{d}x = \\ = \int_\Omega \int_\Omega \frac{2 \left(\sum_{j=1}^N y_j^{(s)} v_j(x') \right) v_k(x')}{|x' - x|} \mathrm{d}x' \left(\sum_{j=1}^N y_j^{(n)} v_j(x) \right) v_i(x) \mathrm{d}x + \\ + \int_\Omega \int_\Omega \frac{\sum_{m=1}^4 \left(\sum_{j=1}^N y_j^{(m)} v_j(x') \right)^2}{|x' - x|} \mathrm{d}x' \, \delta_{ns} v_k(x) v_i(x) \mathrm{d}x$$

Now we can see that we have in there the following term:

$$\int_\Omega \int_\Omega \frac{v_k(x') v_i(x)}{|x' - x|} \mathrm{d}x' \mathrm{d}x$$

which is dense in (ki), as can be easily seen be fixing i and writing

$$\int_\Omega \int_\Omega \frac{v_k(x')}{|x' - x|} \mathrm{d}x' v_i(x) \mathrm{d}x$$

so for each k there is some contribution from the integral $\int_\Omega \frac{v_k(x')}{|x'-x|} \mathrm{d}x'$ for such x where $v_i(x)$ is nonzero, thus making the Jacobian J_{ik} dense.

8.9.9 Thomas-Fermi-Dirac Theory

There are two ways to derive equations for Thomas-Fermi-Dirac theory. One way is to start from grand potential and derive all equations from it. The other way is to start with low level equations and build our way up. Will start with the former.

Top Down Approach

We start with a grand potential for fermions:

$$\begin{aligned}
\Omega[\beta,\mu] = -\sum_i \frac{1}{\beta}\log\left(\sum_{N=0}^{1} e^{-\beta(N\epsilon_i - N\mu)}\right) = \\
= -\sum_i \frac{1}{\beta}\log\left(1+e^{-\beta(\epsilon_i-\mu)}\right) = \\
= -\frac{1}{\beta}\int\int \frac{2\mathrm{d}^3x\mathrm{d}^3p}{(2\pi)^3}\log\left(1+e^{-\beta\left(\frac{p^2}{2}+V(\mathbf{x})-\mu\right)}\right) - E_{ee} - \frac{1}{3}E_{xc} = \\
= -\frac{2}{\beta}\int \mathrm{d}^3x\int_0^\infty \frac{4\pi p^2\mathrm{d}p}{(2\pi)^3}\log\left(1+e^{-\beta\left(\frac{p^2}{2}+V(\mathbf{x})-\mu\right)}\right) - E_{ee} - \frac{1}{3}E_{xc} = \\
= -\frac{1}{\pi^2\beta}\int \mathrm{d}^3x\int_0^\infty p^2\log\left(1+e^{-\beta\left(\frac{p^2}{2}+V(\mathbf{x})-\mu\right)}\right)\mathrm{d}p - E_{ee} - \frac{1}{3}E_{xc} = \\
= -\frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}}\int \mathrm{d}^3x\int_0^\infty \frac{u^{\frac{3}{2}}}{1+e^{u-\beta(\mu-V(\mathbf{x}))}}\mathrm{d}u - E_{ee} - \frac{1}{3}E_{xc} = \\
= -\frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}}\int I_{\frac{3}{2}}\left(\beta\left(\mu-V(\mathbf{x})\right)\right)\mathrm{d}^3x - E_{ee} - \frac{1}{3}E_{xc} = \\
= -\frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}}\int I_{\frac{3}{2}}\left(\beta\left(\mu-V(\mathbf{x})\right)\right)\mathrm{d}^3x - E_{ee} - \frac{1}{3}E_{xc}
\end{aligned}$$

The potential $V(\mathbf{x}) = V_{en}(\mathbf{x}) + V_{ee}(\mathbf{x}) + V_{xc}(\mathbf{x})$ is the total potential that the electrons experience (it contains nuclear, Hartree, and XC terms) and E_{ee} is the Hartree energy:

$$\begin{aligned}
E_{en} &= \int n_e(\mathbf{x})V_{en}(\mathbf{x})\,\mathrm{d}^3x\,, \\
E_{ee} &= \tfrac{1}{2}\int n_e(\mathbf{x})V_{ee}(\mathbf{x})\mathrm{d}^3x\,, \\
E_{xc} &= \frac{3}{4}\int n_e(\mathbf{x})V_{xc}(\mathbf{x})\,\mathrm{d}^3x\,.
\end{aligned}$$

For simplicity, we assume here that V_{xc} only contains the exchange of the homogeneous electron gas. For a general XC functional, the relation is nonlinear and one must simply numerically calculate the XC energy density $e_{xc}(\mathbf{x})$ and calculate the XC energy using:

$$E_{xc} = \int n_e(\mathbf{x})e_{xc}(\mathbf{x})\,\mathrm{d}^3x\,.$$

In our case here, we have $e_{xc} = \frac{3}{4}V_{xc}(\mathbf{x})$, which is only true for the exchange in homogeneous electron gas. Otherwise the relation is nonlinear. In the general case, the correction that must be applied is:

$$\begin{aligned}
E_{en} + E_{ee} + E_{xc} - \int n_e(\mathbf{x})V(\mathbf{x})\,\mathrm{d}^3x = \\
= \int n_e(\mathbf{x})\left((e_{en}-V_{en}) + (e_{ee}-V_{ee}) + (e_{xc}-V_{xc})\right)\mathrm{d}^3x = \\
= \int n_e(\mathbf{x})\left(0 + \left(\tfrac{1}{2}V_{ee} - V_{ee}\right) + \left(\frac{3}{4}V_{xc} - V_{xc}\right)\right)\mathrm{d}^3x = \\
= \int n_e(\mathbf{x})\left(-\tfrac{1}{2}V_{ee} - \frac{1}{4}V_{xc}\right)\mathrm{d}^3x = \\
= -\tfrac{1}{2}\int n_e(\mathbf{x})V_{ee}\mathrm{d}^3x - \frac{1}{3}\frac{3}{4}\int n_e(\mathbf{x})V_{xc}\mathrm{d}^3x = \\
= -E_{ee} - \frac{1}{3}E_{xc}
\end{aligned}$$

The density is a functional derivative with respect to μ:

$$n_e(\mathbf{x}) = -\frac{\delta\Omega[\beta,\mu]}{\delta\mu} = \frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}}\frac{\partial}{\partial\mu}I_{\frac{3}{2}}\left(\beta\left(\mu - V(\mathbf{x})\right)\right) = \frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}}\beta\frac{3}{2}I_{\frac{1}{2}}\left(\beta\left(\mu - V(\mathbf{x})\right)\right) =$$
$$= \frac{\sqrt{2}}{\pi^2\beta^{\frac{3}{2}}}I_{\frac{1}{2}}\left(\beta\left(\mu - V(\mathbf{x})\right)\right)$$

By defining the function $\Phi(n_e(\mathbf{x}))$:

$$\Phi(n_e(\mathbf{x})) = \beta\left(\mu - V(\mathbf{x})\right) = I^{-1}_{\frac{1}{2}}\left(\frac{\pi^2\beta^{\frac{3}{2}}}{\sqrt{2}}n_e(\mathbf{x})\right)$$

we can express the grand potential using n_e as follows:

$$\Omega[\beta, n_e] = -\frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}}\int I_{\frac{3}{2}}(\Phi(n_e(\mathbf{x})))\,\mathrm{d}^3x - \tfrac{1}{2}\int n_e(\mathbf{x})V_{ee}(\mathbf{x})\mathrm{d}^3 - \frac{1}{4}\int n_e(\mathbf{x})V_{xc}(\mathbf{x})\mathrm{d}^3x\,.$$

Now we can calculate the free energy:

$$F_e[\beta, n_e] = \Omega[\beta, n_e] + \mu N = \Omega[\beta, n_e] + \mu\int n_e(\mathbf{x})\,\mathrm{d}^3x =$$
$$= \int\left(-\frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}}I_{\frac{3}{2}}(\Phi(n_e(\mathbf{x}))) + \mu n_e(\mathbf{x}) - n_e(\mathbf{x})\left(\tfrac{1}{2}V_{ee}(\mathbf{x}) + \frac{1}{4}V_{xc}(\mathbf{x})\right)\right)\mathrm{d}^3x =$$
$$= \int\left(-\frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}}I_{\frac{3}{2}}(\Phi(n_e(\mathbf{x}))) + \frac{1}{\beta}n_e(\mathbf{x})\Phi(n_e(\mathbf{x})) + n_e(\mathbf{x})V(\mathbf{x}) - n_e(\mathbf{x})\left(\tfrac{1}{2}V_{ee}(\mathbf{x}) + \frac{1}{4}V_{xc}(\mathbf{x})\right)\right)\mathrm{d}^3x =$$
$$= \int\left(-\frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}}I_{\frac{3}{2}}(\Phi(n_e(\mathbf{x}))) + \frac{1}{\beta}n_e(\mathbf{x})\Phi(n_e(\mathbf{x})) + n_e(\mathbf{x})\left(V_{en}(\mathbf{x}) + \tfrac{1}{2}V_{ee}(\mathbf{x}) + \frac{3}{4}V_{xc}(\mathbf{x})\right)\right)\mathrm{d}^3x\,,$$

where we used the fact that $\mu = \frac{1}{\beta}\Phi(n_e(\mathbf{x})) + V(\mathbf{x})$, i.e. the left hand side μ is a constant, thus the sum of the terms on the right hand side is also constant (even though the individual terms are not).

We can calculate the entropy $S = -\left(\frac{\partial\Omega}{\partial T}\right)_{V,\mu}$ as follows:

$$TS = -T\left(\frac{\partial\Omega}{\partial T}\right)_{V,\mu} =$$
$$= \beta\left(\frac{\partial\Omega}{\partial\beta}\right)_{V,\mu} =$$
$$= \beta\frac{\partial}{\partial\beta}\left(-\frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}}\int I_{\frac{3}{2}}(\Phi(n_e(\mathbf{x})))\,\mathrm{d}^3x - E_{ee} - \frac{1}{3}E_{xc}\right) =$$
$$= \beta\frac{\partial}{\partial\beta}\left(-\frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}}\int I_{\frac{3}{2}}(\Phi(n_e(\mathbf{x})))\,\mathrm{d}^3x\right) =$$
$$= \beta\left(\frac{5}{2}\frac{2\sqrt{2}}{3\pi^2\beta^{\frac{7}{2}}}\int I_{\frac{3}{2}}(\Phi(n_e(\mathbf{x})))\,\mathrm{d}^3x - \frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}}\int\frac{3}{2}I_{\frac{1}{2}}(\Phi(n_e(\mathbf{x})))\frac{\partial\Phi(n_e(\mathbf{x}))}{\partial\beta}\,\mathrm{d}^3x\right) =$$
$$= \beta\left(\frac{5}{2}\frac{2\sqrt{2}}{3\pi^2\beta^{\frac{7}{2}}}\int I_{\frac{3}{2}}(\Phi(n_e(\mathbf{x})))\,\mathrm{d}^3x - \frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}}\int\frac{3}{2}I_{\frac{1}{2}}(\Phi(n_e(\mathbf{x})))(\mu - V(\mathbf{x}))\,\mathrm{d}^3x\right) =$$
$$= \frac{5}{2}\frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}}\int I_{\frac{3}{2}}(\Phi(n_e(\mathbf{x})))\,\mathrm{d}^3x - \int n_e(\mathbf{x})(\mu - V(\mathbf{x}))\,\mathrm{d}^3x =$$
$$= \frac{5}{3}\frac{\sqrt{2}}{\pi^2\beta^{\frac{5}{2}}}\int I_{\frac{3}{2}}(\Phi(n_e(\mathbf{x})))\,\mathrm{d}^3x - \mu N + E_{en} + 2E_{ee} + \frac{4}{3}E_{xc}$$

The total energy is then equal to:

$$E = \Omega + \mu N + TS =$$
$$= \left(-\frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}} \int I_{\frac{3}{2}}(\Phi(n_e(\mathbf{x}))) \, \mathrm{d}^3x - E_{ee} - \frac{1}{3}E_{xc} \right) + \mu N + \frac{5}{3}\frac{\sqrt{2}}{\pi^2\beta^{\frac{5}{2}}} \int I_{\frac{3}{2}}(\Phi(n_e(\mathbf{x}))) \, \mathrm{d}^3x - \mu N + E_{en} + 2E_{ee} + \frac{4}{3}E_{xc} =$$
$$= \frac{\sqrt{2}}{\pi^2\beta^{\frac{5}{2}}} \int I_{\frac{3}{2}}(\Phi(n_e(\mathbf{x}))) \, \mathrm{d}^3x + E_{en} + E_{ee} + E_{xc}$$

From which we can see that the kinetic energy E_{kin} is equal to:

$$E_{kin} = E - (E_{en} + E_{ee} + E_{xc}) =$$
$$= \frac{\sqrt{2}}{\pi^2\beta^{\frac{5}{2}}} \int I_{\frac{3}{2}}(\Phi(n_e(\mathbf{x}))) \, \mathrm{d}^3x$$

The relation between the total energy and free energy can be also written as:

$$E = F + TS = F + \beta \left(\frac{\partial \Omega}{\partial \beta} \right)_{V,\mu} =$$
$$= F + \beta \left(\frac{\partial F}{\partial \beta} \right)_{V,\mu} = \left(\frac{\partial (\beta F)}{\partial \beta} \right)_{V,\mu}$$

But it gives the same result as we obtained above.

To determine the kinetic part of the free energy, we set all potentials equal to zero ($V(\mathbf{x}) = V_{en}(\mathbf{x}) = V_{ee}(\mathbf{x}) = V_{xc}(\mathbf{x}) = 0$) and obtain:

$$F_{kin}[\beta, n_e] = \int \left(-\frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}} I_{\frac{3}{2}}(\Phi(n_e(\mathbf{x}))) + \frac{1}{\beta} n_e(\mathbf{x}) \Phi(n_e(\mathbf{x})) \right) \mathrm{d}^3x \,.$$

If the potentials are zero, then the pressure can be calculated from:

$$P = -\frac{1}{V}\Omega[\beta, n_e] = \frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}V} \int I_{\frac{3}{2}}(\Phi(n_e(\mathbf{x}))) \, \mathrm{d}^3x =$$
$$= \frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}V} \int I_{\frac{3}{2}}(\beta\mu) \, \mathrm{d}^3x = \frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}} I_{\frac{3}{2}}(\beta\mu) \,.$$

If the potentials are not zero, then one can calculate the pressure using:

$$\begin{aligned}
P = -\left(\frac{\partial \Omega}{\partial V}\right)_{\mu,T} = -\left(\frac{\partial F}{\partial V}\right)_{T,N} = \\
= -\frac{\partial}{\partial V}\int f \mathrm{d}^3x = \\
= -\left[f + e_{ee}\right]_b - \int \frac{\partial f}{\partial n_e}\frac{\partial n_e}{\partial V}\mathrm{d}^3x = \\
= -\left[f + e_{ee}\right]_b - \mu\int \frac{\partial n_e}{\partial V}\mathrm{d}^3x = \\
= -\left[f + e_{ee}\right]_b + \mu[n_e]_b = \\
= \left[(-f) - e_{ee} + \mu n_e\right]_b = \\
= \left[\left(\frac{2}{3}e_{kin} + e_{ee} + \frac{1}{3}e_{xc} - \mu n_e\right) - e_{ee} + \mu n_e\right]_b = \\
= \left[\frac{2}{3}e_{kin} + \frac{1}{3}e_{xc}\right]_b = \\
= \frac{1}{3V}\int_b \left(\frac{2}{3}e_{kin} + \frac{1}{3}e_{xc}\right)\mathbf{x}\cdot\mathbf{n}\,\mathrm{d}S = \\
= \frac{1}{3V}\int \left(\frac{2}{3}e_{kin} + \frac{1}{3}e_{xc}\right)\nabla\cdot\mathbf{x}\,\mathrm{d}^3x + \frac{1}{3V}\int \mathbf{x}\cdot\nabla\left(\frac{2}{3}e_{kin} + \frac{1}{3}e_{xc}\right)\mathrm{d}^3x = \\
= \frac{1}{3V}(2E_{kin} + E_{xc}) + \frac{1}{3V}\int \mathbf{x}\cdot\left(-n_e(\mathbf{x})\nabla V(\mathbf{x}) + \nabla\frac{1}{3}e_{xc}\right)\mathrm{d}^3x = \\
= \frac{1}{3V}(2E_{kin} + E_{xc}) + \frac{1}{3V}(E_{en} + E_{ee}) = \\
= \frac{1}{3V}(2E_{kin} + E_{en} + E_{ee} + E_{xc})
\end{aligned}$$

Summary:

$$\begin{aligned}
\Omega &= -\frac{2}{3}E_{kin} - E_{ee} - \frac{1}{3}E_{xc} \\
F_e = \Omega + \mu N &= -\frac{2}{3}E_{kin} - E_{ee} - \frac{1}{3}E_{xc} + \mu N \\
TS &= \frac{5}{3}E_{kin} + E_{en} + 2E_{ee} + \frac{4}{3}E_{xc} - \mu N \\
E = F + TS = \Omega + \mu N + TS &= E_{kin} + E_{en} + E_{ee} + E_{xc}\,,
\end{aligned}$$

where:

$$E_{kin} = \frac{\sqrt{2}}{\pi^2 \beta^{\frac{5}{2}}} \int I_{\frac{3}{2}} \left(\Phi(n_e(\mathbf{x}))\right) \mathrm{d}^3 x$$
$$E_{en} = \int n_e(\mathbf{x}) V_{en}(\mathbf{x}) \, \mathrm{d}^3 x$$
$$E_{ee} = \tfrac{1}{2} \int n_e(\mathbf{x}) V_{ee}(\mathbf{x}) \, \mathrm{d}^3 x$$
$$E_{xc} = \frac{3}{4} \int n_e(\mathbf{x}) V_{xc}(\mathbf{x}) \, \mathrm{d}^3 x$$
$$n_e(\mathbf{x}) = \frac{\sqrt{2}}{\pi^2 \beta^{\frac{3}{2}}} I_{\frac{1}{2}} \left(\beta \left(\mu - V(\mathbf{x})\right)\right)$$
$$\Phi(n_e(\mathbf{x})) = \beta \left(\mu - V(\mathbf{x})\right) = I^{-1}_{\frac{1}{2}} \left(\frac{\pi^2 \beta^{\frac{3}{2}}}{\sqrt{2}} n_e(\mathbf{x}) \right)$$
$$N = \int n_e(\mathbf{x}) \, \mathrm{d}^3 x$$
$$\mu = \frac{1}{\beta} \Phi(n_e(\mathbf{x})) + V(\mathbf{x})$$

and μN is calculated as follows:

$$\mu N = \int \mu n_e(\mathbf{x}) \, \mathrm{d}^3 x =$$
$$= \frac{1}{\beta} \int \Phi(n_e(\mathbf{x})) n_e(\mathbf{x}) \, \mathrm{d}^3 x + \int V(\mathbf{x}) n_e(\mathbf{x}) \, \mathrm{d}^3 x =$$
$$= \frac{1}{\beta} \int \Phi(n_e(\mathbf{x})) n_e(\mathbf{x}) \, \mathrm{d}^3 x + E_{en} + 2E_{ee} + \frac{4}{3} E_{xc} \, .$$

So F_e can also be expressed as:

$$F_e = -\frac{2}{3} E_{kin} - E_{ee} - \frac{1}{3} E_{xc} + \mu N =$$
$$= -\frac{2}{3} E_{kin} + \frac{1}{\beta} \int \Phi(n_e(\mathbf{x})) n_e(\mathbf{x}) \, \mathrm{d}^3 x + E_{en} + E_{ee} + E_{xc} \, .$$

Bottom Up Approach

The other way to derive these equations is to use the following considerations. The number of states in a box of side L is given by:

$$N = \int \frac{\mathrm{d}^3 p}{h^3} 2L^3 = \int \frac{\mathrm{d}^3 p}{(2\pi\hbar)^3} 2L^3 = \int \frac{\mathrm{d}^3 p}{(2\pi)^3} 2L^3 = \int_0^{p_f} \frac{4\pi p^2 \mathrm{d}p}{(2\pi)^3} 2L^3 = \frac{L^3}{\pi^2} \int_0^{p_f} p^2 \mathrm{d}p$$

We use atomic units, so $\hbar = 1$. The electronic particle density is:

$$n_e(\mathbf{x}) = \frac{N}{L^3} = \frac{1}{\pi^2} \int_0^{p_f} p^2 \mathrm{d}p = \frac{p_f^3}{3\pi^2} = \frac{\left[2(E_f - V(\mathbf{x}))\right]^{\frac{3}{2}}}{3\pi^2} \tag{8.9.9.1}$$

where we used the relation for Fermi energy $E_f = \frac{p_f^2}{2} + V(\mathbf{x})$. The potential $V(\mathbf{x})$ is the total potential that the electrons experience (it contains Hartree, nuclear and XC terms). At finite temperature T we need to use the Fermi distribution and this generalizes to:

$$n_e(\mathbf{x}) = \frac{1}{\pi^2} \int_0^{\infty} \frac{p^2 \mathrm{d}p}{e^{\beta(E-\mu)} + 1}$$

Now we use the relation $E = \frac{p^2}{2} + V(\mathbf{x})$ and substitutions $\epsilon = \frac{p^2}{2}$, $y = \beta\epsilon$ to rewrite this using the *Fermi-Dirac Integral*:

$$n_e(\mathbf{x}) = \frac{1}{\pi^2}\int_0^\infty \frac{p^2 \mathrm{d}p}{e^{\beta(E-\mu)}+1} = \frac{1}{\pi^2}\int_0^\infty \frac{p^2 \mathrm{d}p}{e^{\beta(\frac{p^2}{2}+V(\mathbf{x})-\mu)}+1} = \frac{\sqrt{2}}{\pi^2}\int_0^\infty \frac{\sqrt{\epsilon}\mathrm{d}\epsilon}{e^{\beta(\epsilon+V(\mathbf{x})-\mu)}+1} =$$
$$= \frac{\sqrt{2}}{\pi^2\beta^{\frac{3}{2}}}\int_0^\infty \frac{\sqrt{y}\mathrm{d}y}{e^{y-\beta(\mu-V(\mathbf{x}))}+1} = \frac{\sqrt{2}}{\pi^2\beta^{\frac{3}{2}}} I_{\frac{1}{2}}\left(\beta(\mu - V(\mathbf{x}))\right)$$

At low temperature ($T \to 0$) we have $\beta \to \infty$, $I_{\frac{1}{2}}(x) \to \frac{2}{3}x^{\frac{3}{2}}$ and we obtain:

$$n_e(\mathbf{x}) \to \frac{2\sqrt{2}}{3\pi^2\beta^{\frac{3}{2}}}\left(\beta(\mu - V(\mathbf{x}))\right)^{\frac{3}{2}} = \frac{\left[2(\mu - V(\mathbf{x}))\right]^{\frac{3}{2}}}{3\pi^2}$$

Identical with (8.9.9.1). We can see that the chemical potential μ becomes the Fermi energy E_f in the limit $T \to 0$. In the finite-temperature case, μ is determined from the normalization condition for the number of electrons N:

$$N = \int n_e(\mathbf{x})\, d^3x$$

The kinetic energy is

$$E_{kin} = \int \mathrm{d}^3x \int 2\frac{\mathrm{d}^3p}{(2\pi)^3}\frac{p^2}{2}\frac{1}{e^{\beta(E-\mu)}+1} =$$
$$= \int \mathrm{d}^3x \int_0^\infty 2\frac{4\pi p^2 \mathrm{d}p}{(2\pi)^3}\frac{p^2}{2}\frac{1}{e^{\beta(E-\mu)}+1} =$$
$$= \int \mathrm{d}^3x \int_0^\infty 2\frac{4\pi\sqrt{2}\sqrt{\epsilon}\mathrm{d}\epsilon}{(2\pi)^3}\epsilon\frac{1}{e^{\beta(\epsilon+V(\mathbf{x})-\mu)}+1} =$$
$$= \frac{\sqrt{2}}{\pi^2}\int \mathrm{d}^3x \int_0^\infty \frac{\epsilon^{\frac{3}{2}}\mathrm{d}\epsilon}{e^{\beta(\epsilon+V(\mathbf{x})-\mu)}+1} =$$
$$= \frac{\sqrt{2}}{\pi^2\beta^{\frac{5}{2}}}\int \mathrm{d}^3x \int_0^\infty \frac{y^{\frac{3}{2}}\mathrm{d}y}{e^{y-\beta(\mu-V(\mathbf{x}))}+1} =$$
$$= \frac{\sqrt{2}}{\pi^2\beta^{\frac{5}{2}}}\int I_{\frac{3}{2}}\left(\beta(\mu - V(\mathbf{x}))\right)\mathrm{d}^3x$$

From the last formula it can be shown that the kinetic energy is equal to

$$E_{kin} = \frac{3}{2}PV - \frac{1}{2}E_{en} - \frac{1}{2}E_{ee}$$

The potential energy is equal to:

$$E_{pot} = E_{en} + E_{ee}$$

The internal energy E is equal to:

$$E = E_{kin} + E_{pot} = E_{kin} + E_{en} + E_{ee} =$$
$$= \frac{3}{2}PV + \frac{1}{2}E_{en} + \frac{1}{2}E_{ee}$$

The entropy S is equal to:

$$\begin{aligned}
TS &= -\frac{1}{\beta}\sum_i [n_i \log n_i + (1-n_i)\log(1-n_i)] = \\
&= -\frac{1}{\beta}\sum_i \left[n_i \log\left(\frac{n_i}{1-n_i}\right) + \log(1-n_i)\right] = \\
&= \left[\sum_i n_i \epsilon_i\right] + \left[-\sum_i n_i \mu\right] + \left[-\frac{1}{\beta}\sum_i \log(1-n_i)\right] = \\
&= [E_{kin} + E_{en} + 2E_{ee}] + [-\mu N] + \left[\frac{2}{3}E_{kin}\right] = \\
&= \frac{5}{3}E_{kin} + E_{en} + 2E_{ee} - \mu N = \\
&= \frac{5}{2}PV + \frac{1}{6}E_{en} + \frac{7}{6}E_{ee} - \mu N
\end{aligned}$$

where $n_i = \frac{1}{1+e^{\beta(\epsilon_i - \mu)}}$ is the number of states at energy ϵ_i. We used the following calculation expressing one of the sums in terms of the kinetic energy:

$$\begin{aligned}
&-\frac{1}{\beta}\sum_i \log(1-n_i) = \\
&= -\frac{1}{\beta}\int \frac{2\mathrm{d}^3x\mathrm{d}^3p}{(2\pi)^3}\log\frac{e^{\beta(E-\mu)}}{1+e^{\beta(E-\mu)}} = \\
&= -\frac{\sqrt{2}}{\pi^2\beta^{\frac{5}{2}}}\int \mathrm{d}^3x \int_0^\infty \sqrt{y}\,\mathrm{d}y \log\frac{e^{y-\beta(\mu-V(\mathbf{x}))}}{1+e^{y-\beta(\mu-V(\mathbf{x}))}} = \\
&= -\frac{\sqrt{2}}{\pi^2\beta^{\frac{5}{2}}}\int \mathrm{d}^3x \left[-\frac{2}{3}\int_0^\infty \frac{y^{\frac{3}{2}}\mathrm{d}y}{1+e^{y-\beta(\mu-V(\mathbf{x}))}}\right] = \\
&= \frac{2}{3}\frac{\sqrt{2}}{\pi^2\beta^{\frac{5}{2}}}\int I_{\frac{3}{2}}\left(\beta(\mu - V(\mathbf{x}))\right)\mathrm{d}^3x = \\
&= \frac{2}{3}E_{kin}\,,
\end{aligned}$$

where we used $E = \frac{p^2}{2} + V(\mathbf{x})$.

The free energy is equal to:

$$\begin{aligned}
F = E - TS &= -\frac{2}{3}E_{kin} - E_{ee} + \mu N = \\
&= -PV + \frac{1}{3}E_{en} - \frac{2}{3}E_{ee} + \mu N
\end{aligned}$$

The grand potential is equal to:

$$\begin{aligned}
\Omega = F - \mu N &= -\frac{2}{3}E_{kin} - E_{ee} = \\
&= -PV + \frac{1}{3}E_{en} - \frac{2}{3}E_{ee}
\end{aligned}$$

We can now express the free energy functional $F_e[\beta, n_e]$ as a function of the density:

$$\begin{aligned}
F_e[\beta, n_e] = -\frac{2}{3}E_{kin} - E_{ee} + \mu N = \\
= \int \left(-\frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}} I_{\frac{3}{2}}(\Phi(n_e(\mathbf{x}))) - \tfrac{1}{2} n_e(\mathbf{x}) V_{ee}(\mathbf{x}) + \mu n_e(\mathbf{x}) \right) \mathrm{d}^3x = \\
= \int \left(-\frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}} I_{\frac{3}{2}}(\Phi(n_e(\mathbf{x}))) - \tfrac{1}{2} n_e(\mathbf{x}) V_{ee}(\mathbf{x}) + \frac{1}{\beta} n_e(\mathbf{x}) \Phi(n_e(\mathbf{x})) + n_e(\mathbf{x}) V(\mathbf{x}) \right) \mathrm{d}^3x = \\
= \int \left(-\frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}} I_{\frac{3}{2}}(\Phi(n_e(\mathbf{x}))) - \tfrac{1}{2} n_e(\mathbf{x}) V_{ee}(\mathbf{x}) + \frac{1}{\beta} n_e(\mathbf{x}) \Phi(n_e(\mathbf{x})) + n_e(\mathbf{x}) (V_{en}(\mathbf{x}) + V_{ee}(\mathbf{x}) + V_{xc}(\mathbf{x})) \right) \mathrm{d}^3x = \\
= \int \left(-\frac{2\sqrt{2}}{3\pi^2\beta^{\frac{5}{2}}} I_{\frac{3}{2}}(\Phi(n_e(\mathbf{x}))) + \frac{1}{\beta} n_e(\mathbf{x}) \Phi(n_e(\mathbf{x})) + n_e(\mathbf{x}) (V_{en}(\mathbf{x}) + \tfrac{1}{2} V_{ee}(\mathbf{x}) + V_{xc}(\mathbf{x})) \right) \mathrm{d}^3x = \\
= \left(-\frac{2}{3}E_{kin} + \int \frac{1}{\beta} n_e(\mathbf{x}) \Phi(n_e(\mathbf{x})) \, \mathrm{d}^3x \right) + E_{en} + E_{ee} + E_{xc}
\end{aligned}$$

8.9.10 Orbital Free Density Functional Theory

The orbital-free electronic free energy is given by:

$$F_e[n_e] = T_0[n_e] + U_{en}[n_e] + U_{ee}[n_e] + F_{xc}[n_e] \,,$$

where the kinetic energy can be written in a few different equivalent ways as

$$\begin{aligned}
T_0[n_e] = \\
= \frac{\sqrt{2}}{\pi^2\beta^{\frac{5}{2}}} \int \left(I_{\frac{1}{2}} \left(\beta(\mu - V(\mathbf{x})) \right) \beta(\mu - V(\mathbf{x})) - \frac{2}{3} I_{\frac{3}{2}}(\beta(\mu - V(\mathbf{x}))) \right) \mathrm{d}^3x = \\
= \frac{1}{\beta} \int \left(n_e(\mathbf{x}) \beta(\mu - V(\mathbf{x})) - \frac{2\sqrt{2}}{3\pi^2\beta^{\frac{3}{2}}} I_{\frac{3}{2}}(\beta(\mu - V(\mathbf{x}))) \right) \mathrm{d}^3x = \\
= \frac{1}{\beta} \int \left(n_e(\mathbf{x}) \Phi(n_e(\mathbf{x})) - \frac{2\sqrt{2}}{3\pi^2\beta^{\frac{3}{2}}} I_{\frac{3}{2}}(\Phi(n_e(\mathbf{x}))) \right) \mathrm{d}^3x = \\
= \frac{1}{\beta} \int \left(n_e(\mathbf{x}) I^{-1}_{\frac{1}{2}}(y) - \frac{2n_e(\mathbf{x})}{3y} I_{\frac{3}{2}}(I^{-1}_{\frac{1}{2}}(y)) \right) \mathrm{d}^3x = \\
= \frac{1}{\beta} \int n_e(\mathbf{x}) \left(I^{-1}_{\frac{1}{2}}(y) - \frac{2}{3y} I_{\frac{3}{2}}(I^{-1}_{\frac{1}{2}}(y)) \right) \mathrm{d}^3x = \\
= \frac{1}{\beta} \int n_e(\mathbf{x}) f(y) \mathrm{d}^3x \,; \qquad y = \frac{\pi^2}{\sqrt{2}} \beta^{\frac{3}{2}} n_e \,,
\end{aligned}$$

where $f(y)$ is a special function of one variable, composed of a *Fermi-Dirac Integral* of order $\frac{3}{2}$ and its inverse of order $\frac{1}{2}$:

$$f(y) = I^{-1}_{\frac{1}{2}}(y) - \frac{2}{3y} I_{\frac{3}{2}}(I^{-1}_{\frac{1}{2}}(y))$$

the electron-nuclei term has the form

$$U_{en}[n_e] = \int \frac{n_e(\mathbf{x}) n_n(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \mathrm{d}^3x \mathrm{d}^3x' = \int n_e(\mathbf{x}) V_{en}(\mathbf{x}) \mathrm{d}^3x \,,$$

The electron-electron (Hartree) term takes the form:

$$U_{ee}[n_e] = \tfrac{1}{2}\int \frac{n_e(\mathbf{x})n_e(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\mathrm{d}^3x\mathrm{d}^3x' = \tfrac{1}{2}\int n_e(\mathbf{x})V_{ee}(\mathbf{x})\mathrm{d}^3x\,,$$

and the exchange and correlation functional $F_{xc}[n_e]$ is given by the Perdew-Zunger LDA:

$$F_{xc}[n_e] = \int n_e(\mathbf{x})\epsilon_{xc}^{LD}(n_e)\mathrm{d}^3x\,,$$

$n_e(\mathbf{x})$ is the (positive) electron density, $n_n(\mathbf{x})$ is the (positive) nuclei density.

We minimize this free energy under the condition of particle conservation. The constrained functional is (we use $n \equiv n_e$ from now on):

$$\Omega[n] = F_e[n] - \epsilon\left(\int n(\mathbf{x})\mathrm{d}^3x - N\right)$$

The variational solution is:

$$\frac{\delta\Omega[n]}{\delta n} = 0$$

Or:

$$\frac{\delta F_e[n]}{\delta n} = \epsilon \tag{8.9.10.1}$$

Finally we get:

$$H[n] \equiv \frac{\delta F_e[n]}{\delta n} = \frac{\delta T_0[n]}{\delta n} + \frac{\delta U_{en}[n]}{\delta n} + \frac{\delta U_{ee}[n]}{\delta n} + \frac{\delta F_{xc}[n]}{\delta n} = \epsilon \tag{8.9.10.2}$$

The individual terms are:

$$\frac{\delta T_0[n]}{\delta n} = \frac{1}{\beta}\left(f(y) + n(\mathbf{x})\frac{\mathrm{d}f(y)}{\mathrm{d}y}\frac{\pi^2}{\sqrt{2}}\beta^{\frac{3}{2}}\right)$$

and

$$\frac{\delta U_{en}[n]}{\delta n} = \int \frac{n_n(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\mathrm{d}^3x' \equiv V_{en}(\mathbf{x})$$

and

$$\frac{\delta U_{ee}[n]}{\delta n} = \int \frac{n(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\mathrm{d}^3x' \equiv V_{ee}(\mathbf{x})$$

and

$$\frac{\delta F_{xc}[n]}{\delta n} = \epsilon_{xc}^{LD}(n) + n(\mathbf{x})\frac{\mathrm{d}\epsilon_{xc}^{LD}(n)}{\mathrm{d}n} \equiv V_{xc}(\mathbf{x})$$

All together the Hamiltonian is:

$$H[n] = \frac{1}{\beta}\left(f(y) + n(\mathbf{x})\frac{\mathrm{d}f(y)}{\mathrm{d}y}\frac{\pi^2}{\sqrt{2}}\beta^{\frac{3}{2}}\right) + V_{en}(\mathbf{x}) + V_{ee}(\mathbf{x}) + V_{xc}(\mathbf{x})$$

We can also introduce an artificial orbital $\psi(\mathbf{x})$ as follows:

$$n(\mathbf{x}) = \psi^2(\mathbf{x})$$

and minimize Ω with respect to ψ:

$$\frac{\delta\Omega[n]}{\delta\psi} = 0\,.$$

We will use tilde to denote functions in terms of ψ. So $\Omega[n] = \Omega[\psi^2] = \tilde{\Omega}[\psi]$. Using the relation

$$\frac{\delta}{\delta\psi} = \frac{\mathrm{d}n}{\mathrm{d}\psi}\frac{\delta}{\delta n} = 2\psi\frac{\delta}{\delta n}$$

we obtain:

$$\frac{\delta\Omega[n]}{\delta\psi} = 2\psi\frac{\delta\Omega[n]}{\delta n} = 0\,.$$

So the equation (8.9.10.1) gets multiplied by 2ψ:

$$2\psi\frac{\delta F_e[n]}{\delta n} = 2\epsilon\psi$$

as well as the equation (8.9.10.2):

$$\tilde{H}[\psi] \equiv \frac{\delta\tilde{F}_e[\psi]}{\delta\psi} = \frac{\delta F_e[n]}{\delta\psi} = 2\psi\frac{\delta F_e[n]}{\delta n} = 2H[n]\psi = 2\epsilon\psi$$

So the Hamiltonian $H[n]$ expressed using n and the Hamiltonian $\tilde{H}[\psi]$ expressed using ψ are related by $\tilde{H}[\psi] = 2H[n]\psi$.

Free Energy Minimization

For clarity, we will be using $H[n]$ from equation (8.9.10.2) as our main quantity, but we will also write the final relations using $\tilde{H}[\psi]$ for completeness.

We start with some initial guess for $|\psi\rangle$ (it must be normalized as $\langle\psi|\psi\rangle = N$). Let's calculate ϵ:

$$\begin{aligned} H[n] &= \epsilon \\ H[n]\,|\psi\rangle &= \epsilon\,|\psi\rangle \\ \langle\psi|H[n]|\psi\rangle &= \epsilon\,\langle\psi|\psi\rangle \\ \langle\psi|H[n]|\psi\rangle &= \epsilon N \\ \epsilon = \frac{1}{N}\,\langle\psi|H[n]|\psi\rangle \equiv \frac{1}{N}\int H[n]\psi^2(\mathbf{x})\mathrm{d}^3x &= \frac{1}{2N}\int \tilde{H}[\psi]\psi(\mathbf{x})\mathrm{d}^3x \end{aligned}$$

We calculate the steepest-descent (SD) vector $|\chi\rangle$:

$$|\chi\rangle = 2(\epsilon - H[n])\,|\psi\rangle \equiv 2\epsilon\,|\psi\rangle - |\tilde{H}[\psi]\rangle$$

The conjugate-gradient (CG) vector $|\varphi\rangle$ is calculated as:

$$|\varphi\rangle = |\chi\rangle + \frac{\langle\chi|\chi\rangle}{\langle\chi_{k-1}|\chi_{k-1}\rangle}\,|\varphi\rangle$$

To satisfy the normalization constraint of $|\psi\rangle$, the CG vector is further orthogonalized to $|\psi\rangle$ and normalized to N (this step is one particular, but not the only way to impose the normalization constraint):

$$|\varphi'\rangle = \left(1 - \frac{1}{N}\,|\psi\rangle\,\langle\psi|\right)|\varphi\rangle$$

$$|\varphi''\rangle = \sqrt{\frac{N}{\langle\varphi'|\varphi'\rangle}}\,|\varphi'\rangle$$

That is, now $\langle\varphi''|\psi\rangle = 0$ and $\langle\varphi''|\varphi''\rangle = N$. The new CG vector $|\psi_{k+1}\rangle$ is then updated as usual in CG by $|\psi\rangle + \alpha\,|\varphi''\rangle$, but then it must be normalized. As such, equivalently, it is updated by a linear combination of $|\psi\rangle$ and $|\varphi''\rangle$:

$$|\psi_{k+1}\rangle = a\,|\psi\rangle + b\,|\varphi''\rangle$$

such that it remains normalized:

$$\langle\psi_{k+1}|\psi_{k+1}\rangle = (a\,\langle\psi| + b\,\langle\varphi''|)(a\,|\psi\rangle + b\,|\varphi''\rangle) = (a^2+b^2)N = N$$

So a, b are any real numbers satisfying the equation $a^2 + b^2 = 1$, whose parametric solution is $a = \cos\theta$, $b = \sin\theta$ with $0 \le \theta < 2\pi$:

$$|\psi_{k+1}\rangle = \cos\theta\,|\psi\rangle + \sin\theta\,|\varphi''\rangle$$

where θ is determined by minimizing the free energy $F_e[\psi_{k+1}]$ as a function of θ.

8.9.11 References

8.10 Hartree-Fock (HF) Method

8.10.1 Derivation

The interacting Hamiltonian for many body Schrödinger equation is (see the general QFT notes for derivation):

$$i\hbar\partial_t\,|\Psi(t)\rangle = \hat{H}\,|\Psi(t)\rangle$$
$$\hat{H} = \hat{T} + \hat{V} = \sum_{ij} c_i^\dagger\,\langle i|T|j\rangle\,c_j + \tfrac{1}{2}\sum_{ijkl} c_i^\dagger c_j^\dagger\,\langle ij|V|kl\rangle\,c_l c_k$$

where $|i\rangle$ are spin orbitals (thus the integration over ω below) and:

$$\langle i|T|j\rangle = \int \chi_i^*(\mathbf{x})\left(-\tfrac{1}{2}\nabla^2 - \sum_n \frac{Z_n}{|\mathbf{x}-\mathbf{R}_n|}\right)\chi_j(\mathbf{x})\mathrm{d}^3x\,\mathrm{d}\omega$$
$$\langle ij|V|kl\rangle = \int \chi_i^*(\mathbf{x})\chi_j^*(\mathbf{y})\frac{1}{|\mathbf{x}-\mathbf{y}|}\chi_k(\mathbf{x})\chi_l(\mathbf{y})\mathrm{d}^3x\,\mathrm{d}\omega_x\,\mathrm{d}^3y\,\mathrm{d}\omega_y$$

We would like to minimize the energy $E = \langle\Psi|\hat{H}|\Psi\rangle$ using the following basis for Z electrons:

$$|\Psi\rangle = c_1^\dagger c_2^\dagger\cdots c_Z^\dagger\,|0\rangle$$

We express the energy E in this basis:

$$\begin{aligned}
E = \langle\Psi|\hat{H}|\Psi\rangle = \\
= \langle 0|\,c_Z\cdots c_2 c_1 \hat{H} c_1^\dagger c_2^\dagger\cdots c_Z^\dagger\,|0\rangle = \\
= \langle 0|\,c_Z\cdots c_2 c_1\left(\sum_{i,j=1}^{Z} c_i^\dagger\,\langle i|T|j\rangle\,c_j + \tfrac{1}{2}\sum_{i,j,k,l=1}^{Z} c_i^\dagger c_j^\dagger\,\langle ij|V|kl\rangle\,c_l c_k\right) c_1^\dagger c_2^\dagger\cdots c_Z^\dagger\,|0\rangle = \\
= \sum_{i=1}^{Z}\langle i|T|i\rangle + \tfrac{1}{2}\sum_{i,j=1}^{Z}\left(\langle ij|V|ij\rangle - \langle ij|V|ji\rangle\right)
\end{aligned}$$

We minimize it with the constrain $\langle i|j\rangle = \delta_{ij}$:

$$\delta\left(E - \sum_{i,j=1}^{Z}\epsilon_{ij}\,\langle i|j\rangle\right) = 0$$

We obtain:

$$T\left|i\right\rangle + \sum_{j=1}^{Z}\left(\left\langle j|V|ij\right\rangle - \left\langle j|V|ji\right\rangle\right) = \epsilon_i\left|i\right\rangle \tag{8.10.1.1}$$

in the x-representation:

$$\left\langle \mathbf{x}|T|i\right\rangle + \sum_{j=1}^{Z}\left(\left\langle \mathbf{x}\right|\left\langle j|V|ij\right\rangle - \left\langle \mathbf{x}\right|\left\langle j|V|ji\right\rangle\right) = \epsilon_i\left\langle \mathbf{x}|i\right\rangle$$

$$\left\langle \mathbf{x}|T|i\right\rangle + \sum_{j=1}^{Z}\left(\left\langle j\mathbf{x}|V|ij\right\rangle - \left\langle j\mathbf{x}|V|ji\right\rangle\right) = \epsilon_i\left\langle \mathbf{x}|i\right\rangle$$

And writing the individual terms explicitly (in this section, all orbitals are *spin* orbitals):

$$\left\langle \mathbf{x}|i\right\rangle = \psi_i(\mathbf{x})$$

$$\left\langle \mathbf{x}|T|i\right\rangle = \left(-\tfrac{1}{2}\nabla^2 - \sum_n \frac{Z_n}{|\mathbf{x}-\mathbf{R}_n|}\right)\psi_i(\mathbf{x})$$

$$\left\langle j\mathbf{x}|V|ij\right\rangle = \int \psi_j^*(\mathbf{y})\frac{1}{|\mathbf{x}-\mathbf{y}|}\psi_i(\mathbf{x})\psi_j(\mathbf{y})\mathrm{d}^3y = \int \frac{|\psi_j(\mathbf{y})|^2}{|\mathbf{x}-\mathbf{y}|}\mathrm{d}^3y\ \psi_i(\mathbf{x})$$

$$\left\langle j\mathbf{x}|V|ji\right\rangle = \int \psi_j^*(\mathbf{y})\frac{1}{|\mathbf{x}-\mathbf{y}|}\psi_j(\mathbf{x})\psi_i(\mathbf{y})\mathrm{d}^3y = \int \frac{\psi_i(\mathbf{y})\psi_j^*(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\mathrm{d}^3y\ \psi_j(\mathbf{x})$$

we get the Hartree-Fock equations:

$$\left(-\tfrac{1}{2}\nabla^2 - \frac{Z}{|\mathbf{x}|} + \int \frac{\sum_{j=1}^{Z}|\psi_j(\mathbf{y})|^2}{|\mathbf{x}-\mathbf{y}|}\mathrm{d}^3y\right)\psi_i(\mathbf{x}) - \sum_{j=1}^{Z}\int \frac{\psi_i(\mathbf{y})\psi_j^*(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\mathrm{d}^3y\ \psi_j(\mathbf{x}) = \epsilon_i\psi_i(\mathbf{x}) \tag{8.10.1.2}$$

Let's introduce the number density $n(\mathbf{x})$, Hartree potential $V_H(\mathbf{x})$ and nonlocal exchange potential V_x with its kernel $U(\mathbf{x},\mathbf{y})$:

$$n(\mathbf{x}) = \sum_{j=1}^{Z}|\psi_j(\mathbf{y})|^2$$

$$V_H(\mathbf{x}) = \int \frac{\sum_{j=1}^{Z}|\psi_j(\mathbf{y})|^2}{|\mathbf{x}-\mathbf{y}|}\mathrm{d}^3y = \int \frac{n(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\mathrm{d}^3y$$

$$\hat{V}_x f(\mathbf{x}) = -\sum_{j=1}^{Z}\int \frac{f(\mathbf{y})\psi_j^*(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\mathrm{d}^3y\ \psi_j(\mathbf{x}) = \int U(\mathbf{x},\mathbf{y})f(\mathbf{y})\mathrm{d}^3y$$

$$U(\mathbf{x},\mathbf{y}) = -\sum_{j=1}^{Z}\frac{\psi_j(\mathbf{x})\psi_j^*(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}$$

then we can write the HF equations as:

$$\left(-\tfrac{1}{2}\nabla^2 - \frac{Z}{|\mathbf{x}|} + V_H(\mathbf{x}) + \hat{V}_x\right)\psi_i(\mathbf{x}) = \epsilon_i\psi_i(\mathbf{x})$$

$$\left(-\tfrac{1}{2}\nabla^2 - \frac{Z}{|\mathbf{x}|} + V_H(\mathbf{x})\right)\psi_i(\mathbf{x}) + \int U(\mathbf{x},\mathbf{y})\psi_i(\mathbf{y})\mathrm{d}^3y = \epsilon_i\psi_i(\mathbf{x})$$

The Hartree potential can be calculated by solving the Poisson equation:

$$\nabla^2 V_H(\mathbf{x}) = -4\pi n(\mathbf{x})$$

where:

$$n(\mathbf{x}) = \sum_{i=1}^{Z} |\psi_i(\mathbf{x})|^2$$

The application of the exchange potential $\hat{V}_x$ on any function $f(\mathbf{x})$ can be calculated by:

$$\hat{V}_x f(\mathbf{x}) = -\sum_{j=1}^{Z} W_{fj}(\mathbf{x})\psi_j(\mathbf{x})$$
$$W_{fj}(\mathbf{x}) = \int \frac{f(\mathbf{y})\psi_j^*(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\mathrm{d}^3y$$
$$\nabla^2 W_{fj}(\mathbf{x}) = -4\pi f(\mathbf{x})\psi_j^*(\mathbf{x})$$

8.10.2 Roothaan Equations For Closed Shell Systems

Starting from (8.10.1.1) and integrating over spins we get (here i, k are spatial orbitals, not spin orbitals):

$$T\,|i\rangle + \sum_{k=1}^{N/2} \left(2\,\langle k|V|ik\rangle - \langle k|V|ki\rangle\right) = \epsilon_i\,|i\rangle \tag{8.10.2.1}$$

We introduce basis functions $|\mu\rangle$ by (below the greek letters are basis functions, latin letters are spatial orbitals):

$$|i\rangle = \sum_{\nu} C_{\nu i}\,|\nu\rangle$$

substitute into (8.10.2.1) and multiply by $\langle\mu|$ from the left:

$$\sum_{\nu} \langle\mu|T|\nu\rangle\, C_{\nu i} + \sum_{\nu}\sum_{k=1}^{N/2} \left(2\,\langle\mu k|V|\nu k\rangle - \langle\mu k|V|k\nu\rangle\right) C_{\nu i} = \epsilon_i \sum_{\nu} \langle\mu|\nu\rangle\, C_{\nu i} \tag{8.10.2.2}$$

Now we expand the functions $|k\rangle$:

$$\sum_{\nu} \langle\mu|T|\nu\rangle\, C_{\nu i} + \sum_{\nu}\sum_{\alpha\beta} \left(2\sum_{k=1}^{N/2} C_{\alpha k} C_{\beta k}^*\right) \left(\langle\mu\beta|V|\nu\alpha\rangle - \tfrac{1}{2}\,\langle\mu\beta|V|\alpha\nu\rangle\right) C_{\nu i} = \epsilon_i \sum_{\nu} \langle\mu|\nu\rangle\, C_{\nu i} \tag{8.10.2.3}$$

we introduce the density matrix:

$$\hat{\rho} = 2\sum_{k=1}^{N/2} |k\rangle\,\langle k| = \sum_{\alpha\beta} |\alpha\rangle\, 2\sum_{k=1}^{N/2} C_{\alpha k} C_{\beta k}^*\,\langle\beta| = \sum_{\alpha\beta} |\alpha\rangle\, P_{\alpha\beta}\,\langle\beta|$$
$$P_{\alpha\beta} = 2\sum_{k=1}^{N/2} C_{\alpha k} C_{\beta k}^*$$

and get:

$$\sum_{\nu} \left(\langle\mu|T|\nu\rangle + \sum_{\alpha\beta} P_{\alpha\beta}\left(\langle\mu\beta|V|\nu\alpha\rangle - \tfrac{1}{2}\,\langle\mu\beta|V|\alpha\nu\rangle\right)\right) C_{\nu i} = \epsilon_i \sum_{\nu} \langle\mu|\nu\rangle\, C_{\nu i} \tag{8.10.2.4}$$

introducing:

$$\begin{aligned} F_{\mu\nu} &= H_{\mu\nu}^{\text{core}} + G_{\mu\nu} \\ H_{\mu\nu}^{\text{core}} &= \langle \mu | T | \nu \rangle \\ G_{\mu\nu} = \sum_{\alpha\beta} P_{\alpha\beta} \left(\langle \mu\beta | V | \nu\alpha \rangle - \tfrac{1}{2} \langle \mu\beta | V | \alpha\nu \rangle \right) \\ S_{\mu\nu} &= \langle \mu | \nu \rangle \end{aligned}$$

the equation (8.10.2.4) is:

$$\sum_{\nu} F_{\mu\nu} C_{\nu i} = \epsilon_i \sum_{\nu} S_{\mu\nu} C_{\nu i} \tag{8.10.2.5}$$

These are the Roothaan equations. It is a generalized eigenvalue problem.

Total energy is given by (the i, j in the first equation are spin orbitals, in the other equations i, j are spatial orbitals):

$$\begin{aligned} E &= \sum_i \langle i | T | i \rangle + \tfrac{1}{2} \sum_{i,j} \left(\langle ij | V | ij \rangle - \langle ij | V | ji \rangle \right) = \\ &= \sum_i 2I(i) + \sum_{i,j} \left(2J(i,j) - K(i,j) \right) = \\ &= \sum_i 2 \langle i | T | i \rangle + \sum_{i,j} \left(2 \langle ij | V | ij \rangle - \langle ij | V | ji \rangle \right) = \\ &= \sum_i 2 \langle i | T | i \rangle + 2 \sum_{i,j} \left(\langle ij | V | ij \rangle - \tfrac{1}{2} \langle ij | V | ji \rangle \right) = \\ &= \sum_{\mu\nu} \sum_i 2 \langle \mu | T | \nu \rangle C_{\nu i} C_{\mu i}^* + 2 \sum_{\mu\nu} \sum_{\alpha\beta} \sum_{i,j} C_{\nu i} C_{\mu i}^* C_{\alpha j} C_{\beta j}^* \left(\langle \mu\beta | V | \nu\alpha \rangle - \tfrac{1}{2} \langle \mu\beta | V | \alpha\nu \rangle \right) = \\ &= \sum_{\mu\nu} \langle \mu | T | \nu \rangle P_{\nu\mu} + \tfrac{1}{2} \sum_{\mu\nu} \sum_{\alpha\beta} P_{\nu\mu} P_{\alpha\beta} \left(\langle \mu\beta | V | \nu\alpha \rangle - \tfrac{1}{2} \langle \mu\beta | V | \alpha\nu \rangle \right) = \\ &= \sum_{\mu\nu} P_{\nu\mu} (H_{\mu\nu}^{\text{core}} + \tfrac{1}{2} G_{\mu\nu}) = \\ &= \sum_{\mu\nu} P_{\nu\mu} (\tfrac{1}{2} H_{\mu\nu}^{\text{core}} + \tfrac{1}{2} (H_{\mu\nu}^{\text{core}} + G_{\mu\nu})) = \\ &= \tfrac{1}{2} \sum_{\mu\nu} P_{\nu\mu} (H_{\mu\nu}^{\text{core}} + F_{\mu\nu}) \end{aligned}$$

The same thing can be derived in x-representation starting from (8.10.1.2) and introducing spatial orbitals:

$$\left(-\tfrac{1}{2} \nabla^2 - \frac{Z}{|\mathbf{x}|} + \int \frac{2 \sum_{k=1}^{N/2} |\psi_k(\mathbf{y})|^2}{|\mathbf{x} - \mathbf{y}|} \mathrm{d}^3 y \right) \psi_i(\mathbf{x}) - \sum_{k=1}^{N/2} \int \frac{\psi_i(\mathbf{y}) \psi_k^*(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|} \mathrm{d}^3 y \ \psi_k(\mathbf{x}) = \epsilon_i \psi_i(\mathbf{x}) \tag{8.10.2.6}$$

We introduce basis functions ϕ_μ:

$$\psi_i(\mathbf{x}) = \sum_{\nu} C_{\nu i} \phi_\nu(\mathbf{x})$$

substitute into (8.10.2.6) and also multiply the whole equation by ϕ_μ^* and integrate over $\mathbf{x}$:

$$\begin{aligned} &\sum_{\nu} \int \phi_\mu^*(\mathbf{x}) \left(-\tfrac{1}{2} \nabla^2 - \frac{Z}{|\mathbf{x}|} + \int \frac{2 \sum_{k=1}^{N/2} |\psi_k(\mathbf{y})|^2}{|\mathbf{x} - \mathbf{y}|} \mathrm{d}^3 y \right) \phi_\nu(\mathbf{x}) \mathrm{d}^3 x \, C_{\nu i} \\ &- \sum_{\nu} \int \phi_\mu^*(\mathbf{x}) \sum_{k=1}^{N/2} \int \frac{\phi_\nu(\mathbf{y}) \psi_k^*(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|} \mathrm{d}^3 y \ \psi_k(\mathbf{x}) \mathrm{d}^3 x \, C_{\nu i} = \epsilon_i \sum_{\nu} \int \phi_\mu^*(\mathbf{x}) \phi_\nu(\mathbf{x}) \mathrm{d}^3 x \, C_{\nu i} \end{aligned} \tag{8.10.2.7}$$

This can be written as:

$$\sum_\nu F_{\mu\nu} C_{\nu i} = \epsilon_i \sum_\nu S_{\mu\nu} C_{\nu i}$$

$$F_{\mu\nu} = H_{\mu\nu}^{\text{core}} + G_{\mu\nu} = T_{\mu\nu} + V_{\mu\nu} + G_{\mu\nu}$$

where:

$$T_{\mu\nu} = \int \phi_\mu^*(\mathbf{x}) \left(-\tfrac{1}{2}\nabla^2\right) \phi_\nu(\mathbf{x}) \mathrm{d}^3x = \tfrac{1}{2} \int \nabla\phi_\mu^*(\mathbf{x}) \cdot \nabla\phi_\nu(\mathbf{x}) \mathrm{d}^3x$$

$$V_{\mu\nu} = \int \phi_\mu^*(\mathbf{x}) \left(-\frac{Z}{|\mathbf{x}|}\right) \phi_\nu(\mathbf{x}) \mathrm{d}^3x$$

$$G_{\mu\nu} = \int \phi_\mu^*(\mathbf{x}) \left(\int \frac{2\sum_{k=1}^{N/2} |\psi_k(\mathbf{y})|^2}{|\mathbf{x}-\mathbf{y}|} \mathrm{d}^3y \right) \phi_\nu(\mathbf{x}) \mathrm{d}^3x - \int \phi_\mu^*(\mathbf{x}) \sum_{k=1}^{N/2} \int \frac{\phi_\nu(\mathbf{y}) \psi_k^*(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} \mathrm{d}^3y \ \psi_k(\mathbf{x}) \mathrm{d}^3x$$

$$S_{\mu\nu} = \int \phi_\mu^*(\mathbf{x}) \phi_\nu(\mathbf{x}) \mathrm{d}^3x$$

Introducing the density matrix and density:

$$\rho(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x}|\hat{\rho}|\mathbf{y}\rangle = \sum_{\alpha\beta} \langle \mathbf{x}|\alpha\rangle \, P_{\alpha\beta} \, \langle \beta|\mathbf{y}\rangle = \sum_{\alpha\beta} \phi_\alpha(\mathbf{x}) P_{\alpha\beta} \phi_\beta^*(\mathbf{y})$$

$$P_{\alpha\beta} = 2 \sum_{k=1}^{N/2} C_{\alpha k} C_{\beta k}^*$$

$$\rho(\mathbf{x}) = 2\sum_{k=1}^{N/2} |\psi_k(\mathbf{x})|^2 = 2\sum_{k=1}^{N/2} |\, \langle \mathbf{x}|k\rangle \,|^2 = 2\sum_{k=1}^{N/2} \langle \mathbf{x}|k\rangle \, \langle k|\mathbf{x}\rangle = \langle \mathbf{x}|\hat{\rho}|\mathbf{x}\rangle = \sum_{\alpha\beta} \phi_\alpha(\mathbf{x}) P_{\alpha\beta} \phi_\beta^*(\mathbf{x})$$

Expanding the ψ_k functions and using the density matrix we get for $G_{\mu\nu}$:

$$G_{\mu\nu} = \sum_{\alpha\beta} P_{\alpha\beta} \int \phi_\mu^*(\mathbf{x}) \left(\int \frac{\phi_\beta^*(\mathbf{y}) \phi_\alpha(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} \mathrm{d}^3y \right) \phi_\nu(\mathbf{x}) \mathrm{d}^3x - \tfrac{1}{2} \sum_{\alpha\beta} P_{\alpha\beta} \int \phi_\mu^*(\mathbf{x}) \int \frac{\phi_\nu(\mathbf{y}) \phi_\beta^*(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} \mathrm{d}^3y \ \phi_\alpha(\mathbf{x}) \mathrm{d}^3x$$

or

$$G_{\mu\nu} = \sum_{\alpha\beta} P_{\alpha\beta} \int \frac{\phi_\mu^*(\mathbf{x}) \phi_\nu(\mathbf{x}) \phi_\beta^*(\mathbf{y}) \phi_\alpha(\mathbf{y}) - \frac{1}{2} \phi_\mu^*(\mathbf{x}) \phi_\alpha(\mathbf{x}) \phi_\beta^*(\mathbf{y}) \phi_\nu(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} \mathrm{d}^3x \, \mathrm{d}^3y$$

$$\equiv \sum_{\alpha\beta} P_{\alpha\beta} \left(\langle \mu\beta| \frac{1}{r_{12}} |\nu\alpha\rangle - \tfrac{1}{2} \langle \mu\beta| \frac{1}{r_{12}} |\alpha\nu\rangle \right)$$

In physical and chemistry notation this is written as:

$$G_{\mu\nu} = \sum_{\alpha\beta} P_{\alpha\beta} \left(\langle \mu\beta|\nu\alpha\rangle - \tfrac{1}{2} \langle \mu\beta|\alpha\nu\rangle \right) = \sum_{\alpha\beta} P_{\alpha\beta} \left((\mu\nu|\beta\alpha) - \tfrac{1}{2} (\mu\alpha|\beta\nu) \right)$$

Note that this notation implicitly assumes the $\frac{1}{r_{12}}$ factor, so for example $\langle \mu\beta|\nu\alpha\rangle$ actually means $\langle \mu\beta|\frac{1}{r_{12}}|\nu\alpha\rangle$ and one has to understand this from the context.

8.10.3 Two Particle Matrix Element

The two particle matrix element is:

$$\begin{aligned} (ij|kl) &= \int \frac{\psi_i^*(\mathbf{x}) \psi_j(\mathbf{x}) \psi_k^*(\mathbf{x}') \psi_l(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|} \mathrm{d}^3x \mathrm{d}^3x' = \\ &= \langle ik|jl\rangle = \int \frac{\psi_i^*(\mathbf{x}) \psi_k^*(\mathbf{x}') \psi_j(\mathbf{x}) \psi_l(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|} \mathrm{d}^3x \mathrm{d}^3x' \end{aligned} \tag{8.10.3.1}$$

The $\langle ik|jl\rangle$ is called the physicists' notation because the $|jl\rangle$ and $|ik\rangle$ kets are:

$$|jl\rangle = \psi_j(\mathbf{x})\psi_l(\mathbf{x}')$$
$$|ik\rangle = \psi_i(\mathbf{x})\psi_k(\mathbf{x}')$$

The $(ij|kl)$ is called the chemists' notation. From (8.10.3.1) there are two types of symmetries — interchanging of the dummy variables:

$$(ij|kl) = \langle ik|jl\rangle = \int \frac{\psi_i^*(\mathbf{x})\psi_j(\mathbf{x})\psi_k^*(\mathbf{x}')\psi_l(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\mathrm{d}^3x\mathrm{d}^3x' =$$
$$= \int \frac{\psi_i^*(\mathbf{x}')\psi_j(\mathbf{x}')\psi_k^*(\mathbf{x})\psi_l(\mathbf{x})}{|\mathbf{x}'-\mathbf{x}|}\mathrm{d}^3x'\mathrm{d}^3x = (kl|ij) = \langle ki|lj\rangle$$

and taking complex conjugate:

$$(ij|kl)^* = \langle ik|jl\rangle^* = \left(\int \frac{\psi_i^*(\mathbf{x})\psi_k^*(\mathbf{x}')\psi_j(\mathbf{x})\psi_l(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\mathrm{d}^3x\mathrm{d}^3x'\right)^* =$$
$$= \int \frac{\psi_i(\mathbf{x})\psi_k(\mathbf{x}')\psi_j^*(\mathbf{x})\psi_l^*(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\mathrm{d}^3x\mathrm{d}^3x' = \langle jl|ik\rangle = (ji|lk)$$

If the matrix elements are real, then:

$$(ij|kl) = \langle ik|jl\rangle = \langle jl|ik\rangle = (ji|lk)$$

In general those are the only symmetries (4 total).

If however, the $\psi_i(\mathbf{x})$ functions are real, then there are additional symmetries: an exchange $i \leftrightarrow j$ and $k \leftrightarrow l$. The symmetries of $(ij|kl)$ are exchange of i with j or k with l or the ij and kl pairs (8 total):

$$(ij|kl) = (ji|kl) = (ij|lk) = (ji|lk) =$$
$$= (kl|ij) = (lk|ij) = (kl|ji) = (lk|ji)$$

So if we view $(ij|kl)$ as two boxes $(\cdot|\cdot)$ then we can permute the labels in the given box "$\cdot$", as well as exchange the boxes (the only thing we cannot do is to take one particle from one box and put it into the other). As such the box "$\cdot$" is a pair of two electrons (in any order) and the two electron integral assigns a unique number to a pair of such boxes (in any order). The symmetries of the $\langle ik|jl\rangle$ symbol are:

$$\langle ik|jl\rangle = \langle jk|il\rangle = \langle il|jk\rangle = \langle jl|ik\rangle =$$
$$= \langle ki|lj\rangle = \langle li|kj\rangle = \langle kj|li\rangle = \langle lj|ki\rangle$$

Example I: the Slater integral $R^k(i,j,k,l)$ has all 8 symmetries (Slater integral uses physical notation).

Example II: In spherical symmetry, the 2-particle integrals can be written as (see below for derivation):

$$\langle \alpha\beta|\gamma\delta\rangle = \langle n_\alpha l_\alpha m_\alpha\ n_\beta l_\beta m_\beta | n_\gamma l_\gamma m_\gamma\ n_\delta l_\delta m_\delta\rangle = (n_\alpha l_\alpha m_\alpha\ n_\gamma l_\gamma m_\gamma | n_\beta l_\beta m_\beta\ n_\delta l_\delta m_\delta) =$$
$$= \sum_{k=\max(|l_\alpha - l_\gamma|, |l_\beta - l_\delta|, |m_\alpha - m_\gamma|)}^{\min(l_\alpha + l_\gamma, l_\beta + l_\delta)} c^k(l_\alpha, m_\alpha, l_\gamma, m_\gamma) c^k(l_\delta, m_\delta, l_\beta, m_\beta)$$
$$\delta_{m_\alpha + m_\beta - m_\gamma - m_\delta, 0}$$
$$R^k(n_\alpha l_\alpha, n_\beta l_\beta, n_\gamma l_\gamma, n_\delta l_\delta)$$

They only have 4 symmetries, because spherical harmonics are complex. In particular:

$$\langle\beta\alpha|\delta\gamma\rangle = \sum_{k=\max(|l_\beta-l_\delta|,|l_\alpha-l_\gamma|,|m_\beta-m_\delta|)}^{\min(l_\beta+l_\delta,l_\alpha+l_\gamma)} c^k(l_\beta,m_\beta,l_\delta,m_\delta)c^k(l_\gamma,m_\gamma,l_\alpha,m_\alpha)$$
$$\delta_{m_\beta+m_\alpha-m_\delta-m_\gamma,0}$$
$$R^k(n_\beta l_\beta,n_\alpha l_\alpha,n_\delta l_\delta,n_\gamma l_\gamma) =$$
$$= \sum_{k=\max(|l_\beta-l_\delta|,|l_\alpha-l_\gamma|,|m_\beta-m_\delta|)}^{\min(l_\beta+l_\delta,l_\alpha+l_\gamma)} c^k(l_\alpha,m_\alpha,l_\gamma,m_\gamma)c^k(l_\delta,m_\delta,l_\beta,m_\beta)(-1)^{m_\alpha-m_\gamma+m_\beta-m_\delta}$$
$$\delta_{m_\beta+m_\alpha-m_\delta-m_\gamma,0}$$
$$R^k(n_\alpha l_\alpha,n_\beta l_\beta,n_\gamma l_\gamma,n_\delta l_\delta) = \langle\alpha\beta|\gamma\delta\rangle$$

and

$$\langle\gamma\delta|\alpha\beta\rangle = \langle n_\gamma l_\gamma m_\gamma\ n_\delta l_\delta m_\delta|n_\alpha l_\alpha m_\alpha\ n_\beta l_\beta m_\beta\rangle = (n_\gamma l_\gamma m_\gamma\ n_\alpha l_\alpha m_\alpha|n_\delta l_\delta m_\delta\ n_\beta l_\beta m_\beta) =$$
$$= \sum_{k=\max(|l_\gamma-l_\alpha|,|l_\delta-l_\beta|,|m_\gamma-m_\alpha|)}^{\min(l_\gamma+l_\alpha,l_\delta+l_\beta)} c^k(l_\gamma,m_\gamma,l_\alpha,m_\alpha)c^k(l_\beta,m_\beta,l_\delta,m_\delta)$$
$$\delta_{m_\gamma+m_\delta-m_\alpha-m_\beta,0}$$
$$R^k(n_\gamma l_\gamma,n_\delta l_\delta,n_\alpha l_\alpha,n_\beta l_\beta)$$
$$= \langle\alpha\beta|\gamma\delta\rangle$$

We used the symmetries of the Slater integrals as well as the c^k coefficients that change a sign, but thanks to the $\delta_{m_\gamma+m_\delta-m_\alpha-m_\beta,0}$, the overall sign does not change. The other two symmetries are missing, i.e. $\langle\gamma\beta|\alpha\delta\rangle \neq \langle\alpha\beta|\gamma\delta\rangle$ and $\langle\alpha\delta|\gamma\beta\rangle \neq \langle\alpha\beta|\gamma\delta\rangle$.

8.10.4 General Matrix Elements in Spherical Symmetry

Spherical symmetry is this particular choice of a basis:

$$\phi_\mu(\mathbf{x}) = \phi_{n_\mu l_\mu m_\mu}(\mathbf{x}) = \frac{\phi_{n_\mu l_\mu}(r)}{r} Y_{l_\mu m_\mu}(\Omega)$$

It can be shown that the solutions are of the form:

$$\psi_i(\mathbf{x}) = \psi_{nlm}(\mathbf{x}) = \frac{P_{nl}(r)}{r} Y_{lm}(\Omega)$$

We can now write:

$$\psi_i(\mathbf{x}) = \sum_\nu C_{\nu i}\phi_\nu(\mathbf{x})$$

$$\int \phi_\mu(\mathbf{x})\psi_i(\mathbf{x})\mathrm{d}^3x = \sum_\nu S_{\mu\nu}C_{\nu i}$$

$$\sum_\mu S^{-1}_{\mu\nu}\int \phi_\mu(\mathbf{x})\psi_i(\mathbf{x})\mathrm{d}^3x = C_{\nu i}$$

$$\sum_\mu \delta_{l_\mu l_\nu}\delta_{m_\mu m_\nu}\left(S^{l_\mu}_{n_\mu n_\nu}\right)^{-1}\int \phi_{n_\mu l_\mu m_\mu}(\mathbf{x})\psi_{nlm}(\mathbf{x})\mathrm{d}^3x = C_{n_\nu l_\nu m_\nu;nlm}$$

$$\sum_\mu \delta_{l_\mu l_\nu}\delta_{m_\mu m_\nu}\left(S^{l_\mu}_{n_\mu n_\nu}\right)^{-1}\int \frac{\phi_{n_\mu l_\mu}(r)}{r}Y_{l_\mu m_\mu}(\Omega)\frac{P_{nl}(r)}{r}Y_{lm}(\Omega)r^2\mathrm{d}r\mathrm{d}\Omega = C_{n_\nu l_\nu m_\nu;nlm}$$

$$\sum_\mu \delta_{l_\mu l_\nu}\delta_{m_\mu m_\nu}\left(S^{l_\mu}_{n_\mu n_\nu}\right)^{-1}\int \phi_{n_\mu l_\mu}(r)P_{nl}(r)\mathrm{d}r\delta_{ll_\mu}\delta_{mm_\mu} = C_{n_\nu l_\nu m_\nu;nlm}$$

$$\delta_{ll_\nu}\delta_{mm_\nu}\sum_{n_\mu}\left(S^{l}_{n_\mu n_\nu}\right)^{-1}\int \phi_{n_\mu l}(r)P_{nl}(r)\mathrm{d}r = C_{n_\nu l_\nu m_\nu;nlm}$$

$$\delta_{ll_\nu}\delta_{mm_\nu}C^l_{n_\nu n} = C_{n_\nu l_\nu m_\nu;nlm}$$

where

$$C^l_{n_\nu n} = \sum_{n_\mu}\left(S^l_{n_\mu n_\nu}\right)^{-1}\int \phi_{n_\mu l}(r)P_{nl}(r)\mathrm{d}r$$

Also we get:

$$\psi_{nlm}(\mathbf{x}) = \sum_\nu C_{n_\nu l_\nu m_\nu;nlm}\phi_{n_\nu l_\nu m_\nu}(\mathbf{x})$$

$$\frac{P_{nl}(r)}{r}Y_{lm}(\Omega) = \sum_\nu C_{n_\nu l_\nu m_\nu;nlm}\frac{\phi_{n_\nu l_\nu}(r)}{r}Y_{l_\nu m_\nu}(\Omega) =$$

$$= \sum_\nu \delta_{ll_\nu}\delta_{mm_\nu}C^l_{n_\nu n}\frac{\phi_{n_\nu l_\nu}(r)}{r}Y_{l_\nu m_\nu}(\Omega) =$$

$$= \sum_{n_\nu} C^l_{n_\nu n}\frac{\phi_{n_\nu l}(r)}{r}Y_{lm}(\Omega)$$

From which it follows:

$$P_{nl}(r) = \sum_{n_\nu} C^l_{n_\nu n}\phi_{n_\nu l}(r)$$

The μ index runs over all combinations of nlm. In particular, here is an example of one possible way to index the basis of 12 radial functions for each l:

μ	n_μ	l_μ	m_μ
1	1	0	0
2	2	0	0
3	3	0	0
...	...		
12	12	0	0
13	1	1	−1
14	2	1	−1
15	3	1	−1
...	...		
24	12	1	−1
25	1	1	0
26	2	1	0
27	3	1	0
...	...		
36	12	1	0
37	1	1	1
38	2	1	1
39	3	1	1
...	...		
48	12	1	1
49	1	2	−2
50	2	2	−2
51	3	2	−2
...	...		

So the radial index n_μ always starts from 1 for each l_μ.

Overlap

The overlap matrix element

$$S_{\mu\nu} = \int \phi_\mu^*(\mathbf{x})\phi_\nu(\mathbf{x})\mathrm{d}^3x$$

becomes

$$\begin{aligned} S_{\mu\nu} = S_{n_\mu l_\mu m_\mu n_\nu l_\nu m_\nu} = \int \frac{\phi_{n_\mu l_\mu}(r)}{r} Y^*_{l_\mu m_\mu}(\Omega) \frac{\phi_{n_\nu l_\nu}(r)}{r} Y_{l_\nu m_\nu}(\Omega) r^2 \mathrm{d}r\mathrm{d}\Omega = \\ = \delta_{l_\mu l_\nu}\delta_{m_\mu m_\nu} \int_0^\infty \phi_{n_\mu l_\mu}(r)\phi_{n_\nu l_\nu}(r)\mathrm{d}r = \\ = \delta_{l_\mu l_\nu}\delta_{m_\mu m_\nu} \int_0^\infty \phi_{n_\mu l_\mu}(r)\phi_{n_\nu l_\mu}(r)\mathrm{d}r = \\ = \delta_{l_\mu l_\nu}\delta_{m_\mu m_\nu} S^{l_\mu}_{n_\mu n_\nu} \end{aligned}$$

where

$$S^l_{n_\mu n_\nu} = \int_0^\infty \phi_{n_\mu l}(r)\phi_{n_\nu l}(r)\mathrm{d}r$$

Potential

The potential matrix element

$$V_{\mu\nu} = \int \phi_\mu^*(\mathbf{x}) \left(-\frac{Z}{|\mathbf{x}|}\right) \phi_\nu(\mathbf{x}) \mathrm{d}^3x$$

becomes

$$\begin{aligned} V_{\mu\nu} = V_{n_\mu l_\mu m_\mu n_\nu l_\nu m_\nu} = \int \frac{\phi_{n_\mu l_\mu}(r)}{r} Y^*_{l_\mu m_\mu}(\Omega) \left(-\frac{Z}{r}\right) \frac{\phi_{n_\nu l_\nu}(r)}{r} Y_{l_\nu m_\nu}(\Omega) r^2 \mathrm{d}r\mathrm{d}\Omega = \\ = \delta_{l_\mu l_\nu} \delta_{m_\mu m_\nu} \int_0^\infty \phi_{n_\mu l_\mu}(r) \left(-\frac{Z}{r}\right) \phi_{n_\nu l_\nu}(r) \mathrm{d}r = \\ = \delta_{l_\mu l_\nu} \delta_{m_\mu m_\nu} \int_0^\infty \phi_{n_\mu l_\mu}(r) \left(-\frac{Z}{r}\right) \phi_{n_\nu l_\mu}(r) \mathrm{d}r = \\ = \delta_{l_\mu l_\nu} \delta_{m_\mu m_\nu} V^{l_\mu}_{n_\mu n_\nu} \end{aligned}$$

where

$$V^l_{n_\mu n_\nu} = \int_0^\infty \phi_{n_\mu l}(r) \left(-\frac{Z}{r}\right) \phi_{n_\nu l}(r) \mathrm{d}r$$

Kinetic

The kinetic matrix element

$$T_{\mu\nu} = \int \phi_\mu^*(\mathbf{x}) \left(-\tfrac{1}{2}\nabla^2\right) \phi_\nu(\mathbf{x}) \mathrm{d}^3x$$

becomes

$$\begin{aligned} T_{\mu\nu} = T_{n_\mu l_\mu m_\mu n_\nu l_\nu m_\nu} = \int \frac{\phi_{n_\mu l_\mu}(r)}{r} Y^*_{l_\mu m_\mu}(\Omega) \left(\left(-\tfrac{1}{2}\nabla^2\right) \frac{\phi_{n_\nu l_\nu}(r)}{r} Y_{l_\nu m_\nu}(\Omega)\right) r^2 \mathrm{d}r\mathrm{d}\Omega = \\ = \int \frac{\phi_{n_\mu l_\mu}(r)}{r} Y^*_{l_\mu m_\mu}(\Omega) \left(\left(-\tfrac{1}{2}\frac{\partial^2}{\partial r^2} - \frac{1}{r}\frac{\partial}{\partial r} + \frac{l_\nu(l_\nu+1)}{2r^2}\right) \frac{\phi_{n_\nu l_\nu}(r)}{r} Y_{l_\nu m_\nu}(\Omega)\right) r^2 \mathrm{d}r\mathrm{d}\Omega = \\ = \delta_{l_\mu l_\nu} \delta_{m_\mu m_\nu} \int_0^\infty \frac{\phi_{n_\mu l_\mu}(r)}{r} \left(\left(-\tfrac{1}{2}\frac{\partial^2}{\partial r^2} - \frac{1}{r}\frac{\partial}{\partial r} + \frac{l_\mu(l_\mu+1)}{2r^2}\right) \frac{\phi_{n_\nu l_\nu}(r)}{r}\right) r^2 \mathrm{d}r = \\ = \delta_{l_\mu l_\nu} \delta_{m_\mu m_\nu} T^{l_\mu}_{\mu\nu} \end{aligned}$$

where

$$\begin{aligned} T^l_{n_\mu n_\nu} = \int_0^\infty \frac{\phi_{n_\mu l}(r)}{r} \left(\left(-\tfrac{1}{2}\frac{\partial^2}{\partial r^2} - \frac{1}{r}\frac{\partial}{\partial r} + \frac{l(l+1)}{2r^2}\right) \frac{\phi_{n_\nu l}(r)}{r}\right) r^2 \mathrm{d}r = \\ = \int_0^\infty \frac{\phi_{n_\mu l}(r)}{r} \left(\left(-\frac{1}{2r}\frac{\partial^2}{\partial r^2} r + \frac{l(l+1)}{2r^2}\right) \frac{\phi_{n_\nu l}(r)}{r}\right) r^2 \mathrm{d}r = \\ = \int_0^\infty \phi_{n_\mu l}(r) \left(-\tfrac{1}{2}\frac{\partial^2}{\partial r^2} + \frac{l(l+1)}{2r^2}\right) \phi_{n_\nu l}(r) \mathrm{d}r = \\ = \int_0^\infty \left(-\tfrac{1}{2}\phi_{n_\mu l}(r)\phi''_{n_\nu l}(r) + \phi_{n_\mu l}(r)\frac{l(l+1)}{2r^2}\phi_{n_\nu l}(r)\right) \mathrm{d}r = \\ = \int_0^\infty \left(\tfrac{1}{2}\phi'_{n_\mu l}(r)\phi'_{n_\nu l}(r) + \phi_{n_\mu l}(r)\frac{l(l+1)}{2r^2}\phi_{n_\nu l}(r)\right) \mathrm{d}r \end{aligned}$$

G Matrix

The G matrix is given by:

$$G_{\mu\nu} = \sum_{\alpha\beta} P_{\alpha\beta} \left(\langle \mu\beta | \nu\alpha \rangle - \tfrac{1}{2} \langle \mu\beta | \alpha\nu \rangle \right) = \sum_{\alpha\beta} P_{\alpha\beta} \left((\mu\nu | \beta\alpha) - \tfrac{1}{2} (\mu\alpha | \beta\nu) \right)$$

Now we use:

$$\begin{aligned} \langle \alpha\beta | \gamma\delta \rangle = \langle n_\alpha l_\alpha m_\alpha \; n_\beta l_\beta m_\beta | n_\gamma l_\gamma m_\gamma \; n_\delta l_\delta m_\delta \rangle = (n_\alpha l_\alpha m_\alpha \; n_\gamma l_\gamma m_\gamma | n_\beta l_\beta m_\beta \; n_\delta l_\delta m_\delta) = \\ = \sum_{k=\max(|l_\alpha - l_\gamma|, |l_\beta - l_\delta|, |m_\alpha - m_\gamma|)}^{\min(l_\alpha + l_\gamma, l_\beta + l_\delta)} c^k(l_\alpha, m_\alpha, l_\gamma, m_\gamma) c^k(l_\delta, m_\delta, l_\beta, m_\beta) \\ \delta_{m_\alpha + m_\beta - m_\gamma - m_\delta, 0} \\ R^k(n_\alpha l_\alpha, n_\beta l_\beta, n_\gamma l_\gamma, n_\delta l_\delta) \end{aligned}$$

and

$$\begin{aligned} P_{\alpha\beta} &= 2 \sum_i C_{\alpha i} C_{\beta i} \\ P_{n_\alpha l_\alpha m_\alpha; n_\beta l_\beta m_\beta} &= 2 \sum_i C_{n_\alpha l_\alpha m_\alpha; n_i l_i m_i} C_{n_\beta l_\beta m_\beta; n_i l_i m_i} = \\ &= 2 \sum_i \delta_{l_i l_\alpha} \delta_{m_i m_\alpha} \delta_{l_i l_\beta} \delta_{m_i m_\beta} C^{l_i}_{n_\alpha n_i} C^{l_i}_{n_\beta n_i} = \\ &= 2 \delta_{l_\alpha l_\beta} \delta_{m_\alpha m_\beta} \sum_{n_i} C^{l_\alpha}_{n_\alpha n_i} C^{l_\alpha}_{n_\beta n_i} = \\ &= \delta_{l_\alpha l_\beta} \delta_{m_\alpha m_\beta} P^{l_\alpha}_{n_\alpha n_\beta} \end{aligned}$$

where

$$P^l_{n_\alpha n_\beta} = 2 \sum_{n_i} C^l_{n_\alpha n_i} C^l_{n_\beta n_i}$$

Where the sum over all occupied orbitals i can be written as:

$$\sum_i = \sum_{n_i l_i m_i} = \sum_{l=0}^{\infty} \sum_{m_l=-l}^{l} \sum_{n_l=1}^{\infty}$$

Where the sum over l is the outer sum, both $m = m_l$ and $n = n_l$ depend on l. We get:

$$G_{\mu\nu} = G_{n_\mu l_\mu m_\mu n_\nu l_\nu m_\nu} = \sum_{n_\alpha l_\alpha m_\alpha n_\beta l_\beta m_\beta} P_{n_\alpha l_\alpha m_\alpha n_\beta l_\beta m_\beta} \left(\langle \mu\beta | \nu\alpha \rangle - \tfrac{1}{2} \langle \mu\beta | \alpha\nu \rangle \right) = J_{\mu\nu} - K_{\mu\nu}$$

The first part is the direct term, the second part the exchange term. Let's treat the direct term first:

$$J_{\mu\nu} = \sum_{n_\alpha l_\alpha m_\alpha n_\beta l_\beta m_\beta} P_{n_\alpha l_\alpha m_\alpha n_\beta l_\beta m_\beta} \langle \mu\beta|\nu\alpha\rangle =$$

$$= \sum_{n_\alpha l_\alpha m_\alpha n_\beta l_\beta m_\beta} P_{n_\alpha l_\alpha m_\alpha n_\beta l_\beta m_\beta} \sum_{k=\max(|l_\mu - l_\nu|, |l_\beta - l_\alpha|, |m_\mu - m_\nu|)}^{\min(l_\mu + l_\nu, l_\beta + l_\alpha)} c^k(l_\mu, m_\mu, l_\nu, m_\nu) c^k(l_\alpha, m_\alpha, l_\beta, m_\beta)$$

$$\delta_{m_\mu + m_\beta - m_\nu - m_\alpha, 0}$$

$$R^k(n_\mu l_\mu, n_\beta l_\beta, n_\nu l_\nu, n_\alpha l_\alpha) =$$

$$= \sum_{n_\alpha l_\alpha m_\alpha n_\beta l_\beta m_\beta} \delta_{l_\alpha l_\beta} \delta_{m_\alpha m_\beta} P^{l_\alpha}_{n_\alpha n_\beta} \sum_{k=\max(|l_\mu - l_\nu|, |l_\beta - l_\alpha|, |m_\mu - m_\nu|)}^{\min(l_\mu + l_\nu, l_\beta + l_\alpha)} c^k(l_\mu, m_\mu, l_\nu, m_\nu) c^k(l_\alpha, m_\alpha, l_\beta, m_\beta)$$

$$\delta_{m_\mu + m_\beta - m_\nu - m_\alpha, 0}$$

$$R^k(n_\mu l_\mu, n_\beta l_\beta, n_\nu l_\nu, n_\alpha l_\alpha) =$$

$$= \sum_{n_\alpha n_\beta} \sum_{lm} P^l_{n_\alpha n_\beta} \sum_{k=\max(|l_\mu - l_\nu|, |m_\mu - m_\nu|)}^{\min(l_\mu + l_\nu, 2l)} c^k(l_\mu, m_\mu, l_\nu, m_\nu) c^k(l, m, l, m)$$

$$\delta_{m_\mu - m_\nu, 0}$$

$$R^k(n_\mu l_\mu, n_\beta l, n_\nu l_\nu, n_\alpha l) =$$

$$= \delta_{m_\mu m_\nu} \sum_{n_\alpha n_\beta} \sum_{l} P^l_{n_\alpha n_\beta} c^0(l_\mu, m_\mu, l_\nu, m_\nu)(2l+1) R^0(n_\mu l_\mu, n_\beta l, n_\nu l_\nu, n_\alpha l) =$$

$$= \delta_{l_\mu l_\nu} \delta_{m_\mu m_\nu} \sum_{l} (2l+1) \sum_{n_\alpha n_\beta} P^l_{n_\alpha n_\beta} R^0(n_\mu l_\mu, n_\beta l, n_\nu l_\mu, n_\alpha l)$$

For the exchange term we get:

$$K_{\mu\nu} = \tfrac{1}{2} \sum_{n_\alpha l_\alpha m_\alpha n_\beta l_\beta m_\beta} P_{n_\alpha l_\alpha m_\alpha n_\beta l_\beta m_\beta} \langle \mu\beta|\alpha\nu\rangle =$$

$$= \tfrac{1}{2} \sum_{n_\alpha l_\alpha m_\alpha n_\beta l_\beta m_\beta} P_{n_\alpha l_\alpha m_\alpha n_\beta l_\beta m_\beta} \sum_{k=\max(|l_\mu - l_\alpha|, |l_\beta - l_\nu|, |m_\mu - m_\alpha|)}^{\min(l_\mu + l_\alpha, l_\beta + l_\nu)} c^k(l_\mu, m_\mu, l_\alpha, m_\alpha) c^k(l_\nu, m_\nu, l_\beta, m_\beta)$$

$$\delta_{m_\mu + m_\beta - m_\alpha - m_\nu, 0}$$

$$R^k(n_\mu l_\mu, n_\beta l_\beta, n_\alpha l_\alpha, n_\nu l_\nu) =$$

$$= \tfrac{1}{2} \sum_{n_\alpha l_\alpha m_\alpha n_\beta l_\beta m_\beta} \delta_{l_\alpha l_\beta} \delta_{m_\alpha m_\beta} P^{l_\alpha}_{n_\alpha n_\beta} \sum_{k=\max(|l_\mu - l_\alpha|, |l_\beta - l_\nu|, |m_\mu - m_\alpha|)}^{\min(l_\mu + l_\alpha, l_\beta + l_\nu)} c^k(l_\mu, m_\mu, l_\alpha, m_\alpha) c^k(l_\nu, m_\nu, l_\beta, m_\beta)$$

$$\delta_{m_\mu + m_\beta - m_\alpha - m_\nu, 0}$$

$$R^k(n_\mu l_\mu, n_\beta l_\beta, n_\alpha l_\alpha, n_\nu l_\nu) =$$

$$= \tfrac{1}{2} \delta_{m_\mu m_\nu} \sum_{n_\alpha n_\beta} \sum_{lm} P^l_{n_\alpha n_\beta} \sum_{k=\max(|l_\mu - l|, |l - l_\nu|, |m_\mu - m|)}^{\min(l_\mu + l, l + l_\nu)} c^k(l_\mu, m_\mu, l, m) c^k(l_\nu, m_\nu, l, m)$$

$$R^k(n_\mu l_\mu, n_\beta l, n_\alpha l, n_\nu l_\nu) =$$

$$= \tfrac{1}{2} \delta_{m_\mu m_\nu} \sum_{n_\alpha n_\beta} \sum_{lm} P^l_{n_\alpha n_\beta} \sum_{k=\max(|l_\mu - l|, |l - l_\nu|, |m_\mu - m|)}^{\min(l_\mu + l, l + l_\nu)} c^k(l_\mu, m_\mu, l, m) c^k(l_\nu, m_\mu, l, m)$$

$$R^k(n_\mu l_\mu, n_\beta l, n_\alpha l, n_\nu l_\nu)$$

For $l_\mu = l_\nu$ this can be written as:

$$K_{\mu\nu} = \tfrac{1}{2}\delta_{l_\mu l_\nu}\delta_{m_\mu m_\nu}\sum_{n_\alpha n_\beta}\sum_{lm} P^l_{n_\alpha n_\beta}\sum_{k=\max(|l_\mu - l|,|m_\mu - m|)}^{l_\mu + l} c^k(l_\mu, m_\mu, l, m)c^k(l_\mu, m_\mu, l, m)$$
$$R^k(n_\mu l_\mu, n_\beta l, n_\alpha l, n_\nu l_\mu) =$$
$$= \tfrac{1}{2}\delta_{l_\mu l_\nu}\delta_{m_\mu m_\nu}\sum_{n_\alpha n_\beta}\sum_{l} P^l_{n_\alpha n_\beta}\sum_{k=|l_\mu - l|}^{l_\mu + l}\sqrt{\frac{2l+1}{2l_\mu+1}}c^k(l_\mu, 0, l, 0)$$
$$R^k(n_\mu l_\mu, n_\beta l, n_\alpha l, n_\nu l_\mu) =$$
$$= \delta_{l_\mu l_\nu}\delta_{m_\mu m_\nu}\sum_{l}(2l+1)\sum_{n_\alpha n_\beta} P^l_{n_\alpha n_\beta}\sum_{k=|l_\mu - l|}^{l_\mu + l}\tfrac{1}{2}\begin{pmatrix} l_\mu & k & l \\ 0 & 0 & 0\end{pmatrix}^2$$
$$R^k(n_\mu l_\mu, n_\beta l, n_\alpha l, n_\nu l_\mu)$$

All together we get:

$$G_{\mu\nu} = \delta_{l_\mu l_\nu}\delta_{m_\mu m_\nu}\sum_{l}(2l+1)\sum_{n_\alpha n_\beta} P^l_{n_\alpha n_\beta}$$
$$\left(R^0(n_\mu l_\mu, n_\beta l, n_\nu l_\mu, n_\alpha l) - \sum_{k=|l_\mu - l|}^{l_\mu + l}\tfrac{1}{2}\begin{pmatrix} l_\mu & k & l \\ 0 & 0 & 0\end{pmatrix}^2 R^k(n_\mu l_\mu, n_\beta l, n_\alpha l, n_\nu l_\mu)\right) =$$
$$= \delta_{l_\mu l_\nu}\delta_{m_\mu m_\nu}G^{l_\mu}_{n_\mu n_\nu}$$

where

$$G^l_{n_\mu n_\nu} = \sum_{l'}(2l'+1)\sum_{n_\alpha n_\beta} P^{l'}_{n_\alpha n_\beta}$$
$$\left(R^0(n_\mu l, n_\beta l', n_\nu l, n_\alpha l') - \sum_{k=|l - l'|}^{l + l'}\tfrac{1}{2}\begin{pmatrix} l & k & l' \\ 0 & 0 & 0\end{pmatrix}^2 R^k(n_\mu l, n_\beta l', n_\alpha l', n_\nu l)\right)$$

Note: performing the sum over n_α and n_β we get:

$$G^l_{n_\mu n_\nu} = \sum_{l'} 2(2l'+1)\sum_{n'}\sum_{n_\alpha n_\beta} C^{l'}_{n_\alpha n'}C^{l'}_{n_\beta n'}$$
$$\left(R^0(n_\mu l, n_\beta l', n_\nu l, n_\alpha l') - \sum_{k=|l - l'|}^{l + l'}\tfrac{1}{2}\begin{pmatrix} l & k & l' \\ 0 & 0 & 0\end{pmatrix}^2 R^k(n_\mu l, n_\beta l', n_\alpha l', n_\nu l)\right) =$$
$$= \sum_{l'} 2(2l'+1)\sum_{n'}$$
$$\left(R^0(n_\mu l, n'l', n_\nu l, n'l') - \sum_{k=|l - l'|}^{l + l'}\tfrac{1}{2}\begin{pmatrix} l & k & l' \\ 0 & 0 & 0\end{pmatrix}^2 R^k(n_\mu l, n'l', n'l', n_\nu l)\right)$$

Radial Roothaan Equations

The Roothaan equations are:

$$\sum_{\nu}(T+V+G)_{\mu\nu}C_{\nu i} = \epsilon_i \sum_{\nu} S_{\mu\nu}C_{\nu i}$$

$$\sum_{\nu}(T+V+G)_{n_\mu l_\mu m_\mu n_\nu l_\nu m_\nu}C_{n_\nu l_\nu m_\nu n_i l_i m_i} = \epsilon_{n_i l_i m_i} \sum_{\nu} S_{n_\mu l_\mu m_\mu n_\nu l_\nu m_\nu}C_{n_\nu l_\nu m_\nu n_i l_i m_i}$$

$$\sum_{\nu}(T+V+G)_{n_\mu l_\mu m_\mu n_\nu l_\nu m_\nu}\delta_{l_i l_\nu}\delta_{m_i m_\nu}C^{l_i}_{n_\nu n_i} = \epsilon_{n_i l_i m_i} \sum_{\nu} S_{n_\mu l_\mu m_\mu n_\nu l_\nu m_\nu}\delta_{l_i l_\nu}\delta_{m_i m_\nu}C^{l_i}_{n_\nu n_i}$$

$$\sum_{n_\nu}(T+V+G)_{n_\mu l_\mu m_\mu n_\nu l_i m_i}C^{l_i}_{n_\nu n_i} = \epsilon_{n_i l_i m_i} \sum_{n_\nu} S_{n_\mu l_\mu m_\mu n_\nu l_i m_i}C^{l_i}_{n_\nu n_i}$$

$$\delta_{l_\mu l_i}\delta_{m_\mu m_i}\sum_{n_\nu}(T+V+G)^{l_i}_{n_\mu n_\nu}C^{l_i}_{n_\nu n_i} = \delta_{l_\mu l_i}\delta_{m_\mu m_i}\epsilon_{n_i l_i m_i} \sum_{n_\nu} S^{l_i}_{n_\mu n_\nu}C^{l_i}_{n_\nu n_i}$$

The eigenvalues will be degenerate with respect to m_i and so the radial Roothaan equations are to be solved for each l:

$$\sum_{n_\nu}(T+V+G)^{l}_{n_\mu n_\nu}C^{l}_{n_\nu n_i} = \epsilon_{n_i l} \sum_{n_\nu} S^{l}_{n_\mu n_\nu}C^{l}_{n_\nu n_i}$$

The total energy is:

$$E = \tfrac{1}{2}\sum_{\mu\nu} P_{\nu\mu}(H^{\text{core}}_{\mu\nu} + F_{\mu\nu}) =$$

$$= \sum_{\mu\nu} P_{\nu\mu}(F_{\mu\nu} - \tfrac{1}{2}G_{\mu\nu}) =$$

$$= \sum_{\mu\nu} \delta_{l_\mu l_\nu}\delta_{m_\mu m_\nu}P^{l_\mu}_{n_\mu n_\nu}\delta_{l_\mu l_\nu}\delta_{m_\mu m_\nu}(F - \tfrac{1}{2}G)^{l_\mu}_{n_\mu n_\nu} =$$

$$= \sum_{l_\mu}\sum_{m_\mu}\sum_{n_\mu n_\nu} P^{l_\mu}_{n_\mu n_\nu}(F - \tfrac{1}{2}G)^{l_\mu}_{n_\mu n_\nu} =$$

$$= \sum_{l_\mu}\sum_{n_\mu n_\nu} (2l_\mu + 1)P^{l_\mu}_{n_\mu n_\nu}(F - \tfrac{1}{2}G)^{l_\mu}_{n_\mu n_\nu} =$$

$$= \sum_{l}\sum_{n_\mu n_\nu} (2l + 1)P^{l}_{n_\mu n_\nu}(F - \tfrac{1}{2}G)^{l}_{n_\mu n_\nu} =$$

$$= \sum_{l}\sum_{n_\mu n_\nu} (2l+1)\sum_{n} 2C^{l}_{n_\mu n}C^{l}_{n_\nu n}(F^{l}_{n_\mu n_\nu} - \tfrac{1}{2}G^{l}_{n_\mu n_\nu}) =$$

$$= \sum_{l}\sum_{n_\mu n_\nu} (2l+1)\sum_{n} 2C^{l}_{n_\mu n}C^{l}_{n_\nu n}\left(\epsilon_{nl}S^{l}_{n_\mu n_\nu} - \tfrac{1}{2}\sum_{l'} 2(2l'+1)\sum_{n'}\right.$$

$$\left.\left(R^0(n_\mu l, n'l', n_\nu l, n'l') - \sum_{k=|l-l'|}^{l+l'} \tfrac{1}{2}\begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 R^k(n_\mu l, n'l', n'l', n_\nu l)\right)\right) =$$

$$= \sum_{l}\sum_{n} 2(2l+1)\left(\epsilon_{nl} - \sum_{l'}\sum_{n'}(2l'+1)\left(R^0(nl, n'l', nl, n'l') - \tfrac{1}{2}\sum_{k=|l-l'|}^{k=l+l'}\begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 R^k(nl, n'l', n'l', nl)\right)\right)$$

Two particle

We use the following functions for ψ:

$$\psi_i(\mathbf{x}) = \frac{P_{n_1 l_1}(r)}{r} Y_{l_1 m_1}(\Omega)$$
$$\psi_j(\mathbf{x}) = \frac{P_{n_1' l_1'}(r)}{r} Y_{l_1' m_1'}(\Omega)$$
$$\psi_k(\mathbf{x}) = \frac{P_{n_2 l_2}(r)}{r} Y_{l_2 m_2}(\Omega)$$
$$\psi_l(\mathbf{x}) = \frac{P_{n_2' l_2'}(r)}{r} Y_{l_2' m_2'}(\Omega)$$

And the multipole expansion:

$$\frac{1}{|\mathbf{x} - \mathbf{x}'|} = \sum_{k,q} \frac{r_<^k}{r_>^{k+1}} \frac{4\pi}{2k+1} Y_{kq}(\Omega) Y_{kq}^*(\Omega')$$

And we get:

$$
\begin{aligned}
(ij|kl) = \langle ik|jl\rangle = (11'|22') = \langle 12|1'2'\rangle = \\
= \langle l_1 m_1 l_2 m_2 | \frac{1}{|\mathbf{x}-\mathbf{x}'|} | l_1' m_1' l_2' m_2' \rangle = \\
= \int \frac{\psi_i^*(\mathbf{x})\psi_j(\mathbf{x})\psi_k^*(\mathbf{x}')\psi_l(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|} \mathrm{d}^3x \mathrm{d}^3x' = \\
= \int \frac{P_{n_1 l_1}(r)}{r} Y^*_{l_1 m_1}(\Omega) \frac{P_{n_1' l_1'}(r)}{r} Y_{l_1' m_1'}(\Omega) \frac{P_{n_2 l_2}(r')}{r'} Y^*_{l_2 m_2}(\Omega') \frac{P_{n_2' l_2'}(r')}{r'} Y_{l_2' m_2'}(\Omega') \\
\sum_{k,q} \frac{r_<^k}{r_>^{k+1}} \frac{4\pi}{2k+1} Y_{kq}(\Omega) Y^*_{kq}(\Omega') r^2 r'^2 \mathrm{d}r\mathrm{d}r'\mathrm{d}\Omega\mathrm{d}\Omega' = \\
= \sum_{k,q} \int Y^*_{l_1 m_1}(\Omega) Y_{l_1' m_1'}(\Omega) Y_{kq}(\Omega) \mathrm{d}\Omega \int Y^*_{l_2 m_2}(\Omega') Y_{l_2' m_2'}(\Omega') Y^*_{kq}(\Omega') \mathrm{d}\Omega' \\
\int \frac{r_<^k}{r_>^{k+1}} \frac{4\pi}{2k+1} P_{n_1 l_1}(r) P_{n_1' l_1'}(r) P_{n_2 l_2}(r') P_{n_2' l_2'}(r') \mathrm{d}r\mathrm{d}r' = \\
= \sum_{k,q} (-1)^{m_1+m_2+q} \int Y_{l_1,-m_1}(\Omega) Y_{l_1' m_1'}(\Omega) Y_{kq}(\Omega) \mathrm{d}\Omega \int Y_{l_2,-m_2}(\Omega') Y_{l_2' m_2'}(\Omega') Y_{k,-q}(\Omega') \mathrm{d}\Omega' \\
\int \frac{r_<^k}{r_>^{k+1}} \frac{4\pi}{2k+1} P_{n_1 l_1}(r) P_{n_1' l_1'}(r) P_{n_2 l_2}(r') P_{n_2' l_2'}(r') \mathrm{d}r\mathrm{d}r' = \\
= \sum_{k,q} (-1)^{m_1+m_2+q} \sqrt{\frac{(2l_1+1)(2l_1'+1)(2k+1)}{4\pi}} \begin{pmatrix} l_1 & l_1' & k \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & l_1' & k \\ -m_1 & m_1' & q \end{pmatrix} \\
\sqrt{\frac{(2l_2+1)(2l_2'+1)(2k+1)}{4\pi}} \begin{pmatrix} l_2 & l_2' & k \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_2 & l_2' & k \\ -m_2 & m_2' & -q \end{pmatrix} \\
\int \frac{r_<^k}{r_>^{k+1}} \frac{4\pi}{2k+1} P_{n_1 l_1}(r) P_{n_1' l_1'}(r) P_{n_2 l_2}(r') P_{n_2' l_2'}(r') \mathrm{d}r\mathrm{d}r' = \\
= \sum_{k} \sqrt{(2l_1+1)(2l_1'+1)(2l_2+1)(2l_2'+1)} \begin{pmatrix} l_1 & l_1' & k \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_2 & l_2' & k \\ 0 & 0 & 0 \end{pmatrix} \\
\sum_{q=-k}^{k} (-1)^{m_1+m_2+q} \begin{pmatrix} l_1 & l_1' & k \\ -m_1 & m_1' & q \end{pmatrix} \begin{pmatrix} l_2 & l_2' & k \\ -m_2 & m_2' & -q \end{pmatrix} \\
\int \frac{r_<^k}{r_>^{k+1}} P_{n_1 l_1}(r) P_{n_1' l_1'}(r) P_{n_2 l_2}(r') P_{n_2' l_2'}(r') \mathrm{d}r\mathrm{d}r' = \\
= \sum_{k=\max(|l_1-l_1'|,|l_2-l_2'|,|m_1-m_1'|)}^{\min(l_1+l_1',l_2+l_2')} \sqrt{(2l_1+1)(2l_1'+1)(2l_2+1)(2l_2'+1)} \\
(-1)^{m_1+m_2'} \delta_{m_1+m_2-m_1'-m_2',0} \begin{pmatrix} l_1 & l_1' & k \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_2 & l_2' & k \\ 0 & 0 & 0 \end{pmatrix} \\
\begin{pmatrix} l_1 & l_1' & k \\ -m_1 & m_1' & m_1-m_1' \end{pmatrix} \begin{pmatrix} l_2 & l_2' & k \\ -m_2 & m_2' & m_2-m_2' \end{pmatrix} \\
\int \frac{r_<^k}{r_>^{k+1}} P_{n_1 l_1}(r) P_{n_1' l_1'}(r) P_{n_2 l_2}(r') P_{n_2' l_2'}(r') \mathrm{d}r\mathrm{d}r'
\end{aligned}
$$

In the last step we used the fact that the $3j$ symbols are zero unless $-m_1 + m_1' + q = 0$ and $-m_2 + m_2' - q = 0$, from which it follows that $q = m_1 - m_1' = -m_2 + m_2'$ and so one of the $3j$ symbols is zero unless $m_1 + m_2 - m_1' - m_2' = 0$, which is expressed by $\delta_{m_1+m_2-m_1'-m_2',0}$. Given this condition, the sum over q must be such that one q is equal to $m_1 - m_1' = -m_2 + m_2'$, which means that $k \geq |m_1 - m_1'| = |m_2 - m_2'|$ otherwise the $3j$ symbols will be zero.

Finally, k must also satisfy the conditions $|l_1 - l_1'| \leq k \leq l_1 + l_1'$ and $|l_2 - l_2'| \leq k \leq l_2 + l_2'$. The sign factor $(-1)^{m_1+m_2+q} = (-1)^{m_1+m_2+m_1-m_1'} = (-1)^{m_1+m_2-m_2+m_2'}$ is equal to both $(-1)^{m_1+m_2'}$ and $(-1)^{m_2-m_1'}$ so we just used the former.

We can write this using the c^k symbols as:

$$
\begin{aligned}
(ij|kl) = & \sum_{k=\max(|l_1-l_1'|,|l_2-l_2'|,|m_1-m_1'|)}^{\min(l_1+l_1',l_2+l_2')} \sqrt{(2l_1+1)(2l_1'+1)(2l_2+1)(2l_2'+1)} \\
& (-1)^{m_2-m_2'}(-1)^{-m_1}(-1)^{-m_2}\delta_{m_1+m_2-m_1'-m_2',0}
\begin{pmatrix} l_1 & l_1' & k \\ 0 & 0 & 0 \end{pmatrix}
\begin{pmatrix} l_2 & l_2' & k \\ 0 & 0 & 0 \end{pmatrix} \\
& \begin{pmatrix} l_1 & l_1' & k \\ -m_1 & m_1' & m_1 - m_1' \end{pmatrix}
\begin{pmatrix} l_2 & l_2' & k \\ -m_2 & m_2' & m_2 - m_2' \end{pmatrix} \\
& \int \frac{r_<^k}{r_>^{k+1}} P_{n_1 l_1}(r) P_{n_1' l_1'}(r) P_{n_2 l_2}(r') P_{n_2' l_2'}(r') \mathrm{d}r\mathrm{d}r' = \\
= & \sum_{k=\max(|l_1-l_1'|,|l_2-l_2'|,|m_1-m_1'|)}^{\min(l_1+l_1',l_2+l_2')} c^k(l_1, m_1, l_1', m_1') c^k(l_2, m_2, l_2', m_2') \\
& (-1)^{m_2-m_2'}\delta_{m_1+m_2-m_1'-m_2',0} \\
& \int \frac{r_<^k}{r_>^{k+1}} P_{n_1 l_1}(r) P_{n_1' l_1'}(r) P_{n_2 l_2}(r') P_{n_2' l_2'}(r') \mathrm{d}r\mathrm{d}r' = \\
= & \sum_{k=\max(|l_1-l_1'|,|l_2-l_2'|,|m_1-m_1'|)}^{\min(l_1+l_1',l_2+l_2')} c^k(l_1, m_1, l_1', m_1') c^k(l_2', m_2', l_2, m_2) \\
& \delta_{m_1+m_2-m_1'-m_2',0} \\
& \int \frac{r_<^k}{r_>^{k+1}} P_{n_1 l_1}(r) P_{n_1' l_1'}(r) P_{n_2 l_2}(r') P_{n_2' l_2'}(r') \mathrm{d}r\mathrm{d}r' = \\
= & \sum_{k=\max(|l_1-l_1'|,|l_2-l_2'|,|m_1-m_1'|)}^{\min(l_1+l_1',l_2+l_2')} c^k(l_1, m_1, l_1', m_1') c^k(l_2', m_2', l_2, m_2) \\
& \delta_{m_1+m_2-m_1'-m_2',0} \\
& R^k(n_1 l_1, n_2 l_2, n_1' l_1', n_2' l_2')
\end{aligned}
$$

We can also couple the angular momenta as follows:

$$
\begin{aligned}
|l_1 l_2 LM\rangle &= \sum_{m_1 m_2} (l_1 m_1 l_2 m_2 | LM) \, |l_1 m_1\rangle \, |l_2 m_2\rangle = \\
&= \sum_{m_1 m_2} (-1)^{l_1 - l_2 + M} \sqrt{2L+1} \begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & -M \end{pmatrix} |l_1 m_1\rangle \, |l_2 m_2\rangle
\end{aligned}
$$

and we get for the matrix elements:

$$\langle l_1 l_2 LM|\frac{1}{|\mathbf{x}-\mathbf{x}'|}|l_1' l_2' L' M'\rangle =$$

$$= \sum_{m_1 m_2}\sum_{m_1' m_2'}(-1)^{l_1-l_2+l_1'-l_2'+M+M'}\sqrt{(2L+1)(2L'+1)}$$

$$\begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & -M \end{pmatrix}\begin{pmatrix} l_1' & l_2' & L' \\ m_1' & m_2' & -M' \end{pmatrix}$$

$$\langle l_1 m_1|\langle l_2 m_2|\frac{1}{|\mathbf{x}-\mathbf{x}'|}|l_1' m_1'\rangle|l_2 m_2'\rangle =$$

$$= \sum_{m_1 m_2}\sum_{m_1' m_2'}(-1)^{l_1-l_2+l_1'-l_2'+M+M'}\sqrt{(2L+1)(2L'+1)}$$

$$\begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & -M \end{pmatrix}\begin{pmatrix} l_1' & l_2' & L' \\ m_1' & m_2' & -M' \end{pmatrix}$$

$$(ij|kl) =$$

$$= \sum_{m_1 m_2}\sum_{m_1' m_2'}(-1)^{l_1-l_2+l_1'-l_2'+M+M'}\sqrt{(2L+1)(2L'+1)}$$

$$\begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & -M \end{pmatrix}\begin{pmatrix} l_1' & l_2' & L' \\ m_1' & m_2' & -M' \end{pmatrix}$$

$$\sum_k\sqrt{(2l_1+1)(2l_1'+1)(2l_2+1)(2l_2'+1)}\begin{pmatrix} l_1 & l_1' & k \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l_2 & l_2' & k \\ 0 & 0 & 0 \end{pmatrix}$$

$$\sum_{q=-k}^{k}(-1)^{m_1+m_2+q}\begin{pmatrix} l_1 & l_1' & k \\ -m_1 & m_1' & q \end{pmatrix}\begin{pmatrix} l_2 & l_2' & k \\ -m_2 & m_2' & -q \end{pmatrix}$$

$$\int\frac{r_<^k}{r_>^{k+1}}P_{n_1 l_1}(r)P_{n_1' l_1'}(r)P_{n_2 l_2}(r')P_{n_2' l_2'}(r')\mathrm{d}r\mathrm{d}r' =$$

$$= \sum_{m_1 m_2}\sum_{m_1' m_2'}(-1)^{l_1-l_2+l_1'-l_2'}(2L+1)$$

$$\delta_{MM'}\delta_{LL'}\begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & -M \end{pmatrix}\begin{pmatrix} l_1' & l_2' & L \\ m_1' & m_2' & -M \end{pmatrix}$$

$$\sum_k\sqrt{(2l_1+1)(2l_1'+1)(2l_2+1)(2l_2'+1)}\begin{pmatrix} l_1 & l_1' & k \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l_2 & l_2' & k \\ 0 & 0 & 0 \end{pmatrix}$$

$$\sum_{q=-k}^{k}(-1)^{m_1+m_2+q}\begin{pmatrix} l_1 & l_1' & k \\ -m_1 & m_1' & q \end{pmatrix}\begin{pmatrix} l_2 & l_2' & k \\ -m_2 & m_2' & -q \end{pmatrix}$$

$$\int\frac{r_<^k}{r_>^{k+1}}P_{n_1 l_1}(r)P_{n_1' l_1'}(r)P_{n_2 l_2}(r')P_{n_2' l_2'}(r')\mathrm{d}r\mathrm{d}r' =$$

$$= (-1)^{l_1-l_2+l_1'-l_2'}(2L+1)$$

$$\sum_k\sqrt{(2l_1+1)(2l_1'+1)(2l_2+1)(2l_2'+1)}\begin{pmatrix} l_1 & l_1' & k \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l_2 & l_2' & k \\ 0 & 0 & 0 \end{pmatrix}$$

$$\delta_{MM'}\delta_{LL'}(-1)^{l_1+l_1'+L}\begin{Bmatrix} l_1 & l_2 & L \\ l_2' & l_1' & k \end{Bmatrix}$$

$$\int\frac{r_<^k}{r_>^{k+1}}P_{n_1 l_1}(r)P_{n_1' l_1'}(r)P_{n_2 l_2}(r')P_{n_2' l_2'}(r')\mathrm{d}r\mathrm{d}r' =$$

$$= \sum_k\int\frac{r_<^k}{r_>^{k+1}}P_{n_1 l_1}(r)P_{n_1' l_1'}(r)P_{n_2 l_2}(r')P_{n_2' l_2'}(r')\mathrm{d}r\mathrm{d}r'$$

$$(-1)^{L-l_2-l_2'}(2L+1)\delta_{MM'}\delta_{LL'}\sqrt{(2l_1+1)(2l_1'+1)(2l_2+1)(2l_2'+1)}$$

$$\begin{pmatrix} l_1 & l_1' & k \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l_2 & l_2' & k \\ 0 & 0 & 0 \end{pmatrix}\begin{Bmatrix} l_1 & l_2 & L \\ l_2' & l_1' & k \end{Bmatrix} =$$

$$= \sum\int\frac{r_<^k}{r_>^{k+1}}P_{n_1 l_1}(r)P_{n_1' l_1'}(r)P_{n_2 l_2}(r')P_{n_2' l_2'}(r')\mathrm{d}r\mathrm{d}r'$$

Where we used the $6j$ symbol:

$$\begin{Bmatrix} l_1 & l_2 & L \\ l_2' & l_1' & k \end{Bmatrix} = \sum_{m_1 m_2 m_1' m_2' M q} (-1)^{l_1+l_2+l_1'+l_2'+L+k-m_1-m_2-m_1'-m_2'-M-q}$$

$$\begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & -M \end{pmatrix} \begin{pmatrix} l_1 & l_1' & k \\ -m_1 & m_1' & q \end{pmatrix} \begin{pmatrix} l_2' & l_1' & L \\ m_2' & -m_1' & M \end{pmatrix} \begin{pmatrix} l_2' & l_2 & k \\ -m_2' & -m_2 & -q \end{pmatrix} =$$

$$= \sum_{m_1 m_2 m_1' m_2' M q} (-1)^{l_1+l_2+l_1'+l_2'+L+k-m_1-m_2-m_1'-m_2'-M-q}$$

$$\begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & -M \end{pmatrix} \begin{pmatrix} l_1 & l_1' & k \\ -m_1 & m_1' & q \end{pmatrix} \begin{pmatrix} l_1' & l_2' & L \\ m_1' & -m_2' & -M \end{pmatrix} (-1)^{l_2+l_2'+k} \begin{pmatrix} l_2 & l_2' & k \\ -m_2 & -m_2' & -q \end{pmatrix} =$$

$$= \sum_{m_1 m_2 m_1' m_2' q} \delta_{M,m_1'+m_2'} (-1)^{l_1+l_1'+L} (-1)^{m_1+m_2+q}$$

$$\begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & -M \end{pmatrix} \begin{pmatrix} l_1 & l_1' & k \\ -m_1 & m_1' & q \end{pmatrix} \begin{pmatrix} l_1' & l_2' & L \\ m_1' & +m_2' & -M \end{pmatrix} \begin{pmatrix} l_2 & l_2' & k \\ -m_2 & +m_2' & -q \end{pmatrix}$$

Where we have renamed $-m_2'$ to m_2'.

8.10.5 Slater Type Orbitals (STO)

In this section we express the matrix elements in the STO basis. It turns out that all integrals that we need can be expressed in terms of the following simple integral (where $n, \zeta \geq 0$):

$$\int_0^\infty r^n e^{-\zeta r} \mathrm{d}r = \int_0^\infty \left(\frac{x}{\zeta}\right)^n e^{-x} \frac{\mathrm{d}x}{\zeta} = \frac{1}{\zeta^{n+1}} \int_0^\infty x^n e^{-x} \mathrm{d}x = \frac{\Gamma(n+1)}{\zeta^{n+1}} = \frac{n!}{\zeta^{n+1}} \tag{8.10.5.1}$$

The STO basis function for the radial Schrödinger equation for $P(r)$ is:

$$P_{n\zeta}(r) = N_{n\zeta} r^n e^{-\zeta r} \tag{8.10.5.2}$$

Where the normalization constant $N_{n\zeta}$ is such that the STO orbital is normalized as the radial wavefunction $P(r)$:

$$1 = \int_0^\infty P_{n\zeta}^2(r) \mathrm{d}r = N_{n\zeta}^2 \int_0^\infty r^{2n} e^{-2\zeta r} \mathrm{d}r = N_{n\zeta}^2 \frac{(2n)!}{(2\zeta)^{2n+1}}$$

from which we get:

$$N_{n\zeta} = \sqrt{\frac{(2\zeta)^{2n+1}}{(2n)!}}$$

Note that for $R(r) = \frac{P(r)}{r}$ we get the following STO basis function:

$$R_{n\zeta}(r) = \frac{P_{n\zeta}(r)}{r} = N_{n\zeta} r^{n-1} e^{-\zeta r} \tag{8.10.5.3}$$

One uses either (8.10.5.2) or (8.10.5.3) depending on whether one solves the radial Schrödinger equation for P or for $R = \frac{P}{r}$.

Overlap

$$
\begin{aligned}
S_{ij} &= \int P_{n_i\zeta_i}(r) P_{n_j\zeta_j}(r) \mathrm{d}r = \\
&= \int N_{n_i\zeta_i} r^{n_i} e^{-\zeta_i r} N_{n_j\zeta_j} r^{n_j} e^{-\zeta_j r} \mathrm{d}r = \\
&= N_{n_i\zeta_i} N_{n_j\zeta_j} \int r^{n_i+n_j} e^{-(\zeta_i+\zeta_j)r} \mathrm{d}r = \\
&= N_{n_i\zeta_i} N_{n_j\zeta_j} \frac{(n_i+n_j)!}{(\zeta_i+\zeta_j)^{n_i+n_j+1}}
\end{aligned}
$$

Potential

$$
\begin{aligned}
V_{ij} &= \int P_{n_i\zeta_i}(r) \left(-\frac{Z}{r}\right) P_{n_j\zeta_j}(r) \mathrm{d}r = \\
&= \int N_{n_i\zeta_i} r^{n_i} e^{-\zeta_i r} \left(-\frac{Z}{r}\right) N_{n_j\zeta_j} r^{n_j} e^{-\zeta_j r} \mathrm{d}r = \\
&= -Z N_{n_i\zeta_i} N_{n_j\zeta_j} \int r^{n_i+n_j-1} e^{-(\zeta_i+\zeta_j)r} \mathrm{d}r = \\
&= -Z N_{n_i\zeta_i} N_{n_j\zeta_j} \frac{(n_i+n_j-1)!}{(\zeta_i+\zeta_j)^{n_i+n_j}}
\end{aligned}
$$

Kinetic

$$
\begin{aligned}
T_{ij} &= \int \left(\tfrac{1}{2} P'_{n_i\zeta_i}(r) P'_{n_j\zeta_j}(r) + P_{n_i\zeta_i}(r) \frac{l(l+1)}{2r^2} P_{n_j\zeta_j}(r) \right) \mathrm{d}r = \\
&= \tfrac{1}{2} N_{n_i\zeta_i} N_{n_j\zeta_j} \int \left(\frac{\mathrm{d}}{\mathrm{d}r}(r^{n_i} e^{-\zeta_i r}) \frac{\mathrm{d}}{\mathrm{d}r}(r^{n_j} e^{-\zeta_j r}) + r^{n_i} e^{-\zeta_i r} \frac{l(l+1)}{r^2} r^{n_j} e^{-\zeta_j r} \right) \mathrm{d}r = \\
&= \tfrac{1}{2} N_{n_i\zeta_i} N_{n_j\zeta_j} \int \left((n_i r^{n_i-1} e^{-\zeta_i r} - \zeta_i r^{n_i} e^{-\zeta_i r})(n_j r^{n_j-1} e^{-\zeta_j r} - \zeta_j r^{n_j} e^{-\zeta_j r}) + l(l+1) r^{n_i+n_j-2} e^{-(\zeta_i+\zeta_j)r} \right) \mathrm{d}r = \\
&= \tfrac{1}{2} N_{n_i\zeta_i} N_{n_j\zeta_j} \int \left(\left(\frac{n_i n_j}{r^2} - \frac{n_i\zeta_j + n_j\zeta_i}{r} + \zeta_i\zeta_j \right) r^{n_i+n_j} e^{-(\zeta_i+\zeta_j)r} + l(l+1) r^{n_i+n_j-2} e^{-(\zeta_i+\zeta_j)r} \right) \mathrm{d}r = \\
&= \tfrac{1}{2} N_{n_i\zeta_i} N_{n_j\zeta_j} \int \left((n_i n_j + l(l+1)) r^{n_i+n_j-2} e^{-(\zeta_i+\zeta_j)r} - (n_i\zeta_j + n_j\zeta_i) r^{n_i+n_j-1} e^{-(\zeta_i+\zeta_j)r} + \zeta_i\zeta_j r^{n_i+n_j} e^{-(\zeta_i+\zeta_j)r} \right) \mathrm{d}r = \\
&= \tfrac{1}{2} N_{n_i\zeta_i} N_{n_j\zeta_j} \left((n_i n_j + l(l+1)) \frac{(n_i+n_j-2)!}{(\zeta_i+\zeta_j)^{n_i+n_j-1}} - (n_i\zeta_j + n_j\zeta_i) \frac{(n_i+n_j-1)!}{(\zeta_i+\zeta_j)^{n_i+n_j}} + \zeta_i\zeta_j \frac{(n_i+n_j)!}{(\zeta_i+\zeta_j)^{n_i+n_j+1}} \right)
\end{aligned}
$$

Two particle

In this section we also need the following integral:

$$
\int_u^\infty r^n e^{-\zeta r} \mathrm{d}r = \frac{1}{\zeta^{n+1}} \int_{\zeta u}^\infty x^n e^{-x} \mathrm{d}x = \frac{\Gamma(n+1, \zeta u)}{\zeta^{n+1}} = \frac{n!}{\zeta^{n+1}} e^{-\zeta u} \sum_{\nu=0}^{n} \frac{u^\nu \zeta^\nu}{\nu!}
$$

where

$$
\Gamma(n, x) = \int_x^\infty t^{n-1} e^{-t} \mathrm{d}t = (n-1)! e^{-x} \sum_{\nu=0}^{n-1} \frac{x^\nu}{\nu!}
$$

is the incomplete gamma function.

The Slater integral is

$$\begin{aligned}
R^k(i,j,k,l) = \int_0^\infty \int_0^\infty \frac{r_<^k}{r_>^{k+1}} P_i(r)P_k(r)P_j(r')P_l(r')\mathrm{d}r\mathrm{d}r' = \\
= \int_0^\infty \mathrm{d}r \int_0^r \mathrm{d}r' \frac{r'^k}{r^{k+1}} P_i(r)P_k(r)P_j(r')P_l(r') + \\
+ \int_0^\infty \mathrm{d}r \int_r^\infty \mathrm{d}r' \frac{r^k}{r'^{k+1}} P_i(r)P_k(r)P_j(r')P_l(r') = \\
= \int_0^\infty \mathrm{d}r \int_0^r \mathrm{d}r' \frac{r'^k}{r^{k+1}} P_i(r)P_k(r)P_j(r')P_l(r') + \\
+ \int_0^\infty \mathrm{d}r' \int_0^{r'} \mathrm{d}r \frac{r^k}{r'^{k+1}} P_i(r)P_k(r)P_j(r')P_l(r') = \\
= \int_0^\infty \mathrm{d}r \int_0^r \mathrm{d}r' \frac{r'^k}{r^{k+1}} P_i(r)P_k(r)P_j(r')P_l(r') + \\
+ \int_0^\infty \mathrm{d}r \int_0^r \mathrm{d}r' \frac{r'^k}{r^{k+1}} P_i(r')P_k(r')P_j(r)P_l(r) = \\
= R_\triangle^k(i,j,k,l) + R_\triangle^k(j,i,l,k)
\end{aligned}$$

where $R_\triangle^k(i,j,k,l)$ is the integral over the lower triangle (assuming r is the x-axis and r' is the y-axis), that is $r > r'$:

$$\begin{aligned}
R_\triangle^k(i,j,k,l) = \int_0^\infty \mathrm{d}r \int_0^r \mathrm{d}r' \frac{r'^k}{r^{k+1}} P_i(r)P_k(r)P_j(r')P_l(r') = \\
= \int_0^\infty \mathrm{d}r' \int_{r'}^\infty \mathrm{d}r \frac{r'^k}{r^{k+1}} P_i(r)P_k(r)P_j(r')P_l(r') = \\
= N_{n_i\zeta_i} N_{n_j\zeta_j} N_{n_k\zeta_k} N_{n_l\zeta_l} \int_0^\infty \mathrm{d}r' \int_{r'}^\infty \mathrm{d}r \frac{r'^k}{r^{k+1}} r^{n_i+n_k} r'^{n_j+n_l} e^{-(\zeta_i+\zeta_k)r} e^{-(\zeta_j+\zeta_l)r'} = \\
= N_{n_i\zeta_i} N_{n_j\zeta_j} N_{n_k\zeta_k} N_{n_l\zeta_l} \int_0^\infty \mathrm{d}r' r'^k r'^{n_j+n_l} e^{-(\zeta_j+\zeta_l)r'} \frac{(n_i+n_k-k-1)!}{(\zeta_i+\zeta_k)^{n_i+n_k-k}} e^{-(\zeta_i+\zeta_k)r'} \sum_{\nu=0}^{n_i+n_k-k-1} \frac{r'^\nu (\zeta_i+\zeta_k)^\nu}{\nu!} = \\
= N_{n_i\zeta_i} N_{n_j\zeta_j} N_{n_k\zeta_k} N_{n_l\zeta_l} \frac{(n_i+n_k-k-1)!}{(\zeta_i+\zeta_k)^{n_i+n_k-k}} \sum_{\nu=0}^{n_i+n_k-k-1} \frac{(n_j+n_l+k+\nu)!(\zeta_i+\zeta_k)^\nu}{\nu!(\zeta_i+\zeta_j+\zeta_k+\zeta_l)^{n_j+n_l+k+\nu+1}} = \\
= N_{n_i\zeta_i} N_{n_j\zeta_j} N_{n_k\zeta_k} N_{n_l\zeta_l} \frac{(n_i+n_k-k-1)!}{(\zeta_i+\zeta_k)^{n_i+n_k-k}} H_{ijkl}^k
\end{aligned}$$

where:

$$H_{ijkl}^k = \sum_{\nu=0}^{n_i+n_k-k-1} \frac{(n_j+n_l+k+\nu)!(\zeta_i+\zeta_k)^\nu}{\nu!(\zeta_i+\zeta_j+\zeta_k+\zeta_l)^{n_j+n_l+k+\nu+1}}$$

Much more tedious method is the following:

$$\int_0^u r^n e^{-\zeta r}\mathrm{d}r = \int_0^\infty r^n e^{-\zeta r}\mathrm{d}r - \int_u^\infty r^n e^{-\zeta r}\mathrm{d}r = \frac{n!}{\zeta^{n+1}} \left(1 - e^{-\zeta u} \sum_{\nu=0}^{n} \frac{u^\nu \zeta^\nu}{\nu!}\right)$$

The Slater integral is given by:

$$\begin{aligned}
R^k(i,j,k,l) = \int_0^\infty \frac{r_<^k}{r_>^{k+1}} P_i(r)P_k(r)P_j(r')P_l(r')\mathrm{d}r\mathrm{d}r' = \\
= \int_0^\infty P_i(r)P_k(r) \frac{Y^k(P_jP_l,r)}{r}\mathrm{d}r
\end{aligned}$$

where

$$Y^k(f(r), r) = r\int_0^\infty \frac{r_<^k}{r_>^{k+1}} f(r')\mathrm{d}r' = \frac{1}{r^k}\int_0^r r'^k f(r')\mathrm{d}r' + r^{k+1}\int_r^\infty \frac{1}{r'^{k+1}} f(r')\mathrm{d}r'$$

and we get:

$$\begin{aligned}
Y^k(P_j(r)P_l(r), r) &= \frac{1}{r^k}\int_0^r r'^k P_j(r')P_l(r')\mathrm{d}r' + r^{k+1}\int_r^\infty \frac{1}{r'^{k+1}} P_j(r')P_l(r')\mathrm{d}r' = \\
&= \frac{N_{n_j\zeta_j}N_{n_l\zeta_l}}{r^k}\int_0^r r'^k r'^{n_j} e^{-\zeta_j r'} r'^{n_l} e^{-\zeta_l r'}\mathrm{d}r' + N_{n_j\zeta_j}N_{n_l\zeta_l} r^{k+1}\int_r^\infty \frac{1}{r'^{k+1}} r'^{n_j} e^{-\zeta_j r'} r'^{n_l} e^{-\zeta_l r'}\mathrm{d}r' = \\
&= \frac{N_{n_j\zeta_j}N_{n_l\zeta_l}}{r^k}\int_0^r r'^{n_j+n_l+k} e^{-(\zeta_j+\zeta_l)r'}\mathrm{d}r' + N_{n_j\zeta_j}N_{n_l\zeta_l} r^{k+1}\int_r^\infty r'^{n_j+n_l-k-1} e^{-(\zeta_j+\zeta_l)r'}\mathrm{d}r' = \\
&= \frac{N_{n_j\zeta_j}N_{n_l\zeta_l}}{r^k}\frac{(n_j+n_l+k)!}{(\zeta_j+\zeta_l)^{n_j+n_l+k+1}}\left(1 - e^{-(\zeta_j+\zeta_l)r}\sum_{\nu=0}^{n_j+n_l+k}\frac{r^\nu(\zeta_j+\zeta_l)^\nu}{\nu!}\right) + \\
&\quad + N_{n_j\zeta_j}N_{n_l\zeta_l} r^{k+1}\frac{(n_j+n_l-k-1)!}{(\zeta_j+\zeta_l)^{n_j+n_l-k}} e^{-(\zeta_j+\zeta_l)r}\sum_{\nu=0}^{n_j+n_l-k-1}\frac{r^\nu(\zeta_j+\zeta_l)^\nu}{\nu!}
\end{aligned}$$

Putting everything together we get:

$$\begin{aligned}
R^k(i,j,k,l) &= N_{n_i\zeta_i}N_{n_j\zeta_j}N_{n_k\zeta_k}N_{n_l\zeta_l}\int_0^\infty r^{n_i} e^{-\zeta_i r} r^{n_k} e^{-\zeta_k r}\frac{1}{r}\left(\frac{1}{r^k}\frac{(n_j+n_l+k)!}{(\zeta_j+\zeta_l)^{n_j+n_l+k+1}}\left(1 - e^{-(\zeta_j+\zeta_l)r}\sum_{\nu=0}^{n_j+n_l+k}\frac{r^\nu(\zeta_j+\zeta_l)^\nu}{\nu!}\right) + \right. \\
&\quad \left. + r^{k+1}\frac{(n_j+n_l-k-1)!}{(\zeta_j+\zeta_l)^{n_j+n_l-k}} e^{-(\zeta_j+\zeta_l)r}\sum_{\nu=0}^{n_j+n_l-k-1}\frac{r^\nu(\zeta_j+\zeta_l)^\nu}{\nu!}\right)\mathrm{d}r = \\
&= N_{n_i\zeta_i}N_{n_j\zeta_j}N_{n_k\zeta_k}N_{n_l\zeta_l}\left(\frac{(n_j+n_l+k)!}{(\zeta_j+\zeta_l)^{n_j+n_l+k+1}}\int_0^\infty r^{n_i+n_k-k-1} e^{-(\zeta_i+\zeta_k)r}\left(1 - e^{-(\zeta_j+\zeta_l)r}\sum_{\nu=0}^{n_j+n_l+k}\frac{r^\nu(\zeta_j+\zeta_l)^\nu}{\nu!}\right)\mathrm{d}r + \right. \\
&\quad \left. + \int_0^\infty e^{-(\zeta_i+\zeta_j+\zeta_k+\zeta_l)r}\frac{(n_j+n_l-k-1)!}{(\zeta_j+\zeta_l)^{n_j+n_l-k}}\sum_{\nu=0}^{n_j+n_l-k-1}\frac{r^{n_i+n_k+k+\nu}(\zeta_j+\zeta_l)^\nu}{\nu!}\mathrm{d}r\right) = \\
&= N_{n_i\zeta_i}N_{n_j\zeta_j}N_{n_k\zeta_k}N_{n_l\zeta_l}\left(\frac{(n_j+n_l+k)!}{(\zeta_j+\zeta_l)^{n_j+n_l+k+1}}\left(\frac{(n_i+n_k-k-1)!}{(\zeta_i+\zeta_k)^{n_i+n_k-k}} - \sum_{\nu=0}^{n_j+n_l+k}\frac{(\zeta_j+\zeta_l)^\nu(n_i+n_k-k+\nu-1)!}{\nu!(\zeta_i+\zeta_j+\zeta_k+\zeta_l)^{n_i+n_k-k+\nu}}\right) + \right. \\
&\quad \left. + \frac{(n_j+n_l-k-1)!}{(\zeta_j+\zeta_l)^{n_j+n_l-k}}\sum_{\nu=0}^{n_j+n_l-k-1}\frac{(n_i+n_k+k+\nu)!(\zeta_j+\zeta_l)^\nu}{\nu!(\zeta_i+\zeta_j+\zeta_k+\zeta_l)^{n_i+n_k+k+\nu+1}}\right) = \\
&= N_{n_i\zeta_i}N_{n_j\zeta_j}N_{n_k\zeta_k}N_{n_l\zeta_l}\left(\frac{(n_j+n_l+k)!}{(\zeta_j+\zeta_l)^{n_j+n_l+k+1}}\left(\frac{(n_i+n_k-k-1)!}{(\zeta_i+\zeta_k)^{n_i+n_k-k}} - H^{-k-1}_{jilk}\right) + \right. \\
&\quad \left. + \frac{(n_j+n_l-k-1)!}{(\zeta_j+\zeta_l)^{n_j+n_l-k}}H^k_{jilk}\right) = \\
&= N_{n_i\zeta_i}N_{n_j\zeta_j}N_{n_k\zeta_k}N_{n_l\zeta_l}\left(\frac{(n_i+n_k-k-1)!}{(\zeta_i+\zeta_k)^{n_i+n_k-k}}H^k_{ijkl} + \frac{(n_j+n_l-k-1)!}{(\zeta_j+\zeta_l)^{n_j+n_l-k}}H^k_{jilk}\right)
\end{aligned}$$

where:

$$H^k_{ijkl} = \sum_{\nu=0}^{n_i+n_k-k-1}\frac{(n_j+n_l+k+\nu)!(\zeta_i+\zeta_k)^\nu}{\nu!(\zeta_i+\zeta_j+\zeta_k+\zeta_l)^{n_j+n_l+k+\nu+1}}$$

8.10.6 Gaussian Type Orbitals (GTO)

In this section we express the matrix elements in the GTO basis. It turns out that all integrals that we need can be expressed in terms of the following simple integral (where $n, \zeta \geq 0$):

$$\int_0^\infty r^n e^{-\zeta r^2} \mathrm{d}r = \frac{1}{\sqrt{\zeta^{n+1}}} \int_0^\infty x^n e^{-x^2} \mathrm{d}x = \frac{1}{2\sqrt{\zeta^{n+1}}} \int_0^\infty t^{\frac{n-1}{2}} e^{-t} \mathrm{d}t = \frac{\Gamma(\frac{n+1}{2})}{2\zeta^{\frac{n+1}{2}}} = \\ = \begin{cases} (n-1)!! \sqrt{\frac{\pi}{2(2\zeta)^{n+1}}} & \text{for even } n \\ \frac{(\frac{n-1}{2})!}{2\sqrt{\zeta^{n+1}}} & \text{for odd } n \end{cases} \tag{8.10.6.1}$$

The GTO basis function for the radial Schrödinger equation for $P(r)$ is:

$$P_{n\zeta}(r) = N_{n\zeta} r^n e^{-\zeta r^2} \tag{8.10.6.2}$$

However, unlike STO, the GTO functions must satisfy the condition $n = l+2i+1$ where $i = 0, 1, 2, \ldots$ (this condition is later used in (8.10.6.1) to determine whether n is even or odd). The normalization constant $N_{n\zeta}$ is such that the GTO orbital is normalized as the radial wavefunction $P(r)$:

$$1 = \int_0^\infty P_{n\zeta}^2(r) \mathrm{d}r = N_{n\zeta}^2 \int_0^\infty r^{2n} e^{-2\zeta r^2} \mathrm{d}r = N_{n\zeta}^2 (2n-1)!! \sqrt{\frac{\pi}{2(4\zeta)^{2n+1}}}$$

from which we get:

$$N_{n\zeta} = \sqrt{\frac{1}{(2n-1)!!} \sqrt{\frac{2(4\zeta)^{2n+1}}{\pi}}}$$

Note that for $R(r) = \frac{P(r)}{r}$ we get the following GTO basis function:

$$R_{n\zeta}(r) = \frac{P_{n\zeta}(r)}{r} = N_{n\zeta} r^{n-1} e^{-\zeta r^2} \tag{8.10.6.3}$$

One uses either (8.10.5.2) or (8.10.5.3) depending on whether one solves the radial Schrödinger equation for P or for $R = \frac{P}{r}$.

Overlap

$$S_{ij} = \int P_{n_i\zeta_i}(r) P_{n_j\zeta_j}(r) \mathrm{d}r = \\ = \int N_{n_i\zeta_i} r^{n_i} e^{-\zeta_i r^2} N_{n_j\zeta_j} r^{n_j} e^{-\zeta_j r^2} \mathrm{d}r = \\ = N_{n_i\zeta_i} N_{n_j\zeta_j} \int r^{n_i+n_j} e^{-(\zeta_i+\zeta_j) r^2} \mathrm{d}r = \\ = N_{n_i\zeta_i} N_{n_j\zeta_j} \frac{\Gamma(\frac{n_i+n_j+1}{2})}{2(\zeta_i+\zeta_j)^{\frac{n_i+n_j+1}{2}}} = \\ = N_{n_i\zeta_i} N_{n_j\zeta_j} (n_i+n_j-1)!! \sqrt{\frac{\pi}{2(2\zeta_i+2\zeta_j)^{n_i+n_j+1}}}$$

Potential

$$V_{ij} = \int P_{n_i\zeta_i}(r)\left(-\frac{Z}{r}\right)P_{n_j\zeta_j}(r)\mathrm{d}r =$$
$$= \int N_{n_i\zeta_i} r^{n_i} e^{-\zeta_i r^2}\left(-\frac{Z}{r}\right)N_{n_j\zeta_j} r^{n_j} e^{-\zeta_j r^2}\mathrm{d}r =$$
$$= -ZN_{n_i\zeta_i}N_{n_j\zeta_j}\int r^{n_i+n_j-1}e^{-(\zeta_i+\zeta_j)r^2}\mathrm{d}r =$$
$$= -ZN_{n_i\zeta_i}N_{n_j\zeta_j}\frac{\Gamma(\frac{n_i+n_j}{2})}{2(\zeta_i+\zeta_j)^{\frac{n_i+n_j}{2}}} =$$
$$= -ZN_{n_i\zeta_i}N_{n_j\zeta_j}\frac{\left(\frac{n_i+n_j-2}{2}\right)!}{2\sqrt{(\zeta_i+\zeta_j)^{n_i+n_j}}}$$

Kinetic

$$T_{ij} = \int\left(\tfrac{1}{2}P'_{n_i\zeta_i}(r)P'_{n_j\zeta_j}(r) + P_{n_i\zeta_i}(r)\frac{l(l+1)}{2r^2}P_{n_j\zeta_j}(r)\right)\mathrm{d}r =$$
$$= \tfrac{1}{2}N_{n_i\zeta_i}N_{n_j\zeta_j}\int\left(\frac{\mathrm{d}}{\mathrm{d}r}(r^{n_i}e^{-\zeta_i r^2})\frac{\mathrm{d}}{\mathrm{d}r}(r^{n_j}e^{-\zeta_j r^2}) + r^{n_i}e^{-\zeta_i r^2}\frac{l(l+1)}{r^2}r^{n_j}e^{-\zeta_j r^2}\right)\mathrm{d}r =$$
$$= \tfrac{1}{2}N_{n_i\zeta_i}N_{n_j\zeta_j}\int\left((n_i r^{n_i-1}e^{-\zeta_i r^2} - 2\zeta_i r^{n_i+1}e^{-\zeta_i r^2})(n_j r^{n_j-1}e^{-\zeta_j r^2} - 2\zeta_j r^{n_j+1}e^{-\zeta_j r^2}) + l(l+1)r^{n_i+n_j-2}e^{-(\zeta_i+\zeta_j)r^2}\right)\mathrm{d}r =$$
$$= \tfrac{1}{2}N_{n_i\zeta_i}N_{n_j\zeta_j}\int\left(\left(\frac{n_i n_j}{r^2} - 2(n_i\zeta_j + n_j\zeta_i) + 4\zeta_i\zeta_j r^2\right)r^{n_i+n_j}e^{-(\zeta_i+\zeta_j)r^2} + l(l+1)r^{n_i+n_j-2}e^{-(\zeta_i+\zeta_j)r^2}\right)\mathrm{d}r =$$
$$= \tfrac{1}{2}N_{n_i\zeta_i}N_{n_j\zeta_j}\int\left((n_i n_j + l(l+1))r^{n_i+n_j-2}e^{-(\zeta_i+\zeta_j)r^2} - 2(n_i\zeta_j + n_j\zeta_i)r^{n_i+n_j}e^{-(\zeta_i+\zeta_j)r^2} + 4\zeta_i\zeta_j r^{n_i+n_j+2}e^{-(\zeta_i+\zeta_j)r^2}\right)\mathrm{d}r =$$
$$= \tfrac{1}{2}N_{n_i\zeta_i}N_{n_j\zeta_j}\left((n_i n_j + l(l+1))\frac{\Gamma(\frac{n_i+n_j-1}{2})}{2(\zeta_i+\zeta_j)^{\frac{n_i+n_j-1}{2}}} - 2(n_i\zeta_j + n_j\zeta_i)\frac{\Gamma(\frac{n_i+n_j+1}{2})}{2(\zeta_i+\zeta_j)^{\frac{n_i+n_j+1}{2}}} + 4\zeta_i\zeta_j\frac{\Gamma(\frac{n_i+n_j+3}{2})}{2(\zeta_i+\zeta_j)^{\frac{n_i+n_j+3}{2}}}\right)\mathrm{d}r =$$
$$= \tfrac{1}{2}N_{n_i\zeta_i}N_{n_j\zeta_j}\Bigg((n_i n_j + l(l+1))(n_i+n_j-3)!!\sqrt{\frac{\pi}{2(2\zeta_i+2\zeta_j)^{n_i+n_j-1}}} +$$
$$-2(n_i\zeta_j + n_j\zeta_i)(n_i+n_j-1)!!\sqrt{\frac{\pi}{2(2\zeta_i+2\zeta_j)^{n_i+n_j+1}}} +$$
$$+4\zeta_i\zeta_j(n_i+n_j+1)!!\sqrt{\frac{\pi}{2(2\zeta_i+2\zeta_j)^{n_i+n_j+3}}}\Bigg)$$

Two particle

We will need the following integral:

$$\int_u^\infty r^n e^{-\zeta r^2}\mathrm{d}r = \frac{1}{2\zeta^{\frac{n+1}{2}}}\int_{\zeta u^2}^\infty x^{\frac{n-1}{2}}e^{-x}\mathrm{d}x = \frac{\Gamma(\frac{n+1}{2}, \zeta u^2)}{2\zeta^{\frac{n+1}{2}}}$$

Just like for STO, we get:

$$R^k(i,j,k,l) = R^k_\triangle(i,j,k,l) + R^k_\triangle(j,i,l,k)$$

where $R^k_\triangle(i,j,k,l)$ is the integral over the lower triangle (assuming r is the x-axis and r' is the y-axis), that is $r > r'$:

$$R^k_\triangle(i,j,k,l) = \int_0^\infty \mathrm{d}r \int_0^r \mathrm{d}r' \frac{r'^k}{r^{k+1}} P_i(r)P_k(r)P_j(r')P_l(r') =$$
$$= \int_0^\infty \mathrm{d}r' \int_{r'}^\infty \mathrm{d}r \frac{r'^k}{r^{k+1}} P_i(r)P_k(r)P_j(r')P_l(r') =$$
$$= N_{n_i\zeta_i} N_{n_j\zeta_j} N_{n_k\zeta_k} N_{n_l\zeta_l} \int_0^\infty \mathrm{d}r' \int_{r'}^\infty \mathrm{d}r \frac{r'^k}{r^{k+1}} r^{n_i+n_k} r'^{n_j+n_l} e^{-(\zeta_i+\zeta_k)r^2} e^{-(\zeta_j+\zeta_l)r'^2} =$$
$$= N_{n_i\zeta_i} N_{n_j\zeta_j} N_{n_k\zeta_k} N_{n_l\zeta_l} \int_0^\infty \mathrm{d}r' r'^k r'^{n_j+n_l} e^{-(\zeta_j+\zeta_l)r'^2} \frac{\Gamma(\frac{n_i+n_k-k}{2}, (\zeta_i+\zeta_k)r'^2)}{2(\zeta_i+\zeta_k)^{\frac{n_i+n_k-k}{2}}} =$$
$$= N_{n_i\zeta_i} N_{n_j\zeta_j} N_{n_k\zeta_k} N_{n_l\zeta_l} \int_0^\infty \mathrm{d}r' r'^k r'^{n_j+n_l} e^{-(\zeta_i+\zeta_j+\zeta_k+\zeta_l)r'^2} \frac{(\frac{n_i+n_k-k}{2}-1)!}{2(\zeta_i+\zeta_k)^{\frac{n_i+n_k-k}{2}}} \sum_{\nu=0}^{\frac{n_i+n_k-k-2}{2}} \frac{(\zeta_i+\zeta_k)^\nu r'^{2\nu}}{\nu!} =$$
$$= N_{n_i\zeta_i} N_{n_j\zeta_j} N_{n_k\zeta_k} N_{n_l\zeta_l} \frac{(\frac{n_i+n_k-k}{2}-1)!}{2(\zeta_i+\zeta_k)^{\frac{n_i+n_k-k}{2}}} \sum_{\nu=0}^{\frac{n_i+n_k-k-2}{2}} \frac{(\zeta_i+\zeta_k)^\nu}{\nu!} \frac{\Gamma(\frac{n_j+n_l+k+2\nu+1}{2})}{2(\zeta_i+\zeta_j+\zeta_k+\zeta_l)^{\frac{n_j+n_l+k+2\nu+1}{2}}} =$$
$$= N_{n_i\zeta_i} N_{n_j\zeta_j} N_{n_k\zeta_k} N_{n_l\zeta_l} \frac{(\frac{n_i+n_k-k}{2}-1)!}{2(\zeta_i+\zeta_k)^{\frac{n_i+n_k-k}{2}}} \sum_{\nu=0}^{\frac{n_i+n_k-k-2}{2}} \frac{(\zeta_i+\zeta_k)^\nu}{\nu!} \frac{(n_j+n_l+k+2\nu-1)!!\sqrt{\pi}}{2^{\frac{n_j+n_l+k+2\nu+2}{2}}(\zeta_i+\zeta_j+\zeta_k+\zeta_l)^{\frac{n_j+n_l+k+2\nu+1}{2}}} =$$
$$= N_{n_i\zeta_i} N_{n_j\zeta_j} N_{n_k\zeta_k} N_{n_l\zeta_l} \frac{\sqrt{\pi}}{2^{p+2}\sqrt{(\zeta_i+\zeta_j+\zeta_k+\zeta_l)^{2p+1}}} H^k_{ijkl}$$

where $p = \frac{n_i+n_j+n_k+n_l-2}{2}$ and:

$$H^k_{ijkl} = \sum_{\nu=0}^{\frac{n_i+n_k-k}{2}} \frac{(2p-2\nu-1)!!\,(\frac{n_i+n_k-k}{2})!2^\nu(\zeta_i+\zeta_j+\zeta_k+\zeta_l)^\nu}{(n_i+n_k-\nu)!(\lambda_i+\lambda_k)^{1+\nu}}$$

8.10.7 Exchange Integral in Spherical Symmetry

Let's calculate the exchange integral

$$\int \frac{\psi_i^*(\mathbf{x})\psi_j(\mathbf{x})\psi_j^*(\mathbf{x}')\psi_i(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\mathrm{d}^3x\mathrm{d}^3x'$$

for the particular choice of the functions ψ:

$$\psi_i(\mathbf{x}) = \frac{P_{nl}(r)}{r}Y_{lm}(\Omega)$$
$$\psi_j(\mathbf{x}) = \frac{P_{n'l'}(r)}{r}Y_{l'm'}(\Omega)$$

We use multipole expansion:

$$\frac{1}{|\mathbf{x}-\mathbf{x}'|} = \sum_{k,q} \frac{r_<^k}{r_>^{k+1}} \frac{4\pi}{2k+1} Y_{kq}(\Omega)Y^*_{kq}(\Omega')$$

And we get:

$$\int \frac{\psi_i^*(\mathbf{x})\psi_j(\mathbf{x})\psi_j^*(\mathbf{x}')\psi_i(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\mathrm{d}^3x\mathrm{d}^3x' =$$
$$= \int \frac{P_{nl}(r)}{r}Y_{lm}^*(\Omega)\frac{P_{n'l'}(r)}{r}Y_{l'm'}(\Omega)\frac{P_{n'l'}(r')}{r'}Y_{l'm'}^*(\Omega')\frac{P_{nl}(r')}{r'}Y_{lm}(\Omega')$$
$$\sum_{k,q}\frac{r_<^k}{r_>^{k+1}}\frac{4\pi}{2k+1}Y_{kq}(\Omega)Y_{kq}^*(\Omega')r^2r'^2\mathrm{d}r\mathrm{d}r'\mathrm{d}\Omega\mathrm{d}\Omega' =$$
$$= \sum_{k,q}\int Y_{lm}^*(\Omega)Y_{l'm'}(\Omega)Y_{kq}(\Omega)\mathrm{d}\Omega\int Y_{l'm'}^*(\Omega')Y_{lm}(\Omega')Y_{kq}^*(\Omega')\mathrm{d}\Omega'$$
$$\int \frac{r_<^k}{r_>^{k+1}}\frac{4\pi}{2k+1}P_{nl}(r)P_{n'l'}(r)P_{n'l'}(r')P_{nl}(r')\mathrm{d}r\mathrm{d}r' =$$
$$= \sum_{k,q}\int Y_{lm}^*(\Omega)Y_{l'm'}(\Omega)Y_{kq}(\Omega)\mathrm{d}\Omega(-1)^{m+m'+q}\int Y_{l',-m'}(\Omega')Y_{l,-m}^*(\Omega')Y_{k,-q}(\Omega')\mathrm{d}\Omega'$$
$$\int \frac{r_<^k}{r_>^{k+1}}\frac{4\pi}{2k+1}P_{nl}(r)P_{n'l'}(r)P_{n'l'}(r')P_{nl}(r')\mathrm{d}r\mathrm{d}r' =$$
$$= \sum_k c^k(l,m,l',m')\sqrt{\frac{2k+1}{4\pi}}(-1)^{m+m'+m-m'}c^k(l,-m,l',-m')\sqrt{\frac{2k+1}{4\pi}}$$
$$\int \frac{r_<^k}{r_>^{k+1}}\frac{4\pi}{2k+1}P_{nl}(r)P_{n'l'}(r)P_{n'l'}(r')P_{nl}(r')\mathrm{d}r\mathrm{d}r' =$$
$$= \sum_k c^k(l,m,l',m')c^k(l,-m,l',-m')\int \frac{r_<^k}{r_>^{k+1}}P_{nl}(r)P_{n'l'}(r)P_{n'l'}(r')P_{nl}(r')\mathrm{d}r\mathrm{d}r' =$$
$$= \sum_k c^k(l,m,l',m')c^k(l,m,l',m')\int \frac{r_<^k}{r_>^{k+1}}P_{nl}(r)P_{n'l'}(r)P_{n'l'}(r')P_{nl}(r')\mathrm{d}r\mathrm{d}r' =$$
$$= \sum_{k=|l-l'|}^{l+l'}\left(c^k(l,m,l',m')\right)^2\int \frac{r_<^k}{r_>^{k+1}}P_{nl}(r)P_{n'l'}(r)P_{n'l'}(r')P_{nl}(r')\mathrm{d}r\mathrm{d}r'$$

8.10.8 Occupation Numbers

We have a sum over N electron states like this:

$$\sum_{i=1}^{N}A_i(\mathbf{x}) = \sum_{nlms}A_{nlms}(\mathbf{x})$$

where $A_{nlms}(\mathbf{x})$ are some functions that depend on the state numbers (for example squares of the wavefunctions). Then there are two options — either there is a way to sum over the m and s degrees of freedom, so that the sum can be written exactly as:

$$\sum_{nlms}A_{nlms}(\mathbf{x}) = \sum_{nlms}B_{nl}(\mathbf{x})$$

where B_{nl} (that don't depend on m and s) will in general be different to A_{nlms}, but the sum will be the same. Or we have to approximate the sum (for example by averaging over the angles, or in some other way) as:

$$\sum_{nlms}A_{nlms}(\mathbf{x}) \to \sum_{nlms}B_{nl}(\mathbf{x})$$

In either case, the occupation numbers f_{nl} are simply the number of times the functions $B_{nl}(\mathbf{x})$ appear in the sum for the given n and l:

$$\sum_{nlms} B_{nl}(\mathbf{x}) = \sum_{nl} f_{nl} B_{nl}(\mathbf{x})$$

So for closed shells atoms, it is always:

$$f_{nl} = 2(2l+1)$$

because there are two spins, and $2l+1$ possibilities for m, for open shell atoms, f_{nl} is anything between 0 and $2l+1$.

Example I

As an example, let's say that after some calculation for closed shell systems we get exactly:

$$\sum_{nlms} A_{nlms}(\mathbf{x}) = \sum_{nl} 2(2l+1) B_{nl}(\mathbf{x})$$

Then because there are exactly $2(2l+1)$ states in the nl shell, we write the above as:

$$\sum_{nlms} A_{nlms}(\mathbf{x}) = \sum_{nl} 2(2l+1) B_{nl}(\mathbf{x}) = \sum_{nl} f_{nl} B_{nl}(\mathbf{x})$$

Then we do similar calculation for the open shell system, and we have to use some approximations to get the following formula, where the $B_{nl}(\mathbf{x})$ happen to be exactly the same as for the closed shell system:

$$\sum_{nlms} A_{nlms}(\mathbf{x}) \to \sum_{nlm} 2 B_{nl}(\mathbf{x})$$

Then we denote by f_{nl} the number of electrons in the shell nl (at least one of them will be open, for which nl we have $f_{nl} < 2(2l+1)$), and we can write the above as:

$$\sum_{nlms} A_{nlms}(\mathbf{x}) \to \sum_{nlm} 2 B_{nl}(\mathbf{x}) = \sum_{nl} f_{nl} B_{nl}(\mathbf{x})$$

Example II

The usual chemical occupation numbers for the Uranium atom are:

$$\begin{aligned}
f_{1l} &= 2(2l+1) \\
f_{2l} &= 2(2l+1) \\
f_{3l} &= 2(2l+1) \\
f_{4l} &= 2(2l+1) \\
f_{5l} &= 2(2l+1) \qquad \text{for } l \le 2 \\
f_{53} &= 3 \\
f_{60} &= 2 \\
f_{61} &= 6 \\
f_{62} &= 1 \\
f_{70} &= 2
\end{aligned}$$

So the $n=5$, $l=3$ and $n=6$, $l=2$ shells are open, all others are closed. By summing all these f_{nl}, we get 92 states as expected:

$$\sum_{nl} f_{nl} = 2 + (2+6) + (2+6+10) + (2+6+10+14) + (2+6+10) + \\ +3+2+6+1+2 = 92$$

Code:

```
def f_nl(n, l):
    if n < 5 or (n == 5 and l <= 2):
        return 2*(2*l+1)
    else:
        d = {
            (5, 3): 3,
            (6, 0): 2,
            (6, 1): 6,
            (6, 2): 1,
            (7, 0): 2,
            }
        if (n, l) in d:
            return d[n, l]
        else:
            return 0

print "Sum f_nl =", sum([f_nl(n, l) for n in range(8) for l in range(n)])
```

prints:

```
Sum f_nl = 92
```

8.10.9 Hartree Screening Functions

Hartree screening function $Y^k(f,r)$ is defined as:

$$Y^k(f,r) = r\int_0^\infty \frac{r_<^k}{r_>^{k+1}} f(r')\mathrm{d}r'$$

and it occurs in many formulas in the Hartree Fock theory, so this section shows how to calculate it. It depends on k and a function $f(r)$.

We first do the integral:

$$Y^k(f,r) = r\int_0^\infty \frac{r_<^k}{r_>^{k+1}} f(r')\mathrm{d}r' = r\int_0^r \frac{r'^k}{r^{k+1}} f(r')\mathrm{d}r' + r\int_r^\infty \frac{r^k}{r'^{k+1}} f(r')\mathrm{d}r' =$$
$$= \frac{1}{r^k}\int_0^r x^k f(x)\mathrm{d}x + r^{k+1}\int_r^\infty \frac{1}{x^{k+1}} f(x)\mathrm{d}x = Z^k(r) + r^{k+1}\int_r^\infty \frac{1}{x^{k+1}} f(x)\mathrm{d}x$$

where:

$$Z^k(r) = \frac{1}{r^k}\int_0^r x^k f(x)\mathrm{d}x$$
$$\frac{\mathrm{d}Z^k(r)}{\mathrm{d}r} = -\frac{k}{r}Z^k(r) + f(r)$$
$$Z^k(0) = 0$$

Now we differentiate $Y^k(r)$:

$$\begin{aligned}
\frac{\mathrm{d}Y^k(r)}{\mathrm{d}r} &= \frac{\mathrm{d}Z^k(r)}{\mathrm{d}r} + \frac{k+1}{r} r^{k+1} \int_r^\infty \frac{1}{x^{k+1}} f(x)\mathrm{d}x - f(r) = \\
&= -\frac{k}{r} Z^k(r) + f(r) + \frac{k+1}{r} r^{k+1} \int_r^\infty \frac{1}{x^{k+1}} f(x)\mathrm{d}x - f(r) = \\
&= -\frac{k}{r} Z^k(r) + \frac{k+1}{r} r^{k+1} \int_r^\infty \frac{1}{x^{k+1}} f(x)\mathrm{d}x = \\
&= -\frac{k}{r} Z^k(r) + \frac{k+1}{r} (Y^k(r) - Z^k(r)) = \\
&= -\frac{2k+1}{r} Z^k(r) + \frac{k+1}{r} Y^k(r)
\end{aligned}$$

Also $Y^k(\infty) = Z^k(\infty)$, so we get the following set of first order differential equations with boundary conditions:

$$\begin{aligned}
\left(\frac{\mathrm{d}}{\mathrm{d}r} - \frac{k+1}{r}\right) Y^k(r) &= -\frac{2k+1}{r} Z^k(r) \\
\left(\frac{\mathrm{d}}{\mathrm{d}r} + \frac{k}{r}\right) Z^k(r) &= f(r) \\
Y^k(\infty) &= Z^k(\infty) \\
Z^k(0) &= 0
\end{aligned}$$

One way to calculate the Hartree screening function is to integrate the second equation from the left using the boundary condition $Z^k(0) = 0$ and then integrate the first equation from the right, using the boundary condition $Y^k(\infty) = Z^k(\infty)$.

Another way is to obtain one second order equation. Expressing Z^k from the first equation:

$$\begin{aligned}
Z^k(r) &= -\frac{r}{2k+1} \left(\frac{\mathrm{d}}{\mathrm{d}r} - \frac{k+1}{r}\right) Y^k(r) = \\
&= -\frac{r}{2k+1} \frac{\mathrm{d}Y^k(r)}{\mathrm{d}r} + \frac{k+1}{2k+1} Y^k(r)
\end{aligned}$$

and substituting into the second equation we get:

$$\begin{aligned}
-\left(\frac{\mathrm{d}}{\mathrm{d}r} + \frac{k}{r}\right) \left(\frac{r}{2k+1} \frac{\mathrm{d}Y^k(r)}{\mathrm{d}r} + \frac{k+1}{2k+1} Y^k\right) &= f(r) \\
-\frac{r}{2k+1} \left(\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{k(k+1)}{r^2}\right) Y^k(r) &= f(r) \\
\left(-\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{k(k+1)}{r^2}\right) Y^k(r) &= \frac{2k+1}{r} f(r)
\end{aligned}$$

With boundary condition on the left:

$$\begin{aligned}
Z^k(0) = \frac{k+1}{2k+1} Y^k(0) &= 0 \\
Y^k(0) &= 0
\end{aligned}$$

and on the right:

$$\begin{aligned}
Z^k(r) = -\frac{r}{2k+1} \frac{\mathrm{d}Y^k(r)}{\mathrm{d}r} + \frac{k+1}{2k+1} Y^k(r) &= Y^k(r) \\
-\frac{r}{2k+1} \frac{\mathrm{d}Y^k(r)}{\mathrm{d}r} - \frac{k}{2k+1} Y^k(r) &= 0 \\
\frac{\mathrm{d}Y^k(r)}{\mathrm{d}r} + \frac{k}{r} Y^k(r) &= 0
\end{aligned}$$

which for $r \to \infty$ becomes:

$$\left.\frac{\mathrm{d}Y^k(r)}{\mathrm{d}r}\right|_{r=\infty} = 0$$

but in practise, it's better to use the former Newton (Robin) boundary condition. We have obtained one second order equation for $Y^k(r)$

$$\left(-\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{k(k+1)}{r^2}\right) Y^k(r) = \frac{2k+1}{r} f(r)$$

with boundary conditions:

$$Y^k(0) = 0$$
$$\frac{\mathrm{d}Y^k(r)}{\mathrm{d}r} + \frac{k}{r} Y^k(r) = 0$$

The weak formulation is:

$$\int_0^{r_{max}} Y^{k\prime}(r) v'(r) + \frac{k(k+1)}{r^2} Y^k(r) v(r) \mathrm{d}r - [Y^{k\prime}(r) v(r)]_0^{r_{max}} = \int_0^{r_{max}} \frac{2k+1}{r} f(r) v(r) \mathrm{d}r$$

The boundary term can be simplified using the boundary conditions as:

$$-[Y^{k\prime}(r) v(r)]_0^{r_{max}} = -Y^{k\prime}(r_{max}) v(r_{max}) + Y^{k\prime}(0) v(0) = -Y^{k\prime}(r_{max}) v(r_{max}) = \frac{k}{r_{max}} Y^k(r_{max}) v(r_{max})$$

so we get

$$\int_0^{r_{max}} Y^{k\prime}(r) v'(r) + \frac{k(k+1)}{r^2} Y^k(r) v(r) \mathrm{d}r + \frac{k}{r_{max}} Y^k(r_{max}) v(r_{max}) = \int_0^{r_{max}} \frac{2k+1}{r} f(r) v(r) \mathrm{d}r$$

where the test functions $v(r)$ have the constrain $v(0) = 0$ on the left boundary and no constrain on the right.

8.10.10 Hartree Potential in Spherical Symmetry

For both open and closed shell atoms we get exactly:

$$V_H(\mathbf{x}) = \int \frac{n(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\mathrm{d}^3y = \int \frac{\sum_{j=1}^{Z}|\psi_j(\mathbf{y})|^2}{|\mathbf{x}-\mathbf{y}|}\mathrm{d}^3y =$$
$$= 2\sum_{nlm}\int \frac{|Y_{lm}(\Omega')|^2 P_{nl}^2(r')}{|\mathbf{x}-\mathbf{y}|}\mathrm{d}\Omega'\mathrm{d}r' =$$
$$= 2\sum_{nlm}\sum_{l'm'}\int \frac{r_<^{l'}}{r_>^{l'+1}}\frac{4\pi}{2l'+1}Y_{lm}^*(\Omega')Y_{lm}(\Omega')Y_{l'm'}^*(\Omega)Y_{l'm'}(\Omega')P_{nl}^2(r')\mathrm{d}\Omega'\mathrm{d}r' =$$
$$= 2\sum_{nlm}\sum_{l'}\int \frac{r_<^{l'}}{r_>^{l'+1}}\frac{4\pi}{2l'+1}Y_{l'0}^*(\Omega)\sqrt{\frac{2l'+1}{4\pi}}c^{l'}(l,m,l,m)P_{nl}^2(r')\mathrm{d}r' =$$
$$= 2\sum_{nl}\sum_{l'=0}^{2l}\int \frac{r_<^{l'}}{r_>^{l'+1}}\sqrt{\frac{4\pi}{2l'+1}}Y_{l'0}^*(\Omega)\sum_m c^{l'}(l,m,l,m)P_{nl}^2(r')\mathrm{d}r' =$$
$$= 2\sum_{nl}\int \frac{1}{r_>}\sqrt{4\pi}Y_{00}^*(\Omega)\sum_m c^0(l,m,l,m)P_{nl}^2(r')\mathrm{d}r' +$$
$$+2\sum_{nl}\sum_{l'=1}^{2l}\int \frac{r_<^{l'}}{r_>^{l'+1}}\sqrt{\frac{4\pi}{2l'+1}}Y_{l'0}^*(\Omega)\sum_m c^{l'}(l,m,l,m)P_{nl}^2(r')\mathrm{d}r' =$$
$$= \sum_{nl}2\sum_m c^0(l,m,l,m)\int \frac{1}{r_>}P_{nl}^2(r')\mathrm{d}r' +$$
$$+2\sum_{nl}\sum_{l'=1}^{2l}\int \frac{r_<^{l'}}{r_>^{l'+1}}\sqrt{\frac{4\pi}{2l'+1}}Y_{l'0}^*(\Omega)\sum_m c^{l'}(l,m,l,m)P_{nl}^2(r')\mathrm{d}r' =$$
$$= \sum_{nl}2\sum_m \int \frac{1}{r_>}P_{nl}^2(r')\mathrm{d}r' +$$
$$+2\sum_{nl}\sum_{l'=1}^{2l}\int \frac{r_<^{l'}}{r_>^{l'+1}}\sqrt{\frac{4\pi}{2l'+1}}Y_{l'0}^*(\Omega)\sum_m c^{l'}(l,m,l,m)P_{nl}^2(r')\mathrm{d}r' =$$
$$= \sum_{nl} f_{nl}\int \frac{1}{r_>}P_{nl}^2(r')\mathrm{d}r' +$$
$$+2\sum_{nl}\sum_{l'=1}^{2l}\sqrt{\frac{4\pi}{2l'+1}}Y_{l'0}^*(\Omega)\sum_m c^{l'}(l,m,l,m)\int \frac{r_<^{l'}}{r_>^{l'+1}}P_{nl}^2(r')\mathrm{d}r'$$

For closed shell atoms we use the fact, that

$$\sum_{m=-l}^{l} c^{l'}(l,m,l,m) = (2l+1)\delta_{l'0}$$

and the second term disappears, and for open shell atoms we have to use the central field approximation: we average the integral for V_H over the angles:

$$V_H(\mathbf{x}) \to V_H(r) = \frac{1}{4\pi}\int V_H(\mathbf{x})\,\mathrm{d}\Omega$$

and using the fact, that

$$\int Y_{l'0}^*(\Omega)\,\mathrm{d}\Omega = \sqrt{4\pi}\delta_{l'0}$$

the second term disappears as well. We got the same expression for both open shell (with central field approximation) and closed shell (no approximation) atoms. The radial charge density is:

$$n(r) = \frac{1}{4\pi} \sum_{nl} f_{nl} \left(\frac{P_{nl}(r)}{r} \right)^2$$

So we got:

$$V_H(r) = \sum_{nl} f_{nl} \int \frac{1}{r_>} P_{nl}^2(r') \mathrm{d}r' = \int \frac{4\pi n(r') r'^2}{r_>} \mathrm{d}r' = \frac{Y^0(4\pi n(r) r^2, r)}{r}$$

The Hartree screening function $Y^0(4\pi n(r) r^2, r)$ is given by the equation:

$$-\frac{\mathrm{d}^2}{\mathrm{d}r^2} Y^0(r) = \frac{1}{r} 4\pi n(r) r^2$$

So $V_H(r)$ satisfies the radial Poisson equation:

$$\begin{aligned} (V_H(r) r)'' &= -\frac{1}{r} 4\pi n(r) r^2 \\ V_H''(r) r + 2V_H'(r) &= -4\pi n(r) r \\ V_H''(r) + \frac{2}{r} V_H'(r) &= -4\pi n(r) \end{aligned}$$

8.10.11 Nonlocal Exchange Potential in Spherical Symmetry

Similarly, we calculate:

$$\begin{aligned} \sum_{j=1}^{Z} \int \frac{\psi_i(\mathbf{x}') \psi_j^*(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \mathrm{d}^3 x' \ \psi_j(\mathbf{x}) = \\ = \sum_{n'l'm'} \sum_{k,q} \int \frac{P_{nl}(r')}{r'} Y_{lm}(\Omega') \frac{P_{n'l'}(r')}{r'} Y_{l'm'}^*(\Omega') \frac{P_{n'l'}(r)}{r} Y_{l'm'}(\Omega) \\ \frac{r_<^k}{r_>^{k+1}} \frac{4\pi}{2k+1} Y_{kq}(\Omega) Y_{kq}^*(\Omega') r'^2 \mathrm{d}r' \mathrm{d}\Omega' = \\ = \sum_{n'l'm'} \sum_{k,q} \frac{P_{n'l'}(r)}{r} \frac{4\pi}{2k+1} \int Y_{lm}(\Omega') Y_{l'm'}^*(\Omega') Y_{kq}^*(\Omega') Y_{l'm'}(\Omega) Y_{kq}(\Omega) \mathrm{d}\Omega' \int \frac{r_<^k}{r_>^{k+1}} P_{nl}(r') P_{n'l'}(r') \mathrm{d}r' = \\ = \sum_{n'l'} \sum_{k} \frac{P_{n'l'}(r)}{r} \frac{4\pi}{2k+1} \frac{2k+1}{4\pi} \sqrt{\frac{2l'+1}{2l+1}} c^k(l', 0, l, 0) Y_{lm}(\Omega) \int \frac{r_<^k}{r_>^{k+1}} P_{nl}(r') P_{n'l'}(r') \mathrm{d}r' = \\ = \frac{Y_{lm}(\Omega)}{r} \sum_{n'l'} \sum_{k=|l-l'|}^{k=l+l'} \sqrt{\frac{2l'+1}{2l+1}} c^k(l', 0, l, 0) \int \frac{r_<^k}{r_>^{k+1}} P_{nl}(r') P_{n'l'}(r') \mathrm{d}r' \, P_{n'l'}(r) = \\ = \frac{Y_{lm}(\Omega)}{r} \sum_{n'l'} \sum_{k=|l-l'|}^{k=l+l'} (2l'+1) \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 \int \frac{r_<^k}{r_>^{k+1}} P_{nl}(r') P_{n'l'}(r') \mathrm{d}r' \, P_{n'l'}(r) = \\ = \frac{Y_{lm}(\Omega)}{r} \sum_{n'l'} f_{n'l'} \sum_{k=|l-l'|}^{k=l+l'} \tfrac{1}{2} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 \int \frac{r_<^k}{r_>^{k+1}} P_{nl}(r') P_{n'l'}(r') \mathrm{d}r' \, P_{n'l'}(r) \end{aligned}$$

Functions with different spins don't contribute to the sum, so there is no multiplication by 2. We assumed closed shells atoms (we summed over all m' in the above). We used the result of the integral in *Example VI* and also:

$$\sqrt{\frac{2l'+1}{2l+1}} c^k(l', 0, l, 0) = \sqrt{\frac{2l'+1}{2l+1}} \sqrt{(2l'+1)(2l+1)} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 = (2l'+1) \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 \qquad (8.10.11.1)$$

8.10.12 Radial Hartree-Fock Equations

Using the above integrals, the HF equations become:

$$-\tfrac{1}{2}P''_{nl}(r) + \left(\frac{l(l+1)}{2r^2} - \frac{Z}{r} + V_H(r)\right) P_{nl}(r) +$$
$$-\sum_{n'l'} f_{n'l'} \sum_{k=|l-l'|}^{k=l+l'} \tfrac{1}{2} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 \int \frac{r_<^k}{r_>^{k+1}} P_{nl}(r') P_{n'l'}(r') \mathrm{d}r' \, P_{n'l'}(r) = \epsilon_{nl} P_{nl}(r)$$

with:

$$V_H(r) = \sum_{nl} f_{nl} \int \frac{1}{r_>} P_{nl}^2(r') \mathrm{d}r'$$

Using the Hartree screening functions, the HF equations are:

$$-\tfrac{1}{2}P''_{nl}(r) + \left(\frac{l(l+1)}{2r^2} - \frac{Z}{r} + V_H(r)\right) P_{nl}(r) +$$
$$-\sum_{n'l'} f_{n'l'} \sum_{k=|l-l'|}^{k=l+l'} \tfrac{1}{2} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 \frac{Y^k(P_{nl}(r) P_{n'l'}(r), r)}{r} P_{n'l'}(r) = \epsilon_{nl} P_{nl}(r)$$

with:

$$V_H(r) = \sum_{nl} f_{nl} \frac{Y^0(P_{nl}^2(r), r)}{r} = \frac{Y^0(4\pi n(r) r^2, r)}{r}$$

8.10.13 Total Energy

The total energy is:

$$E = \sum_a 2(2l_a+1) \left(\epsilon_a - \sum_b (2l_b+1) \left(R_0(a,b,a,b) - \sum_l \tfrac{1}{2} \begin{pmatrix} l_a & l & l_b \\ 0 & 0 & 0 \end{pmatrix}^2 R_l(a,b,b,a) \right) \right.$$
$$= \sum_{nl} f_{nl} \left(\epsilon_{nl} - \sum_{n'l'} \tfrac{1}{2} f_{n'l'} \left(R_0(nl,n'l',nl,n'l') - \sum_k \tfrac{1}{2} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 R_l(nl,n'l',n'l',nl) \right) \right.$$
$$= \sum_{nl} f_{nl}\epsilon_{nl} - \sum_{nl}\sum_{n'l'} \tfrac{1}{2} f_{nl} f_{n'l'} \left(\int_0^\infty P_{nl}^2(r) \frac{Y^0(P_{n'l'}^2(r), r)}{r} \mathrm{d}r - \sum_k \tfrac{1}{2} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 \int_0^\infty P_{nl}(r) P_{n'l'}(r) \frac{Y^l(P_{nl}(r)P_{n'l'}(r), r)}{r} \mathrm{d}r \right.$$
$$= \sum_{nl} f_{nl}\epsilon_{nl} - \tfrac{1}{2} \int_0^\infty 4\pi n(r) r^2 \frac{Y^0(4\pi n(r) r^2, r)}{r} \mathrm{d}$$
$$+ \sum_{nl}\sum_{n'l'} \tfrac{1}{2} f_{nl} f_{n'l'} \sum_k \tfrac{1}{2} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 \int_0^\infty P_{nl}(r) P_{n'l'}(r) \frac{Y^l(P_{nl}(r)P_{n'l'}(r), r)}{r} \mathrm{d}$$
$$= \sum_{nl} f_{nl}\epsilon_{nl} - 2\pi \int_0^\infty V_H(r) n(r) r^2 \mathrm{d}r + \sum_{nl}\sum_{n'l'} \tfrac{1}{2} f_{nl} f_{n'l'} \sum_k \tfrac{1}{2} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 \int_0^\infty P_{nl}(r) P_{n'l'}(r) \frac{Y^l(P_{nl}(r)P_{n'l'}(r), r}{r}$$

where:

$$R_l(a,b,c,d) = \int_0^\infty P_a(r) P_c(r) \frac{Y^l(P_b(r)P_d(r), r)}{r} \mathrm{d}r$$

Example: Helium

For Helium atom, the only nonzero occupation numbers are:

$$f_{10} = 2$$

and the sum over $n'l'$ simplifies to:

$$\sum_{n'l'} f_{n'l'} \sum_{k=|l-l'|}^{k=l+l'} \tfrac{1}{2} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 = f_{10} \tfrac{1}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}^2 = f_{10} \tfrac{1}{2} = 1$$

so we only need to solve for the $1s$ state and we get:

$$-\tfrac{1}{2} P_{10}''(r) + \left(-\frac{Z}{r} + V_H(r) \right) P_{10}(r) - \frac{Y^0(P_{10}(r)P_{10}(r), r)}{r} P_{10}(r) = \epsilon_{10} P_{10}(r)$$

with:

$$V_H(r) = 2\frac{Y^0(P_{10}^2(r), r)}{r} = \frac{Y^0(4\pi n(r) r^2, r)}{r}$$

We can combine the equations:

$$-\tfrac{1}{2} P_{10}''(r) + \left(-\frac{Z}{r} + 2\frac{Y^0(P_{10}^2(r), r)}{r} \right) P_{10}(r) - \frac{Y^0(P_{10}^2(r), r)}{r} P_{10}(r) = \epsilon_{10} P_{10}(r)$$

and we obtain:

$$-\tfrac{1}{2} P_{10}''(r) + \left(-\frac{Z}{r} + \frac{Y^0(P_{10}^2(r), r)}{r} \right) P_{10}(r) = \epsilon_{10} P_{10}(r)$$

8.10.14 FEM

The weak formulation is ($u(r) = P_{nl}(r)$):

$$\int_0^\infty \left(\tfrac{1}{2} u'(r) v'(r) + \left(\frac{l(l+1)}{2r^2} - \frac{Z}{r} + V_H(r) \right) u(r) v(r) \right) \mathrm{d}r +$$
$$- \sum_{n'l'} f_{n'l'} \sum_{k=|l-l'|}^{k=l+l'} \tfrac{1}{2} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 \int_0^\infty v(r) P_{n'l'}(r) \frac{Y^k(u(r)P_{n'l'}(r), r)}{r} \mathrm{d}r = \epsilon \int_0^\infty u(r) v(r) \mathrm{d}r$$

for closed shell atoms:

$$\int_0^\infty \left(\tfrac{1}{2} u'(r) v'(r) + \left(\frac{l(l+1)}{2r^2} - \frac{Z}{r} + V_H(r) \right) u(r) v(r) \right) \mathrm{d}r +$$
$$- \sum_{n'l'} 2(2l'+1) \sum_{k=|l-l'|}^{k=l+l'} \tfrac{1}{2} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 \int_0^\infty v(r) P_{n'l'}(r) \frac{Y^k(u(r)P_{n'l'}(r), r)}{r} \mathrm{d}r = \epsilon \int_0^\infty u(r) v(r) \mathrm{d}r$$

or (here we use the il index to label all functions u_i for the given l)

$$\int_0^\infty \left(\tfrac{1}{2} u_{il}'(r) v'(r) + \left(\frac{l(l+1)}{2r^2} - \frac{Z}{r} + V_H(r) \right) u_{il}(r) v(r) \right) \mathrm{d}r +$$
$$- \sum_{n'l'} \sum_{k=|l-l'|}^{k=l+l'} \sqrt{\frac{2l'+1}{2l+1}} c^k(l, 0, l', 0) \int_0^\infty v(r) P_{n'l'}(r) \frac{Y^k(u_{il}(r)P_{n'l'}(r), r)}{r} \mathrm{d}r = \epsilon \int_0^\infty u_{il}(r) v(r) \mathrm{d}r$$

Introducing radial basis $\phi_{\mu l}(r)$ (where μ, ν labels all basis functions for the given l):

$$u_{il}(r) = \sum_\nu C_{\nu il}\phi_{\nu l}(r)$$
$$v(r) = \phi_{\mu l}(r)$$

we get (here i is again restricted for the subset corresponding to the given l):

$$\sum_\nu \int_0^\infty \left(\tfrac{1}{2}\phi'_{\mu l}(r)\phi'_{\nu l}(r) + \left(\frac{l(l+1)}{2r^2} - \frac{Z}{r} + V_H(r)\right)\phi_{\mu l}(r)\phi_{\nu l}(r)\right)\mathrm{d}r\, C_{\nu il} +$$
$$-\sum_\nu \sum_{n'l'} \sum_{k=|l-l'|}^{k=l+l'} \sqrt{\frac{2l'+1}{2l+1}} c^k(l,0,l',0) \int_0^\infty \phi_{\mu l}(r) P_{n'l'}(r) \frac{Y^k(\phi_{\nu l}(r)P_{n'l'}(r), r)}{r}\mathrm{d}r\, C_{\nu il} = \epsilon \sum_\nu \int_0^\infty \phi_{\mu l}(r)\phi_{\nu l}(r)\mathrm{d}r\, C_{\nu il}$$

This can be written as

$$\sum_\nu F^l_{\mu\nu} C_{\nu il} = \epsilon_i \sum_\nu S^l_{\mu\nu} C_{\nu il}$$
$$F^l_{\mu\nu} = H^{l\text{core}}_{\mu\nu} + G^l_{\mu\nu} = T^l_{\mu\nu} + V^l_{\mu\nu} + G^l_{\mu\nu}$$

where

$$T^l_{\mu\nu} = \int_0^\infty \tfrac{1}{2}\phi'_{\mu l}(r)\phi'_{\nu l}(r) + \phi_{\mu l}(r)\frac{l(l+1)}{2r^2}\phi_{\nu l}(r)\mathrm{d}r$$
$$V^l_{\mu\nu} = \int_0^\infty \phi_{\mu l}(r)\left(-\frac{Z}{r}\right)\phi_{\nu l}(r)\mathrm{d}r$$
$$G^l_{\mu\nu} = \int_0^\infty \phi_{\mu l}(r) V_H(r) \phi_{\nu l}(r)\mathrm{d}r +$$
$$-\sum_{n'l'} \sum_{k=|l-l'|}^{k=l+l'} \sqrt{\frac{2l'+1}{2l+1}} c^k(l,0,l',0) \int_0^\infty \phi_{\mu l}(r) P_{n'l'}(r) \frac{Y^k(\phi_{\nu l}(r)P_{n'l'}(r), r)}{r}\mathrm{d}r$$
$$S^l_{\mu\nu} = \int_0^\infty \phi_{\mu l}(r)\phi_{\nu l}(r)\mathrm{d}r$$
$$V_H(r) = \sum_{n'l'} 2(2l'+1)\frac{Y^0(P^2_{n'l'}(r), r)}{r}$$

The indices $n'l'$ go over all occupied orbitals $P_{n'l'}$. Introducing density:

$$n(\mathbf{x}) = 2\sum_{k=1}^{N/2} |\psi_k(\mathbf{x})|^2 = 2\sum_{nlm} |\psi_{nlm}(\mathbf{x})|^2 = 2\sum_{nlm} \frac{P^2_{nl}(r)}{r^2}|Y_{lm}(\Omega)|^2 =$$
$$= 2\sum_{nl} \frac{P^2_{nl}(r)}{r^2}\frac{2l+1}{4\pi} = \frac{1}{4\pi}\sum_{nl} 2(2l+1)\frac{P^2_{nl}(r)}{r^2} = n(r)$$

We introduce the density matrix $P^l_{\alpha\beta}$ (where as before α, β run over basis functions for the given l only):

$$P^l_{\alpha\beta} = 2\sum_{nm} C_{\alpha nlm} C_{\beta nlm} = \sum_n 2(2l+1) C_{\alpha nl} C_{\beta nl}$$

where the $C_{\alpha nl}$ coefficients are the same for all m corresponding to the given l. The index n runs over all occupied states for the given l. We can write $n(r)$ as

$$P_{nl}(r) = \sum_{\alpha} C_{\alpha nl}\phi_{\alpha l}(r)$$

$$n(r) = \frac{1}{4\pi}\sum_{nl} 2(2l+1)\frac{P_{nl}^2(r)}{r^2} = \frac{1}{4\pi}\sum_{nl} 2(2l+1)\sum_{\alpha\beta} C_{\alpha nl}C_{\beta nl}\frac{\phi_{\alpha l}(r)\phi_{\beta l}(r)}{r^2} =$$

$$= \frac{1}{4\pi}\sum_{l}\sum_{\alpha\beta}\frac{\phi_{\alpha l}(r)P^l_{\alpha\beta}\phi_{\beta l}(r)}{r^2}$$

Finally we get:

$$V_H(r) = \sum_{nl} 2(2l+1)\frac{Y^0(P_{nl}^2(r), r)}{r} = \frac{Y^0(4\pi n(r)r^2, r)}{r} = \sum_{l}\sum_{\alpha\beta} P^l_{\alpha\beta}\frac{Y^0(\phi_{\alpha l}(r)\phi_{\beta l}(r), r)}{r}$$

$$\int_0^\infty \phi_{\mu l}(r)V_H(r)\phi_{\nu l}(r)\mathrm{d}r = \sum_{l'}\sum_{\alpha\beta} P^{l'}_{\alpha\beta}\int_0^\infty \phi_{\mu l}(r)\phi_{\nu l}(r)\frac{Y^0(\phi_{\alpha l'}(r)\phi_{\beta l'}(r), r)}{r}\mathrm{d}r = \sum_{l'}\sum_{\alpha\beta} P^{l'}_{\alpha\beta}R^0(\mu l, \beta l', \nu l, \alpha l')$$

and

$$-\sum_{n'l'}\sum_{k=|l-l'|}^{k=l+l'}\sqrt{\frac{2l'+1}{2l+1}}c^k(l,0,l',0)\int_0^\infty \phi_{\mu l}(r)P_{n'l'}(r)\frac{Y^k(\phi_{\nu l}(r)P_{n'l'}(r), r)}{r}\mathrm{d}r =$$

$$= -\sum_{n'l'}\sum_{k=|l-l'|}^{k=l+l'} 2(2l'+1)^{\frac{1}{2}}\begin{pmatrix} l & k & l' \\ 0 & 0 & 0\end{pmatrix}^2\sum_{\alpha\beta} C_{\alpha n'l'}C_{\beta n'l'}\int_0^\infty \phi_{\mu l}(r)\phi_{\alpha l'}(r)\frac{Y^k(\phi_{\nu l}(r)\phi_{\beta l'}(r), r)}{r}\mathrm{d}r =$$

$$= -\tfrac{1}{2}\sum_{l'}\sum_{\alpha\beta} P^{l'}_{\alpha\beta}\sum_{k=|l-l'|}^{k=l+l'}\begin{pmatrix} l & k & l' \\ 0 & 0 & 0\end{pmatrix}^2 R^k(\mu l, \beta l', \alpha l', \nu l)$$

So we get:

$$G^l_{\mu\nu} = \sum_{l'}\sum_{\alpha\beta} P^{l'}_{\alpha\beta}R^0(\mu l, \beta l', \nu l, \alpha l') - \tfrac{1}{2}\sum_{l'}\sum_{\alpha\beta} P^{l'}_{\alpha\beta}\sum_{k=|l-l'|}^{k=l+l'}\begin{pmatrix} l & k & l' \\ 0 & 0 & 0\end{pmatrix}^2 R^k(\mu l, \beta l', \alpha l', \nu l) =$$

$$= \sum_{l'=0}^{\infty}\sum_{\alpha\beta} P^{l'}_{\alpha\beta}\left(R^0(\mu l, \beta l', \nu l, \alpha l') - \tfrac{1}{2}\sum_{k=|l-l'|}^{k=l+l'}\begin{pmatrix} l & k & l' \\ 0 & 0 & 0\end{pmatrix}^2 R^k(\mu l, \beta l', \alpha l', \nu l)\right)$$

The density matrix is zero if there are no occupied orbitals for the given l'.

The total energy is:

$$E = \tfrac{1}{2}\sum_{\mu\nu} P_{\mu\nu}(H^{\text{core}}_{\mu\nu} + F_{\mu\nu}) =$$

$$E = \tfrac{1}{2}\sum_{l}\sum_{\mu\nu} P^l_{\mu\nu}(H^{l\text{core}}_{\mu\nu} + F^l_{\mu\nu}) =$$

$$= \sum_{l}\sum_{\mu\nu} P^l_{\mu\nu}(F^l_{\mu\nu} - \tfrac{1}{2}G^l_{\mu\nu}) =$$

$$= \sum_{l}\sum_{n} 2(2l+1)\left(\epsilon_{nl} - \sum_{l'}\sum_{n'}(2l'+1)\left(R^0(nl, n'l', nl, n'l') - \tfrac{1}{2}\sum_{k=|l-l'|}^{k=l+l'}\begin{pmatrix} l & k & l' \\ 0 & 0 & 0\end{pmatrix}^2 R^k(nl, n'l', n'l', nl)\right)\right)$$

where we used:

$$\sum_l \sum_{\mu\nu} P^l_{\mu\nu} F^l_{\mu\nu} =$$
$$= \sum_l \sum_{\mu\nu} \sum_n 2(2l+1) C_{\mu nl} C_{\nu nl} F^l_{\mu\nu} =$$
$$= \sum_l \sum_{\mu\nu} \sum_n 2(2l+1) C_{\mu nl} \epsilon_{nl} S^l_{\mu\nu} C_{\nu nl} =$$
$$= \sum_l \sum_n 2(2l+1) \epsilon_{nl}$$

and

$$\tfrac{1}{2} \sum_l \sum_{\mu\nu} P^l_{\mu\nu} G^l_{\mu\nu} =$$
$$= \tfrac{1}{2} \sum_l \sum_{\mu\nu} P^l_{\mu\nu} \sum_{l'} \sum_{\alpha\beta} P^{l'}_{\alpha\beta} \left(R^0(\mu l, \beta l', \nu l, \alpha l') - \tfrac{1}{2} \sum_{k=|l-l'|}^{k=l+l'} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 R^k(\mu l, \beta l', \alpha l', \nu l) \right) =$$
$$= \tfrac{1}{2} \sum_l \sum_{\mu\nu} \sum_n 2(2l+1) C_{\mu nl} C_{\nu nl} \sum_{l'} \sum_{\alpha\beta} \sum_{n'} 2(2l'+1) C_{\alpha n' l'} C_{\beta n' l'}$$
$$\left(R^0(\mu l, \beta l', \nu l, \alpha l') - \tfrac{1}{2} \sum_{k=|l-l'|}^{k=l+l'} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 R^k(\mu l, \beta l', \alpha l', \nu l) \right) =$$
$$= \sum_l \sum_n 2(2l+1) \sum_{l'} \sum_{n'} (2l'+1)$$
$$\left(R^0(nl, n'l', nl, n'l') - \tfrac{1}{2} \sum_{k=|l-l'|}^{k=l+l'} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 R^k(nl, n'l', n'l', nl) \right)$$

8.10.15 4-Index Transformation

The 4-index transformation is a way to convert the two particle integrals over basis functions $(\alpha\beta|\gamma\delta)$ into two particle integrals over atomic (or molecular) orbitals $(ij|kl)$:

$$(ij|kl) = \sum_{\alpha\beta\gamma\delta} C_{\alpha i} C_{\beta j} C_{\gamma k} C_{\delta l} (\alpha\beta|\gamma\delta)$$

8.10.16 Green's Functions

The self energy up to a second order is given by:

$$\Sigma_{ij}(E) = \sum_{ars} \frac{(ir|as)(2(jr|as) - (js|ar))}{E + \epsilon_a - \epsilon_r - \epsilon_s} + \sum_{abr} \frac{(ia|br)(2(ja|br) - (jb|ar))}{E + \epsilon_r - \epsilon_a - \epsilon_b}$$

The a, b are occupied orbitals, r, s are virtual orbitals. The Dyson equation is:

$$\mathbf{G}(E) = \mathbf{G}_0(E) + \mathbf{G}_0(E) \mathbf{\Sigma}(E) \mathbf{G}(E) \tag{8.10.16.1}$$

where

$$\mathbf{G}_0(E) = \frac{1}{E - \boldsymbol{\epsilon}} = \sum_k \frac{|k\rangle \langle k|}{E - \epsilon_k}$$
$$\mathbf{G}(E) = \frac{1}{E - \mathbf{E}} = \sum_k \frac{|k_{int}\rangle \langle k_{int}|}{E - E_k}$$

The ϵ_k and $|k\rangle$ are HF energies and eigenvectors, E_k and $|k_{int}\rangle$ are interacting energies and eigenvectors. The matrix $\boldsymbol{\epsilon}$ is a diagonal matrix of the ϵ_k eigenvalues, $\mathbf{E}$ is a diagonal matrix of the E_k eigenvalues. Any Green's function $\mathbf{G}(E)$ (interacting or not) can be written using the spectral density function $\mathbf{A}(z)$ as follows:

$$\begin{aligned}\mathbf{G}(E) = \frac{1}{E-\boldsymbol{\epsilon}} = \sum_k \frac{|k\rangle\langle k|}{E-\epsilon_k} = \\ = \sum_k \int_{-\infty}^{\infty} \frac{|k\rangle\langle k|}{E-z}\delta(z-\epsilon_k)\mathrm{d}z = \\ = \int_{-\infty}^{\infty} \frac{\mathbf{A}(z)}{E-z}\mathrm{d}z\end{aligned}$$

where

$$\begin{aligned}\mathbf{A}(z) = \sum_k |k\rangle\langle k|\,\delta(z-\epsilon_k) = \\ = \sum_k |k\rangle\langle k| \lim_{\eta\to 0}\frac{\eta}{\pi}\frac{1}{(z-\epsilon_k)^2+\eta^2} = \\ = \sum_k |k\rangle\langle k| \lim_{\eta\to 0}\frac{i}{2\pi}\left(\frac{1}{z-\epsilon_k+i\eta} - \frac{1}{z-\epsilon_k-i\eta}\right) = \\ = \lim_{\eta\to 0}\frac{i}{2\pi}\left(\mathbf{G}(z+i\eta) - \mathbf{G}(z-i\eta)\right)\end{aligned}$$

From (8.10.16.1) we get:

$$\begin{aligned}\frac{1}{\mathbf{G}_0(E)}\mathbf{G}(E) = 1 + \boldsymbol{\Sigma}(E)\mathbf{G}(E) \\ \frac{1}{\mathbf{G}_0(E)} = \frac{1}{\mathbf{G}(E)} + \boldsymbol{\Sigma}(E) \\ \mathbf{G}(E) = \frac{1}{\frac{1}{\mathbf{G}_0(E)} - \boldsymbol{\Sigma}(E)} = \frac{1}{E-\boldsymbol{\epsilon}-\boldsymbol{\Sigma}(E)}\end{aligned}$$

The poles E_k of the Green's function $\mathbf{G}(E)$ are then given by:

$$D(E) = \det(E - \boldsymbol{\epsilon} - \boldsymbol{\Sigma}(E)) = 0$$

or equivalently:

$$(\boldsymbol{\Sigma}(E_k) + \boldsymbol{\epsilon})\,|v\rangle = E_k\,|v\rangle$$

and from the theory of matrices:

$$\begin{aligned}\operatorname{Tr}\frac{1}{E-\mathbf{h}} = \operatorname{Tr}\frac{\frac{\partial}{\partial E}(E-\mathbf{h})}{E-\mathbf{h}} = \operatorname{Tr}\frac{\partial}{\partial E}\log(E-\mathbf{h}) = \frac{\partial}{\partial E}\operatorname{Tr}\log(E-\mathbf{h}) = \\ = \frac{\partial}{\partial E}\log|\det(E-\mathbf{h})| = \frac{1}{\det(E-\mathbf{h})}\frac{\partial\det(E-\mathbf{h})}{\partial E}\end{aligned}$$

one obtains that

$$G_{ij}(E) = (-1)^{i+j+1}\frac{\partial\log D(E)}{\partial(\boldsymbol{\Sigma}(E_k)+\boldsymbol{\epsilon})_{ji}}$$

and

$$\operatorname{Tr}\mathbf{G}(E) = \sum_k G_{kk}(E) = \frac{\partial\log|D(E)|}{\partial E}$$

The number of particles N can be calculated as follows (a are occupied orbitals, k are all orbitals):

$$\begin{aligned} N &= \sum_a |\langle \mathbf{r}|a\rangle|^2 = \\ &= \sum_a \langle \mathbf{r}|a\rangle \langle a|\mathbf{r}\rangle = \\ &= \operatorname{Tr} \sum_a |a\rangle \langle a| = \\ &= \frac{1}{2\pi i} \operatorname{Tr} \int_C \sum_k \frac{|k\rangle \langle k|}{E - E_k} \mathrm{d}E = \\ &= \frac{1}{2\pi i} \int_C \operatorname{Tr} \mathbf{G}(E)\, \mathrm{d}E = \\ &= \frac{1}{2\pi i} \int_C \frac{\partial \log |D(E)|}{\partial E}\, \mathrm{d}E \end{aligned}$$

The countour C only encloses poles E_a corresponding to occupied orbitals a. Similarly for the total energy E_{tot}:

$$\begin{aligned} E_{tot} &= \sum_a E_a |\langle \mathbf{r}|a\rangle|^2 = \\ &= \sum_a \langle \mathbf{r}|a\rangle E_a \langle a|\mathbf{r}\rangle = \\ &= \operatorname{Tr} \sum_a |a\rangle E_a \langle a| = \\ &= \frac{1}{2\pi i} \operatorname{Tr} \int_C \sum_k \frac{|k\rangle \langle k|}{E - E_k} E\, \mathrm{d}E = \\ &= \frac{1}{2\pi i} \int_C \operatorname{Tr} \mathbf{G}(E) E\, \mathrm{d}E = \\ &= \frac{1}{2\pi i} \int_C \frac{\partial \log |D(E)|}{\partial E} E\, \mathrm{d}E \end{aligned}$$

For doubly filled orbitals we multiply the expressions by 2.

8.11 Projector Augmented-Wave Method (PAW)

We can use the Density Functional Theory (DFT) to reduce the many body problem to solve a single particle Schrödinger equation:

$$H|\psi_n\rangle = \epsilon_n |\psi_n\rangle$$

The wavefunctions contain cusps (and are oscillatory around each nucleus), also one needs to solve this for all core states.

Next step is to use the known behavior around each atom and take advantage of the known physics somehow. There are two general approaches, either one can incorporate the known physic in the basis (for example the partition of unity in the finite element method), or into the equations. PAW method uses the latter approach.

8.11.1 Projectors, Augmentation Spheres and Smooth Wavefunctions

We introduce *smooth wavefunctions* (we'll use "~" for smooth functions) by a linear transformation operator T:

$$|\psi_n\rangle = T|\tilde{\psi}_n\rangle$$

We construct *augmentation spheres* $|\mathbf{r} - \mathbf{R}^a| < r_c^a$ around each atom a (one can imagine a muffin-tin), where r_c^a is a cut-off radius, a is the atom index, $\mathbf{R}^a$ is the atom position.

We write T as:

$$T = \mathbb{1} + \sum_a T^a$$

where T^a only acts in the augmentation sphere. We choose a complete basis $|\phi_i^a\rangle$ (also called *partial waves*) inside the sphere. The smooth partial waves can be obtained using the T operator:

$$\begin{aligned} |\phi_i^a\rangle = T\,|\tilde{\phi}_i^a\rangle = \left(\mathbb{1} + \sum_{a'} T^{a'}\right)|\tilde{\phi}_i^a\rangle \\ = (\mathbb{1} + T^a)\,|\tilde{\phi}_i^a\rangle \end{aligned}$$

Because T^a only acts in the sphere, it follows that

$$|\phi_i^a\rangle = |\tilde{\phi}_i^a\rangle \qquad \text{for } r > r_c^a$$

outside the sphere (i.e. $\langle\mathbf{r}|\phi_i^a\rangle = \langle\mathbf{r}|\tilde{\phi}_i^a\rangle$ for $r > r_c^a$). We can now expand the smooth wavefunctions using the partial waves basis:

$$|\tilde{\psi}_n\rangle = \sum_i P_{ni}^a\,|\tilde{\phi}_i^a\rangle \tag{8.11.1.1}$$

inside the augmentation sphere. Multiplying both sides by T:

$$\begin{aligned} T\,|\tilde{\psi}_n\rangle &= T\sum_i P_{ni}^a\,|\tilde{\phi}_i^a\rangle \\ T\,|\tilde{\psi}_n\rangle &= \sum_i P_{ni}^a\,T\,|\tilde{\phi}_i^a\rangle \\ |\psi_n\rangle &= \sum_i P_{ni}^a\,|\phi_i^a\rangle \end{aligned} \tag{8.11.1.2}$$

So both smooth and non-smooth wavefunctions have the same expansion coefficients P_{ni}^a. We choose smooth *projector functions* $|\tilde{p}_i^a\rangle$ satisfying the following orthogonality and completeness relations inside the augmentation spheres (no restrictions are imposed outside the spheres, so we just define $\langle\mathbf{r}|\tilde{p}_i^a\rangle = 0$):

$$\begin{aligned} \langle\tilde{p}_i^a|\tilde{\phi}_j^a\rangle &= \delta_{ij} \\ \sum_i |\tilde{\phi}_i^a\rangle\,\langle\tilde{p}_i^a| &= \mathbb{1} \end{aligned} \tag{8.11.1.3}$$

then multiplying (8.11.1.1) by $\langle\tilde{p}_i^a|$ and using (8.11.1.3):

$$\langle\tilde{p}_i^a|\tilde{\psi}_n\rangle = \sum_j P_{nj}^a\,\langle\tilde{p}_i^a|\tilde{\phi}_j^a\rangle = \sum_j P_{nj}^a\delta_{ij} = P_{ni}^a$$

we can rewrite (8.11.1.1) and (8.11.1.2):

$$\begin{aligned} |\tilde{\psi}_n\rangle &= \sum_i \langle\tilde{p}_i^a|\tilde{\psi}_n\rangle\,|\tilde{\phi}_i^a\rangle \\ |\psi_n\rangle &= \sum_i \langle\tilde{p}_i^a|\tilde{\psi}_n\rangle\,|\phi_i^a\rangle \end{aligned} \tag{8.11.1.4}$$

Let's write T^a using the projectors:

$$T^a = T^a\mathbb{1} = T^a\sum_i |\tilde{\phi}_i^a\rangle\,\langle\tilde{p}_i^a| = \sum_i (T^a\,|\tilde{\phi}_i^a\rangle)\,\langle\tilde{p}_i^a| = \sum_i (|\phi_i^a\rangle - |\tilde{\phi}_i^a\rangle)\,\langle\tilde{p}_i^a|$$

Note that the right hand side is zero outside the augmentation sphere. Thus

$$T = \mathbb{1} + \sum_a T^a = \mathbb{1} + \sum_a \sum_i (|\phi_i^a\rangle - |\tilde{\phi}_i^a\rangle) \langle \tilde{p}_i^a|$$

In other words, the transformation operator T is completely defined using the smooth and non-smooth partial waves and the projector functions. In terms of the wavefunction:

$$|\psi_n\rangle = T|\tilde{\psi}_n\rangle = |\tilde{\psi}_n\rangle + \sum_a \sum_i (|\phi_i^a\rangle - |\tilde{\phi}_i^a\rangle) \langle \tilde{p}_i^a|\tilde{\psi}_n\rangle =$$

$$= |\tilde{\psi}_n\rangle + \sum_a \left(\sum_i |\phi_i^a\rangle \langle \tilde{p}_i^a|\tilde{\psi}_n\rangle - \sum_i |\tilde{\phi}_i^a\rangle \langle \tilde{p}_i^a|\tilde{\psi}_n\rangle \right)$$

In words, the wavefunction can be decomposed as the sum of the smooth wavefunction and sum over atoms (centers), at each atom we have "1-center all electron" minus "1-center pseudo".

The projection functions can always be written as

$$\langle \tilde{p}_i^a| = \sum_j \left\{ \langle f_k^a|\tilde{\phi}_l^a\rangle \right\}_{ij}^{-1} \langle f_j^a|$$

where $|f_k^a\rangle$ is any set of linearly independent functions.

Note: the n above means all states of interest — either all states, or only the valence states.

8.11.2 Frozen Core Approximation

One can either calculate all electrons in the eigenproblem, or only calculate the valence electrons and treat the core states separately. The simplest option is to introduce a *frozen core approximation*, where

$$|\psi_n\rangle = |\phi_\alpha^{a,\text{core}}\rangle$$

for all core states n, here n runs over (a, α), where a is the atom index and α are the core states of an atom. This approximation can also be relaxed in various ways.

8.11.3 Expectation Values of Local Operators

In the frozen core approximation:

$$\langle O\rangle = \sum_n^{\text{val}} f_n \langle \psi_n|O|\psi_n\rangle + \sum_a \sum_\alpha^{\text{core}} \langle \phi_\alpha^{a,\text{core}}|O|\phi_\alpha^{a,\text{core}}\rangle = \cdots =$$

$$= \sum_n^{\text{val}} f_n \langle \tilde{\psi}_n|O|\tilde{\psi}_n\rangle + \sum_a \sum_{i,j} \left(\langle \phi_i^a|O|\phi_j^a\rangle - \langle \tilde{\phi}_i^a|O|\tilde{\phi}_j^a\rangle \right) D_{ij}^a + \sum_a \sum_\alpha^{\text{core}} \langle \phi_\alpha^{a,\text{core}}|O|\phi_\alpha^{a,\text{core}}\rangle$$

where the tensor D_{ij}^a is:

$$D_{ij}^a = \sum_n f_n \langle \tilde{\psi}_n|\tilde{p}_i^a\rangle \langle \tilde{p}_j^a|\tilde{\psi}_n\rangle$$

Density

$$n(\mathbf{r}) = \sum_n f_n |\psi_n(\mathbf{r})|^2 = \sum_n f_n \langle\psi_n|\mathbf{r}\rangle \langle\mathbf{r}|\psi_n\rangle = \left\langle |\mathbf{r}\rangle \langle\mathbf{r}| \right\rangle =$$
$$= \sum_n^{\text{val}} f_n |\tilde{\psi}_n(\mathbf{r})|^2 + \sum_a \sum_{i,j} \left(\phi_i^a \phi_j^a - \tilde{\phi}_i^a \tilde{\phi}_j^a\right) D_{ij}^a + \sum_a \sum_\alpha^{\text{core}} |\phi_\alpha^{a,\text{core}}|^2$$

The functions $\phi_\alpha^{a,\text{core}}$ are not strictly localized withing the augmentation sphere.

8.11.4 Kohn Sham Equations

We multiply the original equations by $T^\dagger$ from the left and introduce the smooth wavefunctions:

$$H|\psi_n\rangle = \epsilon_n |\psi_n\rangle$$
$$T^\dagger H |\psi_n\rangle = \epsilon_n T^\dagger |\psi_n\rangle$$
$$T^\dagger H T |\tilde{\psi}_n\rangle = \epsilon_n T^\dagger T |\tilde{\psi}_n\rangle$$

The orthogonality of wavefunctions is:

$$\langle\psi_n|\psi_m\rangle = \delta_{nm}$$
$$\langle\tilde{\psi}_n|T^\dagger T|\tilde{\psi}_m\rangle = \delta_{nm}$$

The overlap operator $T^\dagger T$ can be written as:

$$T^\dagger T = \cdots = \mathbb{1} + \sum_a \sum_{i,j} |\tilde{p}_i^a\rangle Q_{ij} \langle\tilde{p}_j^a|$$

where

$$Q_{ij} = \langle\phi_i^a|\phi_j^a\rangle - \langle\tilde{\phi}_i^a|\tilde{\phi}_j^a\rangle$$

The transformed Hamiltonian is

$$H = -\tfrac{1}{2}\nabla^2 + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) + v(\mathbf{r})$$
$$T^\dagger H T = \cdots = -\tfrac{1}{2}\nabla^2 + V_H(\tilde{n}) + V_{xc}(\tilde{n}) + \sum_a \sum_{ij} |\tilde{p}_i^a\rangle H_{ij}^a \langle\tilde{p}_j^a|$$

where

$$H_{ij}^a = \langle\phi_i^a| - \tfrac{1}{2}\nabla^2 + v_{\text{eff}}|\phi_j^a\rangle - \langle\tilde{\phi}_i^a| - \tfrac{1}{2}\nabla^2 + \tilde{v}_{\text{eff}}|\tilde{\phi}_j^a\rangle$$

BIBLIOGRAPHY

[Adams] Adams, J. C. (1878). On the Expression of the Product of Any Two Legendre's Coefficients by Means of a Series of Legendre's Coefficients. Proceedings of the Royal Society of London, 27, 63-71.

[Gaunt] Gaunt, J. A. (1929). The Triplets of Helium. Philosophical Transactions of the Royal Society of London, 228, 151-196.

[Birss] Birss, R.R. (1961). Symmetry and Magnetism. Volume III. North-Holland Publishing Company.

[Bishop] Bishop D.M. (1993). Group theory and chemistry. Courier Dover Publications.

[Elliott] Elliott, J.P. and Dawber, P.G. (1979). Symmetry in Physics. The Macmillan Press Ltd.

[Georgi] Georgi, H. (1990). Lie Algebras in Particle Physics. Addison-Wesley.

[Pilar] Pilar, F.L. (1990). Elementary Quantum Chemistry, McGraw–Hill.

[Sternberg] Sternberg, S. (1994). Group theory and physics. Cambridge University Press

[Blokker72] Blokker, E. International journal of quantum chemistry VI, 925, (1972)

[Cannon69] Cannon, J.J. Communications of the association for computing machinery, **12**, 3 (1969)

[Chillag86] Chillag, D. (1986). Character Values of Finite Groups as Eigenvalues Of Nonnegative Integer Matrices. Proceedings of the American Mathematical Society, 97(3), 565-567.

[Dixon67] Dixon, J.D. Numerische Mathematik **10**, 446 (1967)

[Pask2012] Pask, J. E., Sukumar, N., Mousavi, S. E. (2012). Linear scaling solution of the all-electron Coulomb problem in solids. International Journal for Multiscale Computational Engineering, 10(1), 83–99. doi:10.1615/IntJMultCompEng.2011002201

[DFT] 18. (m) Dreizler, E. K. U. Gross: Density functional theory: an approach to the quantum many-body problem

[pickett] 23. (e) Pickett, Pseudopotential methods in condensed matter applications, Computer Physics reports, Volume 9, Issue 3, April 1989, Pages 115-197, ISSN 0167-7977, DOI: 10.1016/0167-7977(89)90002-6. (http://www.sciencedirect.com/science/article/B6X3V-46R02CR-1J/2/804d9ecaa49469aa5e1050dc007f4a61)

INDEX