Theoretical Physics Reference

Release 0.1

Ondřej Čertík

November 09, 2009

CONTENTS

CHAPTER

ONE

MATHEMATICS

1.1 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.

1.1.1 General Case

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

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

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

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

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

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

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 $n = 3$):

$$\varphi(\theta,\phi) = \begin{pmatrix} r\sin\theta\cos\phi \\ r\sin\theta\sin\phi \\ r\cos\theta \end{pmatrix}$$

$$\mathbf{J} = \begin{pmatrix} -r\sin\theta\sin\phi & r\cos\theta\cos\phi \\ r\sin\theta\cos\phi & r\cos\theta\sin\phi \\ 0 & -r\sin\theta \end{pmatrix}$$

$$\mathbf{G} = \mathbf{J}^T\mathbf{J} = \begin{pmatrix} -r\sin\theta\sin\phi & r\sin\theta\cos\phi & 0 \\ r\cos\theta\cos\phi & r\cos\theta\sin\phi & -r\sin\theta \end{pmatrix} \begin{pmatrix} -r\sin\theta\sin\phi & r\cos\theta\cos\phi \\ r\sin\theta\cos\phi & r\cos\theta\sin\phi \\ 0 & -r\sin\theta \end{pmatrix} = \begin{pmatrix} r^2\sin^2\theta & 0 \\ 0 & r^2 \end{pmatrix}$$

$$\det\mathbf{G} = r^4\sin^2\theta$$

$$\sqrt{\det\mathbf{G}} = r^2\sin\theta$$

$$\int_M f(x,y,z)\mathrm{d}S = \int_{\mathbf{R}^n} f(r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta)\, r^2\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi =$$
$$= \int_0^\pi \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi\, f(r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta)\, r^2\sin\theta$$

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

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

1.1.2 Special Cases

k = n

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

k = 1

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

k = n - 1

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

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

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

y = f(x, z)

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

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

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

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

Implicit Surface

For a surface given explicitly 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:

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

1.1.3 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.

1.2 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:

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

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

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

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:

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

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

$$\begin{aligned}\frac{1}{2\pi i}\int\limits_{|z-z_0|=\epsilon}(z-z_0)^n\mathrm{d}z &= \frac{1}{2\pi i}\int\limits_0^{2\pi}(z_0+\epsilon e^{i\varphi}-z_0)^n i\epsilon e^{i\varphi}\mathrm{d}\varphi = \frac{\epsilon^{n+1}}{2\pi}\int\limits_0^{2\pi} e^{i\varphi(n+1)}\mathrm{d}\varphi = \\ &= \begin{cases}\frac{\epsilon^{n+1}}{2\pi}\left[\frac{e^{i\varphi(n+1)}}{i(n+1)}\right]_0^{2\pi} = 0 & \text{for } n\neq -1 \\ \frac{1}{2\pi}\int\limits_0^{2\pi}\mathrm{d}\varphi = 1 & \text{for } n = -1\end{cases} = \delta_{n,-1}\end{aligned}$$

1.2.1 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:

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

in particular for $m = 1$:

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

for $m = 2$:

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

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

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

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

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

1.2.2 Useful Formulas

Jordan's Lemma

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

$$\left|\int_\Omega g(z)\mathrm{d}z\right| \le \pi R \max_\Omega |g(z)|$$

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

(In the first case the integration path can be extended to the full circle if needed ($0 \le \varphi \le 2\pi$), 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:

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

and

$$\left|\int_\Omega e^{i\alpha z} g(z)\mathrm{d}z\right| = \left|\int_0^\pi e^{i\alpha Re^{i\varphi}} g(Re^{i\varphi}) iRe^{i\varphi}\right| \le$$

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

$$< R \max_\Omega |g(z)| 2 \int_0^{\frac{\pi}{2}} e^{-\alpha R\frac{2}{\pi}\varphi}\mathrm{d}\varphi = \frac{\pi}{\alpha} \max_\Omega |g(z)| (1 - e^{-\alpha R}) =$$

$$< \frac{\pi}{\alpha} \max_\Omega |g(z)|$$

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:

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

1.2.3 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.

1.3 Fourier Transform

The 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:

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

1.4 Laplace Transform

Laplace transform of $f(x)$ 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 $s = \sigma + i\omega$:

$$L[f(x)] \equiv \bar{f}(s) = F[U(x)] \equiv \tilde{U}(\omega) = \int_{-\infty}^{\infty} U(x)e^{-i\omega x}\,\mathrm{d}x = \int_0^{\infty} f(x)e^{-\sigma x}e^{-i\omega x}\,\mathrm{d}x = \int_0^{\infty} f(x)e^{-sx}\,\mathrm{d}x$$

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

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

Where the bar ($\bar{f}$) means the Laplace transform and tilde ($\tilde{U}$) means the Fourier transform. 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 (otherwise the integral in the Laplace transform may not converge). It can be shown that the integral over the left semicircle goes to zero, so the complex integral is equal to the sum of all residues of $\bar{f}(s)e^{sx}$ in the complex plane.

1.5 Polar and Spherical Coordinates

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

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

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

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

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.

1.6 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 $(-\pi, \pi]$, we get the $\mathrm{Arg}(z)$ function:

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

then $\arg z = \mathrm{Arg} z + 2\pi n$, where $n = 0, \pm 1, \pm 2, \dots$. We can then use the following formula to easily calculate $\mathrm{Arg} z$ for any $z = x + iy$:

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

Finally we define $\mathrm{atan2}(y, x)$ as:

$$\mathrm{atan2}(y,x) = \mathrm{Arg}(x+iy)$$

Some properties:

$$\begin{aligned} \tan \mathrm{atan2}(y,x) &= \frac{y}{x} \\ \sin \mathrm{atan2}(y,x) &= \frac{x}{\sqrt{x^2+y^2}} \\ \cos \mathrm{atan2}(y,x) &= \frac{y}{\sqrt{x^2+y^2}} \\ \mathrm{atan2}(ky,kx) = \mathrm{Arg}(kx+iky) &= \mathrm{Arg}(x+iy) = \mathrm{atan2}(y,x) \end{aligned}$$

An example of an application:

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

where

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

1.7 Delta Function

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

$$\int f(x)\delta(x-t)\mathrm{d}x = f(t) \tag{1.1}$$

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

$$\lim_{\alpha\to\infty}\delta_\alpha(x) = \delta(x) \tag{1.2}$$

more precisely:

$$\lim_{\alpha\to\infty}\int f(x)\delta_\alpha(x)\mathrm{d}x = \int f(x)\lim_{\alpha\to\infty}\delta_\alpha(x)\mathrm{d}x = f(t) \tag{1.3}$$

one example of such a sequence is:

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

It's clear that (1.3) 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 (1.1) doesn't "exist", but that's a wrong approach to things, because it is not important if something "exist" or not, but rather if it is clear what we mean by our notation: (1.1) is a shorthand for (1.3) and (1.2) gets a mathematically rigorous meaning when you integrate both sides and use (1.1) to arrive at (1.3). Thus one uses the relations (1.1), (1.2), (1.3) 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 $\hat{\mathbf{r}}(\theta, \phi) = (\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta)$.

$$f(\hat{\mathbf{r}}') = \int \delta(\hat{\mathbf{r}} - \hat{\mathbf{r}}') f(\hat{\mathbf{r}})\, \mathrm{d}\hat{\mathbf{r}} \tag{1.4}$$

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

$$f(\theta', \phi') = \int \delta(\theta - \theta')\delta(\phi - \phi') f(\theta, \phi)\, \mathrm{d}\theta\mathrm{d}\phi \tag{1.5}$$

Expressing (1.4) 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 (1.5) 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 (1.6) for an example of how to deal with more complex expressions involving the delta function like $\delta^2(x)$.

1.8 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):

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

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

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

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:

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

so:

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

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

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

1.9 Variations and Functional Derivatives

Functional derivatives are a common source of confusion and especially the notation. The reason is similar to the delta function — the definition is operational, i.e. it tells you what operations you need to do to get a mathematically precise formula. The notation below is commonly used in physics and in our opinion it is perfectly precise and exact, but some mathematicians may not like it.

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

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

The last equality follows from the fact, that $\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon} f(\mathbf{x}+\varepsilon\mathrm{h})\right|_{\varepsilon=0}$ is a linear function of h. We define $\frac{\partial f}{\partial x_i}$ as

$$\mathrm{a} \equiv \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_N}\right)$$

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

$$\frac{\partial f}{\partial x_i} = a_i = \mathrm{a}\cdot\mathrm{h} = \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon} f(\mathbf{x} + \varepsilon(0, 0, \dots, 1, \dots, 0))\right|_{\varepsilon=0} =$$

$$= \lim_{\varepsilon\to 0} \frac{f(x_1, x_2, \dots, x_i + \varepsilon, \dots, x_N) - f(x_1, x_2, \dots, x_i, \dots, x_N)}{\varepsilon}$$

But this is just the way the partial derivative is usually defined. Every variable can be treated as a function (very simple one):

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

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 $\mathrm{h} = \mathrm{d}\mathbf{x}$ and

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

So $\mathrm{d}\mathbf{x}$ has two meanings — it's either $\mathrm{h} = \mathbf{x} - \mathbf{x}_0$ (a finite change in the independent variable $\mathbf{x}$) or a differential, depending on the context. Even mathematicians use this notation.

Functional $F[f]$ assigns a number to each function $f(x)$. The variation is defined as

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

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

$$\frac{\delta F}{\delta f(x)} = a(x) = \int a(y)\delta(x-y)\mathrm{d}y = \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon} F[f(y) + \varepsilon\delta(x-y)]\right|_{\varepsilon=0} =$$

$$= \lim_{\varepsilon\to 0} \frac{F[f(y)+\varepsilon\delta(x-y)] - F[f(y)]}{\varepsilon}$$

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

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

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 $h(x) = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(f(x)+\varepsilon h(x))\big|_{\varepsilon=0}$ (a finite change in the function f) or a variation of a functional, depending on the context. Mathematicians never write δf in the meaning of $h(x)$, they always write the latter, but it's ridiculous, because it is completely analogous to $\mathrm{d}\mathbf{x}$.

The correspondence between the finite and infinite dimensional case can be summarized as:

$$\begin{aligned}
f(x_i) &\quad\Longleftrightarrow\quad F[f] \\
\mathrm{d}f = 0 &\quad\Longleftrightarrow\quad \delta F = 0 \\
\frac{\partial f}{\partial x_i} = 0 &\quad\Longleftrightarrow\quad \frac{\delta F}{\delta f(x)} = 0 \\
f &\quad\Longleftrightarrow\quad F \\
x_i &\quad\Longleftrightarrow\quad f(x) \\
x &\quad\Longleftrightarrow\quad f \\
i &\quad\Longleftrightarrow\quad x
\end{aligned}$$

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

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

This notation allows us a very convinient computation, as shown in the following examples. First, when computing a variation of some integral, when can interchange δ and $\int$:

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

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

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

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 $\delta(Kf) = K\delta f$.

A few more examples:

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

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

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

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

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

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).

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

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

Some mathematicians would say the above calculation is incorrect, because $\delta^2(x-t)$ is undefined. But that's not true, because 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 $\alpha \to \infty$:

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

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

$$= \int 3f^2(x)\lim_{\alpha\to\infty}\delta_\alpha(x-t)\mathrm{d}x = \int 3f^2(x)\delta(x-t)\mathrm{d}x = 3f^2(t) \tag{1.6}$$

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.

1.10 Spherical Harmonics

Are defined 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 defined by the 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

$$\sum_{l=0}^{\infty} \frac{2l+1}{2} P_l(x') P_l(x) = \delta(x - x') \tag{1.7}$$

The spherical harmonics are ortonormal:

$$\int Y_{lm}\, Y^*_{l'm'}\, \mathrm{d}\Omega = \int_0^{2\pi} \int_0^{\pi} Y_{lm}(\theta,\phi)\, Y^*_{l'm'}(\theta,\phi) \sin\theta\, \mathrm{d}\theta\, \mathrm{d}\phi = \delta_{mm'}\delta_{ll'} \tag{1.8}$$

and complete (both in the l-subspace and the whole space):

$$\sum_{m=-l}^{l} |Y_{lm}(\theta,\phi)|^2 = \frac{2l+1}{4\pi} \tag{1.9}$$

$$\sum_{l=0}^{\infty} \sum_{m=-l}^{l} Y_{lm}(\theta,\phi) Y^*_{lm}(\theta',\phi') = \frac{1}{\sin\theta} \delta(\theta-\theta')\delta(\phi-\phi') = \delta(\hat{\mathbf{r}} - \hat{\mathbf{r}}') \tag{1.10}$$

The relation (1.9) is a special case of an addition theorem for spherical harmonics

$$\sum_{m=-l}^{l} Y_{lm}(\theta,\phi) Y^*_{lm}(\theta',\phi') = \frac{4\pi}{2l+1} P_l(\cos\gamma) \tag{1.11}$$

where γ is the angle between the unit vectors given by $\hat{\mathbf{r}} = (\theta,\phi)$ and $\hat{\mathbf{r}}' = (\theta',\phi')$:

$$\cos\gamma = \cos\theta\cos\theta' + \sin\theta\sin\theta'\cos(\phi-\phi') = \hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'$$

1.11 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 (1.8) 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 (1.10) 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 (1.11):

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

Using the completeness relation (1.7):

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

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

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

$$f(\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}') = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} Y_{lm}(\hat{\mathbf{r}}) f_l Y^*_{lm}(\hat{\mathbf{r}}') \tag{1.12}$$

1.12 Homogeneous functions

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:

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

1.12.1 Example

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 obviously is true.

1.13 Differential Geometry

1.13.1 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:

$$U^\lambda_0 = (1, 0, 0, 0)$$

$$\frac{\mathrm{d}S^i}{\mathrm{d}t} = 0$$

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 $\vec e^\mu_0 \equiv U^\mu$, 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 $U^\beta \partial_\beta U^\alpha = \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda}\frac{\partial}{\partial x^\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda} = \frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}\lambda^2}$):

$$\frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}\lambda^2} + \Gamma^\alpha_{\beta\gamma}\frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\gamma}{\mathrm{d}\lambda} = 0$$

Curvature

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 $[\nabla_\alpha, \nabla_\beta]V^\mu$), see how it changes and that's the curvature:

$$[\nabla_\alpha, \nabla_\beta]V^\mu \equiv R^\mu{}_{\nu\alpha\beta} V^\nu$$

That's all there is to it. Expanding the left hand side:

$$[\nabla_\alpha, \nabla_\beta]V^\mu = \left(\partial_\alpha \Gamma^\mu_{\beta\nu} - \partial_\beta \Gamma^\mu_{\alpha\nu} + \Gamma^\mu_{\alpha\sigma}\Gamma^\sigma_{\beta\nu} - \Gamma^\mu_{\beta\sigma}\Gamma^\sigma_{\alpha\nu}\right) V^\nu$$

we get

$$R^\mu{}_{\nu\alpha\beta} = \partial_\alpha \Gamma^\mu_{\beta\nu} - \partial_\beta \Gamma^\mu_{\alpha\nu} + \Gamma^\mu_{\alpha\sigma}\Gamma^\sigma_{\beta\nu} - \Gamma^\mu_{\beta\sigma}\Gamma^\sigma_{\alpha\nu}$$

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 in spherical and cylindrical coordinates), but for coordinate bases they are. It can be proven, that

$$\Gamma^\mu_{\alpha\beta} = \frac{1}{2}g^{\mu\sigma}\left(\partial_\beta g_{\sigma\alpha} + \partial_\alpha g_{\sigma\beta} - \partial_\sigma g_{\alpha\beta} + c_{\alpha\sigma\beta} + c_{\beta\sigma\alpha} - c_{\sigma\alpha\beta}\right)$$

and for coordinate bases ${c^\alpha}_{\mu\nu} = 0$, so

$$\Gamma^\mu_{\alpha\beta} = \Gamma^\mu_{\beta\alpha}$$

$$\Gamma^\mu_{\alpha\beta} = \frac{1}{2}g^{\mu\sigma}\left(\partial_\beta g_{\sigma\alpha} + \partial_\alpha g_{\sigma\beta} - \partial_\sigma g_{\alpha\beta}\right)$$

As a special case:

$$\Gamma^\mu_{\mu\beta} = \frac{1}{2}g^{\mu\sigma}\left(\partial_\beta g_{\sigma\mu} + \partial_\mu g_{\sigma\beta} - \partial_\sigma g_{\mu\beta}\right) = \frac{1}{2}g^{\mu\sigma}\partial_\beta g_{\sigma\mu} =$$

$$= \frac{1}{2}\mathrm{Tr}\, g^{-1}\partial_\beta g = \frac{1}{2}\mathrm{Tr}\,\partial_\beta \log g = \frac{1}{2}\partial_\beta \mathrm{Tr}\,\log g = \frac{1}{2}\partial_\beta \log|\det g| = \partial_\beta \log\sqrt{|\det g|} =$$

$$= \frac{1}{2\det g}\partial_\beta \det g = \frac{1}{\sqrt{|\det g|}}\partial_\beta\sqrt{|\det g|}$$

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 unsymmetric. Later we used the identity Tr $\log g = \log|\det g|$, which follows from the well-known identity $\det\exp A = \exp\mathrm{Tr}\,A$ by substituting $A = \log g$ and taking the logarithm of both sides.

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.

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

$$\sqrt{|\det g_{ij}|} = h_1 h_2 h_3$$
$$g^{ij} = \begin{pmatrix} \frac{1}{h_1^2} & 0 & 0 \\ 0 & \frac{1}{h_2^2} & 0 \\ 0 & 0 & \frac{1}{h_3^2} \end{pmatrix}$$

and

$$\nabla \cdot \mathbf{A} = \nabla_i A^i = \frac{1}{h_1 h_2 h_3} \partial_i \left(h_1 h_2 h_3 A^i\right)$$

Laplace Operator

$$\nabla^2 \varphi = \nabla_\mu \nabla^\mu \varphi = \partial_\mu \nabla^\mu \varphi + \Gamma^\mu_{\mu\sigma} \nabla^\sigma \varphi = \partial_\mu \partial^\mu \varphi + \Gamma^\mu_{\mu\sigma} \partial^\sigma \varphi =$$
$$= \partial_\mu \partial^\mu \varphi + \frac{1}{\sqrt{|\det g|}} \left(\partial_\sigma \sqrt{|\det g|}\right) \partial^\sigma \varphi =$$
$$= \frac{1}{\sqrt{|\det g|}} \partial_\mu \left(\sqrt{|\det g|}\, \partial^\mu \varphi\right) = \frac{1}{\sqrt{|\det g|}} \partial_\mu \left(\sqrt{|\det g|}\, g^{\mu\sigma} \partial_\sigma \varphi\right)$$

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:

$$\int_\Omega \nabla_\mu u^\mu \sqrt{|g|}\mathrm{d}^4 x = \int_\Omega \frac{1}{\sqrt{|g|}}\partial_\mu\left(\sqrt{|g|}u^\mu\right)\sqrt{|g|}\mathrm{d}^4 x = \int_\Omega \partial_\mu\left(\sqrt{|g|}u^\mu\right)\mathrm{d}^4 x =$$
$$= \int_{\partial\Omega} \sqrt{|g|}u^\mu n_\mu \mathrm{d}^3 x = \int_{\partial\Omega} u^\mu n_\mu \sqrt{|g|}\mathrm{d}^3 x$$

where $\partial\Omega$ is the boundary (surface) of Ω and n_ν is the normal vector to this surface.

1.13.2 Examples

Weak Formulation of Laplace Equation

As an example, we write the weak formulation of the Laplace equation in arbitrary coordintes:

$$\nabla^2\varphi - f = 0$$
$$\int \left(\nabla^2\varphi v - fv\right)\sqrt{|g|}d^3 x = 0$$
$$\int \left(\frac{1}{\sqrt{|g|}}\partial_i\left(\sqrt{|g|}g^{ij}\partial_j\varphi\right)v - fv\right)\sqrt{|g|}d^3 x = 0$$
$$\int \left(\partial_i\left(\sqrt{|g|}g^{ij}\partial_j\varphi\right)v - fv\sqrt{|g|}\right)d^3 x = 0$$

Now we apply per-partes (assuming the boundary integral vanishes):

$$\int \left(-\sqrt{|g|}g^{ij}\partial_j\varphi\partial_i v - fv\sqrt{|g|}\right)d^3 x = 0$$
$$\int \left(-g^{ij}\partial_j\varphi\partial_i v - fv\right)\sqrt{|g|}d^3 x = 0$$

For diagonal metric this evaluates to:

$$\int \left(-\sum_i \frac{1}{h_i^2}\partial_i\varphi\partial_i v - fv\right)h_1 h_2 h_3 d^3 x = 0$$

Cylindrical Coordinates

$$x = \rho\cos\phi$$

$$y = \rho\sin\phi$$

$$z = z$$

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 $x^i = (\rho,\phi,z)$:

$$g_{ij} = \frac{\partial\hat{x}^a}{\partial x^i}\frac{\partial\hat{x}^b}{\partial x^j}\hat{g}_{ab} = \left(\frac{\partial\hat{x}}{\partial x}\right)^T\hat{g}\frac{\partial\hat{x}}{\partial x} =$$

$$= \left(\frac{\partial(x,y,z)}{\partial(\rho,\phi,z)}\right)^T\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\frac{\partial(x,y,z)}{\partial(\rho,\phi,z)} =$$

$$= \begin{pmatrix}\cos\phi & \sin\phi & 0\\ -\rho\sin\phi & \rho\cos\phi & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\cos\phi & -\rho\sin\phi & 0\\ \sin\phi & \rho\cos\phi & 0\\ 0 & 0 & 1\end{pmatrix} = \begin{pmatrix}1 & 0 & 0\\ 0 & \rho^2 & 0\\ 0 & 0 & 1\end{pmatrix}$$

$$g^{ij} = \begin{pmatrix}1 & 0 & 0\\ 0 & \frac{1}{\rho^2} & 0\\ 0 & 0 & 1\end{pmatrix}$$

$$\det g = \det g_{ij} = \rho^2$$

$$\nabla^i\nabla_i\varphi = \frac{1}{\sqrt{|\det g|}}\partial_i\left(\sqrt{|\det g|}\,g^{ij}\partial_j\varphi\right) =$$

$$= \frac{1}{\rho}\partial_i\left(\rho g^{ij}\partial_j\varphi\right) = \frac{1}{\rho}\partial_\rho\left(\rho\partial_\rho\varphi\right) + \frac{1}{\rho}\partial_\phi\left(\rho\frac{1}{\rho^2}\partial_\phi\varphi\right) + \frac{1}{\rho}\partial_z\left(\rho\partial_z\varphi\right) =$$
$$= \frac{1}{\rho}\partial_\rho\left(\rho\partial_\rho\varphi\right) + \frac{1}{\rho^2}\partial_\phi\partial_\phi\varphi + \partial_z\partial_z\varphi =$$

$$= \partial_\rho\partial_\rho\varphi + \frac{1}{\rho}\partial_\rho\varphi + \frac{1}{\rho^2}\partial_\phi\partial_\phi\varphi + \partial_z\partial_z\varphi$$

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

$$x=\rho\sin\theta\cos\phi$$

$$y=\rho\sin\theta\sin\phi$$

$$z=\rho\cos\theta$$

The transformation matrix is

$$\frac{\partial(x,y,z)}{\partial(\rho,\theta,\phi)}=\begin{pmatrix}\sin\theta\cos\phi & \rho\cos\theta\cos\phi & -\rho\sin\theta\sin\phi\\ \sin\theta\sin\phi & \rho\cos\theta\sin\phi & \rho\sin\theta\cos\phi\\ \cos\phi & -\rho\sin\theta & 0\end{pmatrix}$$

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 spherical coordinates $x^i=(\rho,\theta,\phi)$:

$$g_{ij}=\frac{\partial\hat{x}^a}{\partial x^i}\frac{\partial\hat{x}^b}{\partial x^j}\hat{g}_{ab}=\left(\frac{\partial\hat{x}}{\partial x}\right)^T\hat{g}\frac{\partial\hat{x}}{\partial x}=$$

$$=\left(\frac{\partial(x,y,z)}{\partial(\rho,\theta,\phi)}\right)^T\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}\frac{\partial(x,y,z)}{\partial(\rho,\theta,\phi)}=$$

$$=\begin{pmatrix}\sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta\\ \rho\cos\theta\cos\phi & \rho\cos\theta\sin\phi & -\rho\sin\theta\\ -\rho\sin\theta\sin\phi & \rho\sin\theta\cos\phi & 0\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}\sin\theta\cos\phi & \rho\cos\theta\cos\phi & -\rho\sin\theta\sin\phi\\ \sin\theta\sin\phi & \rho\cos\theta\sin\phi & \rho\sin\theta\cos\phi\\ \cos\phi & -\rho\sin\theta & 0\end{pmatrix}=$$

$$=\begin{pmatrix}1&0&0\\0&\rho^2&0\\0&0&\rho^2\sin^2\theta\end{pmatrix}$$

$$g^{ij}=\begin{pmatrix}1&0&0\\0&\frac{1}{\rho^2}&0\\0&0&\frac{1}{\rho^2\sin^2\theta}\end{pmatrix}$$

$$\det g=\det g_{ij}=\rho^4\sin^2\theta$$

$$\nabla^i\nabla_i\varphi = \partial^i\partial_i\varphi + \frac{1}{2\det g}\partial_j(\det g)\ g^{jk}\partial_k\varphi =$$

$$= g^{ij}\partial_i\partial_j\varphi + \frac{1}{2\rho^4\sin^2\theta}\left(\partial_\rho(\rho^4\sin^2\theta)\ g^{\rho\rho}\partial_\rho\varphi + \partial_\theta(\rho^4\sin^2\theta)\ g^{\theta\theta}\partial_\theta\varphi\right)$$

$$= g^{ij}\partial_i\partial_j\varphi + \frac{2}{\rho}\partial_\rho\varphi + \frac{\cos\theta}{\rho^2\sin\theta}\partial_\theta\varphi =$$

$$= \partial_\rho\partial_\rho\varphi + \frac{1}{\rho^2}\partial_\theta\partial_\theta\varphi + \frac{1}{\rho^2\sin^2\theta}\partial_\phi\partial_\phi\varphi + \frac{2}{\rho}\partial_\rho\varphi + \frac{\cos\theta}{\rho^2\sin\theta}\partial_\theta\varphi$$

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:

$$\mathbf{F} = -m\nabla\phi = m\,(x', y', 0)\,\omega^2 = m\mathbf{r}'\omega^2$$

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:

$$\frac{\mathrm{d}^2x^0}{\mathrm{d}\lambda^2} = 0$$

$$\frac{\mathrm{d}^2x^1}{\mathrm{d}\lambda^2} + \Gamma^1_{00}\left(\frac{\mathrm{d}x^0}{\mathrm{d}\lambda}\right)^2 + 2\Gamma^1_{20}\frac{\mathrm{d}x^2}{\mathrm{d}\lambda}\frac{\mathrm{d}x^0}{\mathrm{d}\lambda} = 0$$

$$\frac{\mathrm{d}^2x^2}{\mathrm{d}\lambda^2} + \Gamma^2_{00}\left(\frac{\mathrm{d}x^0}{\mathrm{d}\lambda}\right)^2 + 2\Gamma^2_{10}\frac{\mathrm{d}x^1}{\mathrm{d}\lambda}\frac{\mathrm{d}x^0}{\mathrm{d}\lambda} = 0$$

$$\frac{\mathrm{d}^2x^3}{\mathrm{d}\lambda^2} = 0$$

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):

$$\nabla_0 V'^\mu = \partial_0 V'^\mu + \Gamma^\mu_{0\alpha}V'^\alpha$$

$$\nabla_0\begin{pmatrix}V'^0\\V'^1\\V'^2\\V'^3\end{pmatrix} = \begin{pmatrix}\partial_0 V'^0\\ \partial_0 V'^1 + \Gamma^1_{00}V'^0 + \Gamma^1_{02}V'^2\\ \partial_0 V'^2 + \Gamma^2_{00}V'^0 + \Gamma^2_{01}V'^1\\ \partial_0 V'^3\end{pmatrix} = \begin{pmatrix}\partial_0 V'^0\\ \partial_0 V'^1 - x'\omega^2 V'^0 - \omega V'^2\\ \partial_0 V'^2 - y'\omega^2 V'^0 + \omega V'^1\\ \partial_0 V'^3\end{pmatrix} =$$

$$= \partial_0\begin{pmatrix}V'^0\\V'^1\\V'^2\\V'^3\end{pmatrix} + \begin{pmatrix}0 & 0 & 0 & 0\\ -x'\omega^2 & 0 & -\omega & 0\\ -y'\omega^2 & \omega & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}\begin{pmatrix}V'^0\\V'^1\\V'^2\\V'^3\end{pmatrix} \tag{1.13}$$

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:

$$\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t} = \boldsymbol{\omega}\times\mathbf{A} + \frac{\mathrm{d}'\mathbf{A}}{\mathrm{d}t} \tag{1.14}$$

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:

$$(d\mathbf{G})_{\rm space} = (d\mathbf{G})_{\rm body} + (d\mathbf{G})_{\rm rot}$$

Now consider a vector fixed in the rigid body. Then $(d\mathbf{G})_{\rm body} = 0$ and

$$(d\mathbf{G})_{\rm rot} = (d\mathbf{G})_{\rm space} = d\mathbf{\Omega} \times \mathbf{G}$$

For an arbitrary vector, the change relative to the space axes is the sum of the two effects:

$$(d\mathbf{G})_{\rm space} = (d\mathbf{G})_{\rm body} + d\mathbf{\Omega} \times \mathbf{G}$$

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:

$$\boldsymbol{\omega} = \frac{d\mathbf{\Omega}}{dt}$$

we get

$$\left(\frac{d\mathbf{G}}{dt}\right)_{\rm space} = \left(\frac{d\mathbf{G}}{dt}\right)_{\rm body} + \boldsymbol{\omega} \times \mathbf{G}$$

CHAPTER

TWO

CLASSICAL MECHANICS, SPECIAL AND GENERAL RELATIVITY

2.1 Newtonian Physics

2.1.1 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.

2.1.2 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 footnote{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:

$$\mathbf{F} = -G\frac{m_1 m_2}{r^2}\mathbf{n}$$

}:

$$\nabla\cdot\mathbf{F} = -4\pi Gm\,\rho(t,x,y,z)$$

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

$$\nabla^2\phi = 4\pi G\rho \tag{2.1}$$

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:

$$[r] = [\mathbf{x}] = \mathrm{m}$$

$$[m] = \mathrm{kg}$$

$$[\rho] = \mathrm{kg\,m^{-3}}$$

$$[F] = \mathrm{kg\,m\,s^{-2}}$$

$$[G] = \mathrm{kg^{-1}\,m^3\,s^{-2}}$$

$$[\phi] = \mathrm{m^2\,s^{-2}}$$

2.1.3 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

$$x^\mu = \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}$$

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:

$$\frac{\mathrm{d}^2 x^0}{\mathrm{d}\lambda^2} = 0$$

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}\lambda^2} + \delta^{ij}\partial_j\phi\left(\frac{\mathrm{d}x^0}{\mathrm{d}\lambda}\right)^2 = 0$$

But this is exactly the geodesic equation for the following Christoffel symbols:

$$\Gamma^i_{00} = \delta^{ij}\partial_j\phi \tag{2.2}$$

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:

$$R^j{}_{0k0} = -R^j{}_{00k} = \delta^{ji}\partial_i\partial_k\phi$$

we calculate the $R_{\alpha\beta}$ by contracting:

$$R_{00} = R^\mu{}_{0\mu0} = R^i{}_{0i0} = \delta^{ij}\partial_i\partial_j\phi$$

$$R_{ij} = 0$$

and we see that the Newton gravitation law is

$$R_{00} = 4\pi G\rho$$

$$R_{ij} = 0$$

Thus we have reformulated the Newton's laws in a frame invariant way — the matter curves the geometry using the equations:

$$R_{00} = 4\pi G\rho$$

$$R_{ij} = 0$$

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.

2.1.4 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 $c \to \infty$):

$$\Gamma^0_{\mu\nu} = \begin{pmatrix} - \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{- 2 \phi(t,x,y,z) - c^{2}} & - \frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{- 2 \phi(t,x,y,z) - c^{2}} & - \frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{- 2 \phi(t,x,y,z) - c^{2}} & - \frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{- 2 \phi(t,x,y,z) - c^{2}} \\ - \frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{- 2 \phi(t,x,y,z) - c^{2}} & \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{c^{2} \left(- 2 \phi(t,x,y,z) - c^{2}\right)} & 0 & 0 \\ - \frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{- 2 \phi(t,x,y,z) - c^{2}} & 0 & \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{c^{2} \left(- 2 \phi(t,x,y,z) - c^{2}\right)} & 0 \\ - \frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{- 2 \phi(t,x,y,z) - c^{2}} & 0 & 0 & \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{c^{2} \left(- 2 \phi(t,x,y,z) - c^{2}\right)} \end{pmatrix}$$

$$\Gamma^1_{\mu\nu} = \begin{pmatrix} \frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}} & - \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} & 0 & 0 \\ - \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} & - \frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} & - \frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} & - \frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} \\ 0 & - \frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} & \frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} & 0 \\ 0 & - \frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} & 0 & \frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} \end{pmatrix}$$

$$\Gamma^2_{\mu\nu} = \begin{pmatrix} \frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}} & 0 & - \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} & 0 \\ 0 & \frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} & - \frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} & 0 \\ - \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} & - \frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} & - \frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} & - \frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} \\ 0 & 0 & - \frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} & \frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} \end{pmatrix}$$

$$\Gamma^3_{\mu\nu} = \begin{pmatrix} \frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}} & 0 & 0 & - \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} \\ 0 & \frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} & 0 & - \frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} \\ 0 & 0 & \frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} & - \frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} \\ - \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} & - \frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} & - \frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} & - \frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{c^{2} \left(1 - 2 \frac{\phi(t,x,y,z)}{c^{2}}\right)} \end{pmatrix}$$

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

$$\frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}\lambda^2} + \Gamma^\alpha_{\beta\gamma} \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\gamma}{\mathrm{d}\lambda} = 0$$

becomes

$$\frac{\mathrm{d}^2 x^0}{\mathrm{d}\lambda^2} = 0$$

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}\lambda^2} + \delta^{ij} \partial_j \phi \left(\frac{\mathrm{d}x^0}{\mathrm{d}\lambda} \right)^2 = 0$$

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 textit{is} the equation of geodesics.

2.1.5 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 equations:

$$R_{\alpha\beta} - \tfrac{1}{2} R g_{\alpha\beta} = \frac{8\pi G}{c^4} T_{\alpha\beta}$$

or

$$R_{\alpha\beta} = \frac{8\pi G}{c^4}(T_{\alpha\beta} - \tfrac{1}{2} T g_{\alpha\beta})$$

$$R^\alpha{}_\beta = \frac{8\pi G}{c^4}(T^\alpha{}_\beta - \tfrac{1}{2} T)$$

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 (2.1).

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 textit{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 $c = 1$):

$$\Gamma^0_{\mu\nu} = \left(\begin{matrix}- \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & - \frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & - \frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & - \frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)}\\- \frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & 0 & 0\\- \frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & 0 & \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & 0\\- \frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & 0 & 0 & \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)}\end{matrix}\right)$$

$$\Gamma^1_{\mu\nu} = \left(\begin{matrix}\frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & - \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & 0 & 0\\- \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & - \frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & - \frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & - \frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)}\\0 & - \frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & \frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & 0\\0 & - \frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & 0 & \frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)}\end{matrix}\right)$$

$$\Gamma^2_{\mu\nu} = \left(\begin{matrix}\frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & 0 & - \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & 0\\0 & \frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & - \frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & 0\\- \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & - \frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & - \frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & - \frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)}\\0 & 0 & - \frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & \frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)}\end{matrix}\right)$$

$$\Gamma^3_{\mu\nu} = \left(\begin{matrix}\frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & 0 & 0 & - \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)}\\0 & \frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & 0 & - \frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)}\\0 & 0 & \frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & - \frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)}\\- \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & - \frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & - \frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)} & - \frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{1 - 2 \phi(t,x,y,z)}\end{matrix}\right)$$

and in the second case, only $\Gamma^0_{\mu\nu}$ is different:

$$\Gamma^0_{\mu\nu} = \left(\begin{matrix}\frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{1 + 2 \phi(t,x,y,z)} & \frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{1 + 2 \phi(t,x,y,z)} & \frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{1 + 2 \phi(t,x,y,z)} & \frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{1 + 2 \phi(t,x,y,z)}\\\frac{\frac{\partial}{\partial x} \phi(t,x,y,z)}{1 + 2 \phi(t,x,y,z)} & - \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{1 + 2 \phi(t,x,y,z)} & 0 & 0\\\frac{\frac{\partial}{\partial y} \phi(t,x,y,z)}{1 + 2 \phi(t,x,y,z)} & 0 & - \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{1 + 2 \phi(t,x,y,z)} & 0\\\frac{\frac{\partial}{\partial z} \phi(t,x,y,z)}{1 + 2 \phi(t,x,y,z)} & 0 & 0 & - \frac{\frac{\partial}{\partial t} \phi(t,x,y,z)}{1 + 2 \phi(t,x,y,z)}\end{matrix}\right)$$

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 (2.2) 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 (2.2) 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.

2.1.6 Intertial 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 textit{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 intertial 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).

2.1.7 Lorentz Group

The Lorentz group is O(3,1), e.g. all matrices satisfying:

$$g = \Lambda^T g \Lambda \tag{2.3}$$

with $g = \mathrm{diag}(-c^2, 1, 1, 1)$. Taking the determinant of (2.3) we get $(\det \Lambda)^2 = 1$ or $\det \Lambda = \pm 1$. Writing the 00 component of (2.3) 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 space inversion (also called space reflection or parity transformation):

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

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 $A^0{}_0 \le -1$. 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 $A^0{}_0 \ge 1$. 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 $A^0{}_0 \le -1$. 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 $\Lambda^0{}_0 \le -1$ can be written using $-\mathbb{1}$ and a matrix with $\Lambda^0{}_0 \ge 1$. 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:

$$\begin{aligned}
\mathrm{diag}(-1,0,0,0) &= T\\
\mathrm{diag}(0,-1,0,0) &= PR_1(\pi)\\
\mathrm{diag}(0,0,-1,0) &= PR_2(\pi)\\
\mathrm{diag}(0,0,0,-1) &= PR_3(\pi)\\
\mathrm{diag}(0,-1,-1,-1) &= P\\
\mathrm{diag}(-1,0,-1,-1) &= TR_1(\pi)\\
\mathrm{diag}(-1,-1,0,-1) &= TR_2(\pi)\\
\mathrm{diag}(-1,-1,-1,0) &= TR_3(\pi)
\end{aligned}$$

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:

$$\det A = \det e^L = e^{\mathrm{Tr}\, L} = 1$$

so $\mathrm{Tr}\, L = 0$ and L is a real, traceless matrix. Rewriting (2.3):

$$g = A^T g A$$

$$A^{-1} = g^{-1} A^T g$$

$$e^{-L} = g^{-1} e^{L^T} g = e^{g^{-1} L^T g}$$

$$-L = g^{-1} L^T g$$

$$-gL = (gL)^T$$

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 $\mathrm{diag}(-1,1,1,1)$, 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 $C = \mathrm{diag}(c, 1, 1, 1)$ 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 $\zeta = \frac{v}{c}\,\mathrm{atanh}\left(\frac{v}{c}\right)$, 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:

$$A = \begin{pmatrix} \cosh\zeta & -\frac{1}{c}\sinh\zeta & 0 & 0 \\ -c\sinh\zeta & \cosh\zeta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & -\frac{\frac{v}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 \\ -\frac{v}{\sqrt{1-\frac{v^2}{c^2}}} & \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

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

$$w = \frac{u + v}{1 + \frac{uv}{c^2}}$$

2.1.8 O(4) Group

The group of rotations in 4 dimensions is O(4), e.g. all matrices satisfying:

$$g = \Lambda^T g \Lambda \tag{2.4}$$

with $g = \mathrm{diag}(c^2, 1, 1, 1)$. Taking the determinant of (2.4) we get $(\det \Lambda)^2 = 1$ or $\det \Lambda = \pm 1$. Writing the 00 component of (2.4) 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.

2.1.9 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.

2.1.10 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.

2.1.11 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 $g_{00} = -c^2$?)

What are all the allowed transformations between intertial 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.

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.

CHAPTER

THREE

FLUID DYNAMICS

3.1 Fluid Dynamics

3.1.1 Stress-Energy Tensor

In general, the stress energy tensor is the flux of momentum p^μ over the surface x^ν. The Navier-Stokes equations can be derived from the conservation law:

$$\partial_\nu T^{\mu\nu} + f^\mu = 0$$

3.1.2 Navier-Stokes Equations

When we write the above conservation law in a nonrelativistic limit, we get 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 Newtonean 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 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:

$$\nabla\cdot\mathbb{T} = \mu\nabla^2\mathbf{v} + (\mu+\lambda)\nabla\nabla\cdot\mathbf{v}$$

For incompressible fluid we have $\nabla\cdot\mathbf{v} = \mathbf{0}$, so we get:

$$\nabla\cdot\mathbb{T} = \mu\nabla^2\mathbf{v}$$

and for a perfect fluid we have no viscosity, e.g. $\mu = 0$, so:

$$\nabla\cdot\mathbb{T} = 0$$

and the equations are then called Euler equations (for perfect fluid).

3.1.3 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:

$$\frac{v^2}{2} + gz + \frac{p}{\rho} = \mathrm{const.}$$

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:

$$\frac{v^2}{2} + gz + \frac{p}{\rho} = \mathrm{const.}$$

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.

3.2 MHD Equations

3.2.1 Introduction

The magnetohydrodynamics (MHD) equations are:

$$\frac{\partial\rho}{\partial t}+\nabla\cdot(\rho\mathbf{v})=0 \tag{3.1}$$

$$\rho\left(\frac{\partial\mathbf{v}}{\partial t}+(\mathbf{v}\cdot\nabla)\mathbf{v}\right)=-\nabla p+\frac{1}{\mu}(\nabla\times\mathbf{B})\times\mathbf{B}+\rho\mathbf{g} \tag{3.2}$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) + \eta \nabla^2 \mathbf{B} \tag{3.3}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{3.4}$$

assuming η is constant. See the next section for a derivation. We can now apply the following identities (we use the fact that $\nabla \cdot \mathbf{B} = 0$):

$$\begin{aligned}
[(\nabla \times \mathbf{B}) \times \mathbf{B}]_i = \varepsilon_{ijk}(\nabla \times \mathbf{B})_j B_k = \varepsilon_{ijk}\varepsilon_{jlm}(\partial_l B_m)B_k = (\delta_{kl}\delta_{im} - \delta_{km}\delta_{il})(\partial_l B_m)B_k = \\
= (\partial_k B_i)B_k - (\partial_i B_k)B_k = \left[(\mathbf{B} \cdot \nabla)\mathbf{B} - \frac{1}{2}\nabla|\mathbf{B}|^2\right]_i \\
(\nabla \times \mathbf{B}) \times \mathbf{B} = (\mathbf{B} \cdot \nabla)\mathbf{B} - \frac{1}{2}\nabla|\mathbf{B}|^2 = (\mathbf{B} \cdot \nabla)\mathbf{B} + \mathbf{B}(\nabla \cdot \mathbf{B}) - \frac{1}{2}\nabla|\mathbf{B}|^2 = \nabla \cdot (\mathbf{B}\mathbf{B}^T) - \frac{1}{2}\nabla|\mathbf{B}|^2 \\
\nabla \times (\mathbf{v} \times \mathbf{B}) = (\mathbf{B} \cdot \nabla)\mathbf{v} - \mathbf{B}(\nabla \cdot \mathbf{v}) + \mathbf{v}(\nabla \cdot \mathbf{B}) - (\mathbf{v} \cdot \nabla)\mathbf{B} = \nabla \cdot (\mathbf{B}\mathbf{v}^T - \mathbf{v}\mathbf{B}^T) \\
\nabla \cdot (\rho \mathbf{v}\mathbf{v}^T) = (\nabla \cdot (\rho \mathbf{v}))\,\mathbf{v} + \rho(\mathbf{v} \cdot \nabla)\mathbf{v} = -\mathbf{v}\frac{\partial \rho}{\partial t} + \rho(\mathbf{v} \cdot \nabla)\mathbf{v}
\end{aligned}$$

So the MHD equations can alternatively be written as:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \tag{3.5}$$

$$\frac{\partial \rho \mathbf{v}}{\partial t} + \nabla \cdot (\rho \mathbf{v}\mathbf{v}^T) = -\nabla p + \frac{1}{\mu}\left(\nabla \cdot (\mathbf{B}\mathbf{B}^T) - \frac{1}{2}\nabla|\mathbf{B}|^2\right) + \rho \mathbf{g} \tag{3.6}$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \cdot (\mathbf{B}\mathbf{v}^T - \mathbf{v}\mathbf{B}^T) + \eta \nabla^2 \mathbf{B} \tag{3.7}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{3.8}$$

One can also introduce a new variable $p^* = p + \frac{1}{2}\nabla|\mathbf{B}|^2$, that simplifies (3.6) a bit.

3.2.2 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:

$$\nabla \cdot \mathbf{B} = 0$$

3.2.3 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):

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} - (\mathbf{B} \cdot \nabla)\mathbf{B} + \nabla p = 0 \tag{3.9}$$

$$\frac{\partial \mathbf{B}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{B} - (\mathbf{B} \cdot \nabla)\mathbf{u} = 0 \tag{3.10}$$

$$\nabla \cdot \mathbf{u} = 0 \tag{3.11}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{3.12}$$

If the equation (3.12) 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 (3.10), resp. (3.3)) 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 (3.9), (3.10) and (3.11), 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:

$$\frac{\mathbf{u}^n - \mathbf{u}^{n-1}}{\tau} + (\mathbf{u}^{n-1} \cdot \nabla)\mathbf{u}^n - (\mathbf{B}^{n-1} \cdot \nabla)\mathbf{B}^n + \nabla p = 0 \tag{3.13}$$

$$\frac{\mathbf{B}^n - \mathbf{B}^{n-1}}{\tau} + (\mathbf{u}^{n-1}\cdot\nabla)\mathbf{B}^n - (\mathbf{B}^{n-1}\cdot\nabla)\mathbf{u}^n = 0 \tag{3.14}$$

$$\nabla\cdot\mathbf{u}^n = 0 \tag{3.15}$$

Testing (3.13) by the test functions (v_1, v_2), (3.14) by the functions (C_1, C_2) and (3.15) by the test function q, we obtain the following weak formulation:

$$\begin{aligned}
\int_\Omega \frac{u_1 v_1}{\tau} + (\mathbf{u}^{n-1}\cdot\nabla)u_1 v_1 - (\mathbf{B}^{n-1}\cdot\nabla)B_1 v_1 - p\frac{\partial v_1}{\partial x}\,\mathrm{d}\mathbf{x} &= \int_\Omega \frac{u_1^{n-1} v_1}{\tau}\,\mathrm{d}\mathbf{x} \\
\int_\Omega \frac{u_2 v_2}{\tau} + (\mathbf{u}^{n-1}\cdot\nabla)u_2 v_2 - (\mathbf{B}^{n-1}\cdot\nabla)B_2 v_2 - p\frac{\partial v_2}{\partial y}\,\mathrm{d}\mathbf{x} &= \int_\Omega \frac{u_2^{n-1} v_2}{\tau}\,\mathrm{d}\mathbf{x}
\end{aligned} \tag{3.16}$$

$$\begin{aligned}
\int_\Omega \frac{B_1 C_1}{\tau} + (\mathbf{u}^{n-1}\cdot\nabla)B_1 C_1 - (\mathbf{B}^{n-1}\cdot\nabla)u_1 C_1\,\mathrm{d}\mathbf{x} &= \int_\Omega \frac{B_1^{n-1} C_1}{\tau}\,\mathrm{d}\mathbf{x} \\
\int_\Omega \frac{B_2 C_2}{\tau} + (\mathbf{u}^{n-1}\cdot\nabla)B_2 C_2 - (\mathbf{B}^{n-1}\cdot\nabla)u_2 C_2\,\mathrm{d}\mathbf{x} &= \int_\Omega \frac{B_2^{n-1} C_2}{\tau}\,\mathrm{d}\mathbf{x}
\end{aligned} \tag{3.17}$$

$$\int_\Omega \frac{\partial u_1}{\partial x} q + \frac{\partial u_2}{\partial y} q\,\mathrm{d}\mathbf{x} = 0 \tag{3.18}$$

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:

$$\begin{array}{llllllll}
+A(u_1,v_1) & & -X(p,v_1) & -B(B_1,v_1) & & = & l_1(v_1) \\
& +A(u_2,v_2) & -Y(p,v_2) & & -B(B_2,v_2) & = & l_2(v_2) \\
+X(q,u_1) & +Y(q,u_2) & & & & = & 0 \\
-B(u_1,C_1) & & & +A(B_1,C_1) & & = & l_4(C_1) \\
& -B(u_2,C_2) & & & +A(B_2,C_2) & = & l_5(C_2)
\end{array}$$

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 $l1$, ..., $l5$ are the same as above.

3.3 Compressible Euler Equations

The compressible Euler equations are:

$$\begin{gathered}
\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{u}) = 0\\
\frac{\partial(\rho\mathbf{u})}{\partial t} + \nabla\cdot(\rho\mathbf{u}\mathbf{u}^T) + \nabla p - \mathbf{f} = 0\\
\frac{\partial E}{\partial t} + \nabla\cdot(\mathbf{u}(E+p)) = 0
\end{gathered}$$

where

$$E = \rho e + \tfrac{1}{2}\rho u^2$$

is the total energy per unit volume ($\frac{1}{2}\rho u^2$ is the kinetic energy per unit volume), e is the internal energy per unit mass ($e = \frac{U}{nM}$) and we use the ideal gas equations, so:

$$
\begin{aligned}
e &= Tc_v \\
p &= \frac{n}{V}\bar{R}T = \frac{nM}{V}\frac{\bar{R}}{M}T = \rho RT = \rho R\frac{e}{c_v} = \frac{R}{c_v}(E - \tfrac{1}{2}\rho u^2)
\end{aligned}
$$

where n is the number of moles of gas, M is the molar mass of the gas (e.g. a mass of one mole of the gas), $\rho = \frac{nM}{V}$ is the density of the gas, $\bar{R}$ is the ideal gas constant, $R = \frac{\bar{R}}{M}$ is the specific ideal gas constant, c_v is the specific heat capacity at constant volume (e.g. a heat capacity per unit volume), 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).

We use the substitution:

$$
\begin{aligned}
\mathbf{U} &= \rho\mathbf{u} \\
p &= \frac{R}{c_v}(E - \tfrac{1}{2}\rho u^2) = \frac{R}{c_v}\left(E - \frac{\mathbf{U}^2}{2\rho}\right)
\end{aligned}
$$

and we get:

$$
\begin{aligned}
\frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{U} &= 0 \\
\frac{\partial\mathbf{U}}{\partial t} + \nabla\cdot\left(\frac{\mathbf{U}\mathbf{U}^T}{\rho} + p\mathbb{1}\right) - \mathbf{f} &= 0 \\
\frac{\partial E}{\partial t} + \nabla\cdot\left(\frac{\mathbf{U}}{\rho}(E+p)\right) &= 0
\end{aligned}
$$

Now we write $\mathbf{U} = (U, V, W)$ and we get:

$$
\frac{\partial}{\partial t}\begin{pmatrix} \varrho \\ U \\ V \\ W \\ E \end{pmatrix} + \frac{\partial}{\partial x}\begin{pmatrix} U \\ \frac{U^2}{\varrho} + p \\ \frac{UV}{\varrho} \\ \frac{UW}{\varrho} \\ \frac{U}{\varrho}(E+p) \end{pmatrix} + \frac{\partial}{\partial y}\begin{pmatrix} V \\ \frac{VU}{\varrho} \\ \frac{V^2}{\varrho} + p \\ \frac{VW}{\varrho} \\ \frac{V}{\varrho}(E+p) \end{pmatrix} + \frac{\partial}{\partial z}\begin{pmatrix} W \\ \frac{WU}{\varrho} \\ \frac{WV}{\varrho} \\ \frac{W^2}{\varrho} + p \\ \frac{W}{\varrho}(E+p) \end{pmatrix} + \begin{pmatrix} 0 \\ -f_x \\ -f_y \\ -f_z \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}
$$

$$
p = \frac{R}{c_v}\left(E - \frac{U^2 + V^2 + W^2}{2\rho}\right)
$$

We solve for the unknowns ρ, U, V, W and E 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.

After introducing:

$$\mathbf{w} = \begin{pmatrix} \varrho \\ U \\ V \\ W \\ E \end{pmatrix} = \begin{pmatrix} \varrho \\ \rho u_1 \\ \rho u_2 \\ \rho u_3 \\ E \end{pmatrix} = \begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix}$$

$$\mathbf{f}_x = \begin{pmatrix} U \\ \frac{U^2}{\varrho} + p \\ \frac{UV}{\varrho} \\ \frac{UW}{\varrho} \\ \frac{U}{\varrho}(E+p) \end{pmatrix} = \begin{pmatrix} w_1 \\ \frac{w_1^2}{w_0} + p \\ \frac{w_1 w_2}{w_0} \\ \frac{w_1 w_3}{w_0} \\ \frac{w_1}{w_0}(w_4+p) \end{pmatrix}$$

$$\mathbf{f}_y = \begin{pmatrix} V \\ \frac{VU}{\varrho} \\ \frac{V^2}{\varrho} + p \\ \frac{VW}{\varrho} \\ \frac{V}{\varrho}(E+p) \end{pmatrix} = \begin{pmatrix} w_2 \\ \frac{w_2 w_1}{w_0} \\ \frac{w_2^2}{w_0} + p \\ \frac{w_2 w_3}{w_0} \\ \frac{w_2}{w_0}(w_4+p) \end{pmatrix}$$

$$\mathbf{f}_z = \begin{pmatrix} W \\ \frac{WU}{\varrho} \\ \frac{WV}{\varrho} \\ \frac{W^2}{\varrho} + p \\ \frac{W}{\varrho}(E+p) \end{pmatrix} = \begin{pmatrix} w_3 \\ \frac{w_3 w_1}{w_0} \\ \frac{w_3 w_2}{w_0} \\ \frac{w_3^2}{w_0} + p \\ \frac{w_3}{w_0}(w_4+p) \end{pmatrix}$$

$$\mathbf{g} = \begin{pmatrix} 0 \\ -f_x \\ -f_y \\ -f_z \\ 0 \end{pmatrix}$$

$$p = \frac{R}{c_v}\left(E - \frac{U^2+V^2+W^2}{2\rho}\right) = \frac{R}{c_v}\left(w_4 - \frac{w_1^2+w_2^2+w_3^2}{2w_0}\right)$$

we can then write the equations as:

$$\frac{\partial \mathbf{w}}{\partial t} + \frac{\partial \mathbf{f}_x}{\partial x} + \frac{\partial \mathbf{f}_y}{\partial y} + \frac{\partial \mathbf{f}_z}{\partial z} + \mathbf{g} = 0$$

Note: $\mathbf{U} \equiv \mathbf{j}$, 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.

Now we write the spatial derivatives using 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):

$$\mathbf{A}_x(\mathbf{w}) = \frac{\partial \mathbf{f}_x}{\partial \mathbf{w}} = \left(\begin{matrix} 0 & 1 & 0 \\ -\frac{w_1^2}{w_0^2} + \frac{R}{c_v}\frac{w_1^2+w_2^2+w_3^2}{2w_0^2} & \frac{2w_1}{w_0} - \frac{R}{c_v}\frac{w_1}{w_0} & -\frac{R}{c_v}\frac{w_2}{w_0} \\ -\frac{w_1 w_2}{w_0^2} & \frac{w_2}{w_0} & \frac{w_1}{w_0} \\ -\frac{w_1 w_3}{w_0^2} & \frac{w_3}{w_0} & 0 \\ -\frac{w_1 w_4}{w_0^2} - \frac{w_1}{w_0^2}\frac{R}{c_v}\left(w_4 - \frac{w_1^2+w_2^2+w_3^2}{2w_0}\right) + \frac{w_1}{w_0}\frac{R}{c_v}\frac{w_1^2+w_2^2+w_3^2}{2w_0^2} & \frac{w_4}{w_0} + \frac{1}{w_0}\frac{R}{c_v}\left(w_4 - \frac{w_1^2+w_2^2+w_3^2}{2w_0}\right) - \frac{R}{c_v}\frac{w_1^2}{w_0^2} & -\frac{R}{c_v}\frac{w_1 w_2}{w_0^2} \end{matrix} \right.$$

$$\mathbf{A}_y(\mathbf{w}) = \frac{\partial \mathbf{f}_y}{\partial \mathbf{w}} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ -\frac{w_2 w_1}{w_0^2} & \frac{w_2}{w_0} & \frac{w_1}{w_0} & 0 & 0 \\ -\frac{w_2^2}{w_0^2} + \frac{R}{c_v}\frac{w_1^2+w_2^2+w_3^2}{2w_0^2} & -\frac{R}{c_v}\frac{w_1}{w_0} & \frac{2w_2}{w_0} - \frac{R}{c_v}\frac{w_2}{w_0} & -\frac{R}{c_v}\frac{w_3}{w_0} & \frac{R}{c_v} \\ -\frac{w_2 w_3}{w_0^2} & 0 & \frac{w_3}{w_0} & \frac{w_2}{w_0} & 0 \\ . & . & . & . & . \end{pmatrix}$$

$$\mathbf{A}_z(\mathbf{w}) = \frac{\partial \mathbf{f}_z}{\partial \mathbf{w}} = \left(\begin{matrix} 0 & 0 & 0 & 1 \\ -\frac{w_3 w_1}{w_0^2} & \frac{w_3}{w_0} & 0 & \frac{w_1}{w_0} \\ -\frac{w_3 w_2}{w_0^2} & 0 & \frac{w_3}{w_0} & \frac{w_2}{w_0} \\ -\frac{w_3^2}{w_0^2} + \frac{R}{c_v}\frac{w_1^2+w_2^2+w_3^2}{2w_0^2} & -\frac{R}{c_v}\frac{w_1}{w_0} & -\frac{R}{c_v}\frac{w_2}{w_0} & \frac{2w_3}{w_0} - \frac{R}{c_v}\frac{w_3}{w_0} \\ -\frac{w_3 w_4}{w_0^2} - \frac{w_3}{w_0^2}\frac{R}{c_v}\left(w_4 - \frac{w_1^2+w_2^2+w_3^2}{2w_0}\right) + \frac{w_3}{w_0}\frac{R}{c_v}\frac{w_1^2+w_2^2+w_3^2}{2w_0^2} & -\frac{R}{c_v}\frac{w_3 w_1}{w_0^2} & -\frac{R}{c_v}\frac{w_3 w_2}{w_0^2} & \frac{w_4}{w_0} + \frac{1}{w_0}\frac{R}{c_v}\left(w_4 - \frac{w_1^2+w_2^2+w_3^2}{2w_0}\right. \end{matrix} \right.$$

3.3.1 FEM formulation

The Euler equations:

$$\frac{\partial \mathbf{w}}{\partial t} + \mathbf{A}_x(\mathbf{w}) \frac{\partial \mathbf{w}}{\partial x} + \mathbf{A}_y(\mathbf{w}) \frac{\partial \mathbf{w}}{\partial y} + \mathbf{A}_z(\mathbf{w}) \frac{\partial \mathbf{w}}{\partial z} + \mathbf{g} = 0$$

are nonlinear. The simplest approximation is to linearize them by:

$$\frac{\mathbf{w}^{n+1}-\mathbf{w}^n}{\tau}+\mathbf{A}_x(\mathbf{w}^n)\frac{\partial\mathbf{w}^{n+1}}{\partial x}+\mathbf{A}_y(\mathbf{w}^n)\frac{\partial\mathbf{w}^{n+1}}{\partial y}+\mathbf{A}_z(\mathbf{w}^n)\frac{\partial\mathbf{w}^{n+1}}{\partial z}+\mathbf{g}=0$$

Then we multiply by the test functions (one by one):

$$\begin{pmatrix}\varphi^0\\0\\0\\0\\0\end{pmatrix},\begin{pmatrix}0\\\varphi^1\\0\\0\\0\end{pmatrix},\begin{pmatrix}0\\0\\\varphi^2\\0\\0\end{pmatrix},\begin{pmatrix}0\\0\\0\\\varphi^3\\0\end{pmatrix},\begin{pmatrix}0\\0\\0\\0\\\varphi^4\end{pmatrix}$$

and integrate over the 3D domain Ω, so we get (here the index $i=0,1,2,3,4$ is numbering the 5 equations, so we are *not* summing over it):

$$\int_\Omega\frac{w_i^{n+1}-w_i^n}{\tau}\varphi^i+(\mathbf{A}_x(\mathbf{w}^n))_{ij}\frac{\partial w_j^{n+1}}{\partial x}\varphi^i+(\mathbf{A}_y(\mathbf{w}^n))_{ij}\frac{\partial w_j^{n+1}}{\partial y}\varphi^i+(\mathbf{A}_z(\mathbf{w}^n))_{ij}\frac{\partial w_j^{n+1}}{\partial z}\varphi^i+g_i\varphi^i\,\mathrm{d}^3x=0$$

Now we integrate by parts and use the homogeneity property ($w_j\frac{\partial(\mathbf{A}_z(\mathbf{w}^n))_{ij}}{\partial x}\varphi^i=0$):

$$\int_\Omega\frac{w_i^{n+1}-w_i^n}{\tau}\varphi^i-(\mathbf{A}_x(\mathbf{w}^n))_{ij}\,w_j^{n+1}\frac{\partial\varphi^i}{\partial x}-(\mathbf{A}_y(\mathbf{w}^n))_{ij}\,w_j^{n+1}\frac{\partial\varphi^i}{\partial y}-(\mathbf{A}_z(\mathbf{w}^n))_{ij}\,w_j^{n+1}\frac{\partial\varphi^i}{\partial z}+g_i\varphi^i\,\mathrm{d}^3x+$$
$$+\int_{\partial\Omega}(\mathbf{A}_x(\mathbf{w}^n))_{ij}\,w_j^{n+1}\varphi^i\,n_x+(\mathbf{A}_y(\mathbf{w}^n))_{ij}\,w_j^{n+1}\varphi^i\,n_y+(\mathbf{A}_z(\mathbf{w}^n))_{ij}\,w_j^{n+1}\varphi^i\,n_z\,\mathrm{d}^2x=0$$

where $\mathbf{n}=(n_x,n_y,n_z)$ is the outward surface normal to $\partial\Omega$. Rearranging:

$$\int_\Omega\frac{w_i^{n+1}}{\tau}\varphi^i-(\mathbf{A}_x(\mathbf{w}^n))_{ij}\,w_j^{n+1}\frac{\partial\varphi^i}{\partial x}-(\mathbf{A}_y(\mathbf{w}^n))_{ij}\,w_j^{n+1}\frac{\partial\varphi^i}{\partial y}-(\mathbf{A}_z(\mathbf{w}^n))_{ij}\,w_j^{n+1}\frac{\partial\varphi^i}{\partial z}\,\mathrm{d}^3x+$$
$$+\int_{\partial\Omega}(\mathbf{A}_x(\mathbf{w}^n))_{ij}\,w_j^{n+1}\varphi^i\,n_x+(\mathbf{A}_y(\mathbf{w}^n))_{ij}\,w_j^{n+1}\varphi^i\,n_y+(\mathbf{A}_z(\mathbf{w}^n))_{ij}\,w_j^{n+1}\varphi^i\,n_z\,\mathrm{d}^2x=\int_\Omega\frac{w_i^n}{\tau}\varphi^i-g_i\varphi^i\,\mathrm{d}^3x$$

3.3.2 Sea Breeze Modeling

In our model we make the following assumptions:

$$\begin{aligned}f_x&=0\\f_y&=0\\f_z&=-\rho g=-w_0g\\V&=0\\\frac{\partial U}{\partial y}=\frac{\partial V}{\partial y}&=\frac{\partial W}{\partial y}=\frac{\partial E}{\partial y}=0\end{aligned}$$

so we get a 2D model:

$$\frac{\partial}{\partial t}\begin{pmatrix}\varrho\\U\\0\\W\\E\end{pmatrix}+\frac{\partial}{\partial x}\begin{pmatrix}U\\\frac{U^2}{\varrho}+p\\0\\\frac{UW}{\varrho}\\\frac{U}{\varrho}(E+p)\end{pmatrix}+\frac{\partial}{\partial z}\begin{pmatrix}W\\\frac{WU}{\varrho}\\0\\\frac{W^2}{\varrho}+p\\\frac{W}{\varrho}(E+p)\end{pmatrix}+\begin{pmatrix}0\\0\\0\\\rho g\\0\end{pmatrix}=\begin{pmatrix}0\\0\\0\\0\\0\end{pmatrix}$$
$$p=\frac{R}{c_v}\left(E-\frac{U^2+W^2}{2\rho}\right)$$

where we prescribe R, c_v, g and solve for ρ, U, W and E as functions of (t, x, z). We delete the row for y, which only contains zeros anyway and introduce:

$$\mathbf{w} = \begin{pmatrix} \varrho \\ \rho u_1 \\ \rho u_3 \\ E \end{pmatrix} = \begin{pmatrix} w_0 \\ w_1 \\ w_3 \\ w_4 \end{pmatrix}$$

$$\mathbf{A}_x(\mathbf{w}) = \frac{\partial \mathbf{f}_x}{\partial \mathbf{w}} = \left(\begin{matrix} 0 & 1 & 0 & 0 \\ -\frac{w_1^2}{w_0^2} + \frac{R}{c_v}\frac{w_1^2+w_3^2}{2w_0^2} & \frac{2w_1}{w_0} - \frac{R}{c_v}\frac{w_1}{w_0} & -\frac{R}{c_v}\frac{w_3}{w_0} & \frac{R}{c_v} \\ -\frac{w_1 w_3}{w_0^2} & \frac{w_3}{w_0} & \frac{w_1}{w_0} & 0 \\ -\frac{w_1 w_4}{w_0^2} - \frac{w_1}{w_0^2}\frac{R}{c_v}\left(w_4 - \frac{w_1^2+w_3^2}{2w_0}\right) + \frac{w_1}{w_0}\frac{R}{c_v}\frac{w_1^2+w_3^2}{2w_0^2} & \frac{w_4}{w_0} + \frac{1}{w_0}\frac{R}{c_v}\left(w_4 - \frac{w_1^2+w_3^2}{2w_0}\right) - \frac{R}{c_v}\frac{w_1^2}{w_0^2} & -\frac{R}{c_v}\frac{w_1 w_3}{w_0^2} & \frac{w_1}{w_0} + \frac{R}{c_v}\frac{w_1}{w_0} \end{matrix} \right.$$

$$\mathbf{A}_z(\mathbf{w}) = \frac{\partial \mathbf{f}_z}{\partial \mathbf{w}} = \left(\begin{matrix} 0 & 0 & 1 & 0 \\ -\frac{w_3 w_1}{w_0^2} & \frac{w_3}{w_0} & \frac{w_1}{w_0} & 0 \\ -\frac{w_3^2}{w_0^2} + \frac{R}{c_v}\frac{w_1^2+w_3^2}{2w_0^2} & -\frac{R}{c_v}\frac{w_1}{w_0} & \frac{2w_3}{w_0} - \frac{R}{c_v}\frac{w_3}{w_0} & \frac{R}{c_v} \\ -\frac{w_3 w_4}{w_0^2} - \frac{w_3}{w_0^2}\frac{R}{c_v}\left(w_4 - \frac{w_1^2+w_3^2}{2w_0}\right) + \frac{w_3}{w_0}\frac{R}{c_v}\frac{w_1^2+w_3^2}{2w_0^2} & -\frac{R}{c_v}\frac{w_3 w_1}{w_0^2} & \frac{w_4}{w_0} + \frac{1}{w_0}\frac{R}{c_v}\left(w_4 - \frac{w_1^2+w_3^2}{2w_0}\right) - \frac{R}{c_v}\frac{w_3^2}{w_0^2} & \frac{w_3}{w_0} + \frac{R}{c_v}\frac{w_3}{w_0} \end{matrix} \right.$$

The weak formulation in 2D is (here $i = 0, 1, 3, 4$):

$$\int_\Omega \frac{w_i^{n+1}}{\tau}\varphi^i - (\mathbf{A}_x(\mathbf{w}^n))_{ij}\, w_j^{n+1}\frac{\partial \varphi^i}{\partial x} - (\mathbf{A}_z(\mathbf{w}^n))_{ij}\, w_j^{n+1}\frac{\partial \varphi^i}{\partial z}\,\mathrm{d}^2 x +$$
$$+ \int_{\partial\Omega} (\mathbf{A}_x(\mathbf{w}^n))_{ij}\, w_j^{n+1}\varphi^i\, n_x + (\mathbf{A}_z(\mathbf{w}^n))_{ij}\, w_j^{n+1}\varphi^i\, n_z\,\mathrm{d}x = \int_\Omega \frac{w_i^n}{\tau}\varphi^i - g_i\varphi^i\,\mathrm{d}^2 x$$

In the boundary (line) integral we prescribe w_4^{n+1} using a Dirichlet condition and calculate it at each iteration using:

$$w_4^{n+1} = E = \rho T c_v + \tfrac{1}{2}\rho u^2 = w_0 T c_v + \frac{w_1^2 + w_3^2}{2w_0}$$

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 $w_3^{n+1} = 0$ at the top and bottom.

In order to specify the input forms for Hermes, we'll write the weak formulation as:

$$\begin{aligned}
B_{00}(w_0,\varphi^0) + B_{01}(w_1,\varphi^0) + B_{03}(w_3,\varphi^0) + B_{04}(w_4,\varphi^0) &= l_0(\varphi^0) \\
B_{10}(w_0,\varphi^1) + B_{11}(w_1,\varphi^1) + B_{13}(w_3,\varphi^1) + B_{14}(w_4,\varphi^1) &= l_1(\varphi^1) \\
B_{30}(w_0,\varphi^3) + B_{31}(w_1,\varphi^3) + B_{33}(w_3,\varphi^3) + B_{34}(w_4,\varphi^3) &= l_3(\varphi^3) \\
B_{40}(w_0,\varphi^4) + B_{41}(w_1,\varphi^4) + B_{43}(w_3,\varphi^4) + B_{44}(w_4,\varphi^4) &= l_4(\varphi^4)
\end{aligned}$$

where the forms are (we write w_i instead of w_i^{n+1}):

$$\begin{aligned}
l_0(\varphi^0) &= \int_\Omega \frac{w_0^n\varphi^0}{\tau}\,\mathrm{d}^2 x \\
l_1(\varphi^1) &= \int_\Omega \frac{w_1^n\varphi^1}{\tau}\,\mathrm{d}^2 x \\
l_3(\varphi^3) &= \int_\Omega \frac{w_3^n\varphi^3}{\tau} + \rho g\varphi^3\,\mathrm{d}^2 x \\
l_4(\varphi^4) &= \int_\Omega \frac{w_4^n\varphi^4}{\tau}\,\mathrm{d}^2 x \\
B_{ij}(w_j,\varphi^i) &= \int_\Omega \frac{w_i}{\tau}\varphi^i\delta_{ij} - (\mathbf{A}_x(\mathbf{w}^n))_{ij}\, w_j\frac{\partial\varphi^i}{\partial x} - (\mathbf{A}_z(\mathbf{w}^n))_{ij}\, w_j\frac{\partial\varphi^i}{\partial z}\,\mathrm{d}^2 x
\end{aligned}$$

In the last expression we do *not* sum over i nor j. In particular:

$$B_{00}(w_0,\varphi^0) = \int_\Omega \frac{w_0}{\tau}\varphi^0 - (\mathbf{A}_x(\mathbf{w}^n))_{00}\, w_0 \frac{\partial\varphi^0}{\partial x} - (\mathbf{A}_z(\mathbf{w}^n))_{00}\, w_0 \frac{\partial\varphi^0}{\partial z}\, \mathrm{d}^2x = \int_\Omega \frac{w_0}{\tau}\varphi^0\, \mathrm{d}^2x$$
$$B_{01}(w_1,\varphi^0) = \int_\Omega - (\mathbf{A}_x(\mathbf{w}^n))_{01}\, w_1 \frac{\partial\varphi^0}{\partial x} - (\mathbf{A}_z(\mathbf{w}^n))_{01}\, w_1 \frac{\partial\varphi^0}{\partial z}\, \mathrm{d}^2x = \int_\Omega - (\mathbf{A}_x(\mathbf{w}^n))_{01}\, w_1 \frac{\partial\varphi^0}{\partial x}\, \mathrm{d}^2x$$
$$B_{03}(w_3,\varphi^0) = \int_\Omega - (\mathbf{A}_x(\mathbf{w}^n))_{03}\, w_3 \frac{\partial\varphi^0}{\partial x} - (\mathbf{A}_z(\mathbf{w}^n))_{03}\, w_3 \frac{\partial\varphi^0}{\partial z}\, \mathrm{d}^2x = \int_\Omega - (\mathbf{A}_z(\mathbf{w}^n))_{03}\, w_3 \frac{\partial\varphi^0}{\partial z}\, \mathrm{d}^2x$$
$$B_{04}(w_4,\varphi^0) = \int_\Omega - (\mathbf{A}_x(\mathbf{w}^n))_{04}\, w_4 \frac{\partial\varphi^0}{\partial x} - (\mathbf{A}_z(\mathbf{w}^n))_{04}\, w_4 \frac{\partial\varphi^0}{\partial z}\, \mathrm{d}^2x = 0$$
$$B_{10}(w_0,\varphi^1) = \int_\Omega - (\mathbf{A}_x(\mathbf{w}^n))_{10}\, w_0 \frac{\partial\varphi^1}{\partial x} - (\mathbf{A}_z(\mathbf{w}^n))_{10}\, w_0 \frac{\partial\varphi^1}{\partial z}\, \mathrm{d}^2x$$
$$B_{11}(w_1,\varphi^1) = \int_\Omega \frac{w_1}{\tau}\varphi^1 - (\mathbf{A}_x(\mathbf{w}^n))_{11}\, w_1 \frac{\partial\varphi^1}{\partial x} - (\mathbf{A}_z(\mathbf{w}^n))_{11}\, w_1 \frac{\partial\varphi^1}{\partial z}\, \mathrm{d}^2x$$
$$\dots$$

3.3.3 Older notes

Author: Pavel Solin

Governing Equations and Boundary Conditions

$$\frac{\partial}{\partial t}\begin{pmatrix} \varrho \\ U \\ W \\ \theta \end{pmatrix} + \frac{\partial}{\partial x}\begin{pmatrix} U \\ \frac{U^2}{\varrho} + R\theta \\ \frac{UW}{\varrho} \\ \frac{\theta U}{\varrho} \end{pmatrix} + \frac{\partial}{\partial z}\begin{pmatrix} W \\ \frac{UW}{\varrho} \\ \frac{W^2}{\varrho} + R\theta \\ \frac{\theta W}{\varrho} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \varrho g \\ \frac{R\theta}{c_v}\mathrm{div}\mathbf{v} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{3.19}$$

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 (3.19) 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

$$F_i^\varrho(Y^{n+1}) = \int_\Omega \frac{\varrho^{n+1}}{\tau}\varphi_i^\varrho - \int_\Omega \frac{\varrho^n}{\tau}\varphi_i^\varrho + \int_\Omega \frac{\partial U^{n+1}}{\partial x}\varphi_i^\varrho + \int_\Omega \frac{\partial W^{n+1}}{\partial z}\varphi_i^\varrho = 0 \tag{3.20}$$

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

$$\varrho^{n+1} = \sum_{k=1}^{N^\varrho} y_k^\varrho \varphi_k^\varrho, \quad U^{n+1} = \sum_{k=1}^{N^U} y_k^U \varphi_k^U, \quad W^{n+1} = \sum_{k=1}^{N^W} y_k^W \varphi_k^W, \quad \theta^{n+1} = \sum_{k=1}^{N^\theta} y_k^\theta \varphi_k^\theta. \tag{3.21}$$

It follows from (3.20) and (3.21) 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 (3.19) 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 (3.19) 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 (3.19) 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

FOUR

QUANTUM FIELD THEORY AND QUANTUM MECHANICS

4.1 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.

4.2 Standard Model

4.2.1 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:

$$\partial^\mu \to D^\mu = \partial^\mu - \frac{i}{2} g \tau_k A^\mu_k - \frac{i}{2} g' Y B^\mu$$

and second adding the kinetic terms

$$-\frac{1}{4} F^a_{\mu\nu} F^{a\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}$$

of the vector gauge particles to the lagrangian.

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g\epsilon^{abc} A^b_\mu A^c_\nu$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$$

$$\Phi = e^{\frac{i}{v}\pi^a(x)\tau^a} \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(v + H(x)) \end{pmatrix}$$

This breaks the gauge invariance. The $\partial^\mu \pi^a$ are going to be added to A^a_μ 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):

$$\mathcal{L}_{Higgs} = \frac{1}{2} \Phi^+ (\overleftarrow{\partial}_\mu + igA^a_\mu \frac{\tau^a}{2} + ig'YB_\mu)(\overrightarrow{\partial}^\mu + igA^{a\mu}\frac{\tau^a}{2} + ig'YB^\mu)\Phi - \lambda(\Phi^+\Phi - \frac{v^2}{2})^2 =$$

$$= \Phi^+_U (\overleftarrow{\partial}_\mu + igA^a_\mu \frac{\tau^a}{2} + ig'YB_\mu)(\overrightarrow{\partial}^\mu + igA^{a\mu}\frac{\tau^a}{2} + ig'YB\mu)\Phi_U - \lambda(\Phi^+_U\Phi_U - \frac{v^2}{2})^2 =$$

$$= \frac{1}{2}\partial_\mu H \partial^\mu H - \lambda v^2 H^2 - \lambda v H^3 - \frac{1}{4}\lambda H^4 +$$

$$+ \frac{1}{8}(v+H)^2 \left(2g^2 \frac{A^1_\mu + iA^2_\mu}{\sqrt{2}} \frac{A^{1\mu} - iA^{2\mu}}{\sqrt{2}} + (g^2 + 4Y^2 g'^2) \frac{gA^3_\mu - 2Yg'B_\mu}{\sqrt{g^2 + 4Y^2 g'^2}} \frac{gA^{3\mu} - 2Yg'B^\mu}{\sqrt{g^2 + 4Y^2 g'^2}} \right) =$$

$$= \frac{1}{2}\partial_\mu H \partial^\mu H - \lambda v^2 H^2 - \lambda v H^3 - \frac{1}{4}\lambda H^4 + \frac{1}{8}(v+H)^2 \left(2g^2 W^-_\mu W^{+\mu} + \frac{g^2}{\cos^2\theta_W} Z_\mu Z^\mu \right) =$$

$$= \frac{1}{2}\partial_\mu H \partial^\mu H - \lambda v^2 H^2 + \frac{1}{4} g^2 v^2 W^-_\mu W^{+\mu} + \frac{g^2 v^2}{8\cos^2\theta_W} Z_\mu Z^\mu - \lambda v H^3 - \frac{1}{4}\lambda H^4 +$$

$$+\frac{1}{2}vg^2W^-_\mu W^{+\mu}H+\frac{g^2}{4\cos\theta_W}vZ_\mu Z^\mu H+\frac{1}{4}g^2W^-_\mu W^{+\mu}H^2+\frac{g^2}{8\cos\theta_W}Z_\mu Z^\mu H^2$$

Where we put

$$W^\pm_\mu=\frac{1}{\sqrt{2}}(A^1_\mu\mp iA^2_\mu)$$

$$Z_\mu=\frac{g}{\sqrt{g^2+4Y^2g'^2}}A^3_\mu-\frac{2Yg'}{\sqrt{g^2+4Y^2g'^2}}B_\mu$$

we defined θ_W by the relation

$$\cos\theta_W=\frac{g}{\sqrt{g^2+4Y^2g'^2}}$$

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

$$\mathcal{L}_{Yukawa}=-h_e\bar L\Phi e_R-\text{h.c.}=-h_e\bar L\Phi_U e_R-\text{h.c.}=$$

$$=-\frac{1}{\sqrt{2}}h_e(v+H)(\bar e_Le_R+\bar e_Re_L)=-\frac{1}{\sqrt{2}}h_e(v+H)\bar ee=$$

$$=-\frac{1}{\sqrt{2}}h_ev\bar ee-\frac{1}{\sqrt{2}}h_e\bar eeH$$

The term $\bar L\Phi e_R$ is $U(1)$ (hypercharge) invariant, so

$$-Y_L+Y+Y_R=0$$

Leptonic Terms

$$\mathcal{L}=i\bar L\gamma^\mu\partial_\mu L+i\bar e_R\gamma^\mu\partial_\mu e_R\to$$

$$\to i\bar L\gamma^\mu(\partial_\mu-igA^a_\mu\frac{\tau^a}{2}-ig'Y_LB_\mu)L+i\bar e_R\gamma^\mu(\partial_\mu-ig'Y_RB_\mu)e_R=$$

$$= i\bar L\gamma^\mu\partial_\mu L + i\bar e_R\gamma^\mu\partial_\mu e_R + g\bar L\gamma^\mu\frac{\tau^a}{2}LA^a_\mu + g'Y_L\bar L\gamma^\mu LB_\mu + g'Y_R\bar e_R\gamma^\mu e_RB_\mu =$$

$$= i\bar L\gamma^\mu\partial_\mu L + i\bar e_R\gamma^\mu\partial_\mu e_R + \frac{g}{\sqrt2}(\bar\nu_L\gamma^\mu e_LW^+_\mu + \text{h.c.}) + \frac12 g\bar L\gamma^\mu\tau^3LA^3_\mu + g'Y_L\bar L\gamma^\mu LB_\mu + g'Y_R\bar e_R\gamma^\mu e_RB_\mu =$$

$$= i\bar\nu_L\gamma^\mu\partial_\mu\nu_L + i\bar e\gamma^\mu\partial_\mu e + \frac{g}{\sqrt2}(\bar\nu_L\gamma^\mu e_LW^+_\mu + \text{h.c.}) + \frac12 g\bar\nu_L\gamma^\mu\nu_LA^3_\mu - \frac12 g\bar e_L\gamma^\mu e_LA^3_\mu$$

$$+g'Y_L\bar\nu_L\gamma^\mu\nu_LB_\mu + g'Y_L\bar e_L\gamma^\mu e_LB_\mu + g'Y_R\bar e_R\gamma^\mu e_RB_\mu =$$

$$= i\bar\nu_L\gamma^\mu\partial_\mu\nu_L + i\bar e\gamma^\mu\partial_\mu e + \frac{g}{\sqrt2}(\bar\nu_L\gamma^\mu e_LW^+_\mu + \text{h.c.})$$

$$+\left[(\tfrac12 g\sin\theta_W + Y_Lg'\cos\theta_W)\bar\nu_L\gamma^\mu\nu_L + (-\tfrac12 g\sin\theta_W + Y_Lg'\cos\theta_W)\bar e_L\gamma^\mu e_L + Y_Rg'\cos\theta_W\bar e_R\gamma^\mu e_R\right]A_\mu$$

$$+\left[(\tfrac12 g\cos\theta_W - Y_Lg'\sin\theta_W)\bar\nu_L\gamma^\mu\nu_L + (-\tfrac12 g\cos\theta_W - Y_Lg'\sin\theta_W)\bar e_L\gamma^\mu e_L - 2Y_Lg'\sin\theta_W\bar e_R\gamma^\mu e_R\right]Z_\mu$$

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 $\frac12 g\sin\theta_W + Y_Lg'\cos\theta_W = 0$ 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:

$$\tfrac12 g\sin\theta_W + Y_Lg'\cos\theta_W = 0$$

$$-\tfrac12 g\sin\theta_W + Y_Lg'\cos\theta_W = -g\sin\theta_W$$

$$Y_Rg'\cos\theta_W = -g\sin\theta_W$$

$$\tan\theta_W = -2Y_L\frac{g'}{g}$$

and the Lagrangian can be further simplified:

$$\mathcal{L} = i\bar\nu_L\gamma^\mu\partial_\mu\nu_L + i\bar e\gamma^\mu\partial_\mu e + \frac{g}{\sqrt2}(\bar\nu_L\gamma^\mu e_LW^+_\mu + \text{h.c.})$$

$$-g\sin\theta_W(\bar e_L\gamma^\mu e_L + \bar e_R\gamma^\mu e_R)A_\mu$$

$$+\frac{g}{\cos\theta_W}\left[\tfrac12\bar\nu_L\gamma^\mu\nu_L + (-\tfrac12+\sin^2\theta_W)\bar e_L\gamma^\mu e_L + \sin^2\theta_W\bar e_R\gamma^\mu e_R\right]Z_\mu =$$

$$= i\bar\nu_L\gamma^\mu\partial_\mu\nu_L + i\bar e\gamma^\mu\partial_\mu e + \frac{g}{2\sqrt2}(\bar\nu\gamma^\mu(1-\gamma_5)eW^+_\mu + \text{h.c.}) - g\sin\theta_W\bar e\gamma^\mu eA_\mu$$

$$+\frac{g}{2\cos\theta_W}\left[\bar\nu\gamma^\mu(1-\gamma_5)\nu + \bar e\gamma^\mu(-\tfrac12 + 2\sin^2\theta_W + \tfrac12\gamma_5)e\right]Z_\mu$$

Where we used the relations $\bar\nu_L\gamma^\mu e_L = \frac12\bar\nu\gamma^\mu(1-\gamma_5)e$ and $\bar\nu_R\gamma^\mu e_R = \frac12\bar\nu\gamma^\mu(1+\gamma_5)e$.

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:

$$\mathcal{L} = \frac{1}{2}\partial_\mu H \partial^\mu H - \lambda v^2 H^2 + \frac{1}{4} g^2 v^2 W^-_\mu W^{+\mu} + \frac{g^2 v^2}{8\cos^2\theta_W} Z_\mu Z^\mu - \lambda v H^3 - \frac{1}{4}\lambda H^4 +$$

$$+\frac{1}{2} v g^2 W^-_\mu W^{+\mu} H + \frac{g^2}{4\cos\theta_W} v Z_\mu Z^\mu H + \frac{1}{4} g^2 W^-_\mu W^{+\mu} H^2 + \frac{g^2}{8\cos\theta_W} Z_\mu Z^\mu H^2$$

$$-\frac{1}{\sqrt{2}} h_e v \bar{e} e - \frac{1}{\sqrt{2}} h_e \bar{e} e H$$

$$-\frac{1}{2} W^-_{\mu\nu} W^{+\mu\nu} - \frac{1}{4} A_{\mu\nu} A^{\mu\nu} - \frac{1}{4} Z_{\mu\nu} Z^{\mu\nu} + \mathcal{L}_{WW\gamma} + L_{WWZ} + L_{WW\gamma\gamma} + L_{WWWW} + L_{WWZZ} + L_{WWZ\gamma}$$

$$+i\bar{\nu}_L \gamma^\mu \partial_\mu \nu_L + i\bar{e}\gamma^\mu \partial_\mu e + \frac{g}{2\sqrt{2}} (\bar{\nu}\gamma^\mu (1-\gamma_5) e W^+_\mu + \text{h.c.}) - g \sin\theta_W \bar{e}\gamma^\mu e A_\mu$$

$$+\frac{g}{2\cos\theta_W} \left[\bar{\nu}\gamma^\mu (1-\gamma_5)\nu + \bar{e}\gamma^\mu (-\tfrac{1}{2} + 2\sin^2\theta_W + \tfrac{1}{2}\gamma_5) e\right] Z_\mu$$

$$+(e, \nu_e, h_e \leftrightarrow \mu, \nu_\mu, h_\mu) + (e, \nu_e, h_e \leftrightarrow \tau, \nu_\tau, h_\tau)$$

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

$$\mathcal{L} = -\frac{1}{2} m_H^2 H^2 + m_W^2 W^-_\mu W^{+\mu} + \frac{1}{2} m_Z^2 Z_\mu Z^\mu - m_e \bar{e} e + \cdots$$

comparing to the above:

$$\mathcal{L} = -\lambda v^2 H^2 + \frac{1}{4} g^2 v^2 W^-_\mu W^{+\mu} + \frac{g^2 v^2}{8\cos^2\theta_W} Z_\mu Z^\mu - \frac{1}{\sqrt{2}} h_e v \bar{e} e + \cdots$$

we get

$$m_W = \tfrac{1}{2} g v$$

$$m_Z = \frac{g v}{2\cos\theta_W} = \frac{m_W}{\cos\theta_W}$$

$$m_H = v\sqrt{2\lambda}$$

$$m_e = \frac{1}{\sqrt{2}} h_e v$$

Quarks

$$\mathcal{L}_{fermion} {+}{=} \sum_{q=d,s,b} i\bar{L}_0^{(q)}\gamma^\mu\partial_\mu L_0^{(q)} + \sum_{q=d,u,s,c,b,t} i\bar{q}_{0R}\gamma^\mu\partial_\mu q_{0R}$$

$$\mathcal{L}_{Yukawa} {+}{=} -\sum_{\substack{q=d,s,b\\ q'=d,s,b}} h_{qq'} i\bar{L}_0^{(q)}\Phi q'_{0R} + \text{h.c.} - \sum_{\substack{q=d,s,b\\ q'=u,c,t}} \tilde{h}_{qq'} i\bar{L}_0^{(q)}\tilde{\Phi} q'_{0R} + \text{h.c.}$$

4.2.2 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:

$$\phi_I(x) = \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_\mathbf{p}}} \left(a_\mathbf{p} e^{-ip\cdot x} + a_\mathbf{p}^\dagger e^{ip\cdot x}\right)$$

$$\pi_I(x) = \partial_t \phi_I(x) = \int \frac{\mathrm{d}^3 p}{(2\pi)^3} (-i)\sqrt{\frac{E_\mathbf{p}}{2}} \left(a_\mathbf{p} e^{-ip\cdot x} - a_\mathbf{p}^\dagger e^{ip\cdot x}\right)$$

where:

$$\left[a_\mathbf{p}, a_\mathbf{q}^\dagger\right] = (2\pi)^3 \delta^{(3)}(\mathbf{p}-\mathbf{q})$$

(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

$$H = \int \frac{\mathrm{d}^3 p}{(2\pi)^3} E_\mathbf{p} a_\mathbf{p}^\dagger a_\mathbf{p}$$

Fermions:

$$\psi_I(x) = \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_\mathbf{p}}} \sum_{s=1}^{2} \left(b_\mathbf{p}^s u^s(\mathbf{p}) e^{-ip\cdot x} + d_\mathbf{p}^{s\dagger} v^s(\mathbf{p}) e^{ip\cdot x}\right)$$

$$\bar{\psi}_I(x) = \psi_I^\dagger(x)\gamma^0 = \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_\mathbf{p}}} \sum_{s=1}^{2} \left(d_\mathbf{p}^s \bar{v}^s(\mathbf{p}) e^{-ip\cdot x} + b_\mathbf{p}^{s\dagger} \bar{u}^s(\mathbf{p}) e^{ip\cdot x}\right)$$

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:

$$\left\{\psi_a(\mathbf{x}), \psi^\dagger_b(\mathbf{y})\right\} = \delta^{(3)}(\mathbf{x}-\mathbf{y})\delta_{ab}$$
$$\left\{\psi_a(\mathbf{x}), \psi_b(\mathbf{y})\right\} = \left\{\psi^\dagger_a(\mathbf{x}), \psi^\dagger_b(\mathbf{y})\right\} = 0$$

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:

$$
\begin{aligned}
a^\dagger_{\mathbf{k}\text{ in}} &\to i\int \mathrm{d}^4x e^{ikx}\left(\partial^2+m^2\right)\phi(x) = \frac{k^2-m^2}{i}\tilde\phi(-k) = \frac{1}{\tilde D(k)}\tilde\phi(-k)\\
a_{\mathbf{k}\text{ out}} &\to i\int \mathrm{d}^4x e^{-ikx}\left(\partial^2+m^2\right)\phi(x) = \frac{k^2-m^2}{i}\tilde\phi(k) = \frac{1}{\tilde D(k)}\tilde\phi(k)\\
b^{s\dagger}_{\mathbf{k}\text{ in}} &\to i\int \mathrm{d}^4x \bar\psi(x)\left(i\overleftarrow{\not\partial}+m\right)u^s(\mathbf{k})e^{ikx} = \tilde{\bar\psi}(-k)\frac{-\not k-m}{i}u^s(\mathbf{k}) = \tilde{\bar\psi}(-k)\frac{1}{\tilde S(-k)}u^s(\mathbf{k})\\
b^{s}_{\mathbf{k}\text{ out}} &\to i\int \mathrm{d}^4x e^{-ikx}\bar u^s(\mathbf{k})\left(-i\not\partial+m\right)\psi(x) = \bar u^s(\mathbf{k})\frac{\not k-m}{i}\psi(k) = \bar u^s(\mathbf{k})\frac{1}{\tilde S(k)}\psi(k)\\
d^{s\dagger}_{\mathbf{k}\text{ in}} &\to -i\int \mathrm{d}^4x e^{ikx}\bar v^s(\mathbf{k})\left(-i\not\partial+m\right)\psi(x) = -\bar v^s(\mathbf{k})\frac{\not k-m}{i}\psi(-k) = -\bar v^s(\mathbf{k})\frac{1}{\tilde S(k)}\psi(-k)\\
d^{s}_{\mathbf{k}\text{ out}} &\to -i\int \mathrm{d}^4x \bar\psi(x)\left(i\overleftarrow{\not\partial}+m\right)v^s(\mathbf{k})e^{-ikx} = -\tilde{\bar\psi}(k)\frac{-\not k-m}{i}v^s(\mathbf{k}) = -\tilde{\bar\psi}(k)\frac{1}{\tilde S(-k)}v^s(\mathbf{k})\\
a^{r\dagger}_{\mathbf{k}\text{ in}} &\to i\epsilon^{r*}_\mu(\mathbf{k})\int \mathrm{d}^4x e^{ikx}\partial^2A^\mu(x) = \epsilon^{r*}_\mu(\mathbf{k})\frac{k^2}{i}\tilde A^\mu(-k)\\
a^{r}_{\mathbf{k}\text{ out}} &\to i\epsilon^{r}_\mu(\mathbf{k})\int \mathrm{d}^4x e^{-ikx}\partial^2A^\mu(x) = \epsilon^{r}_\mu(\mathbf{k})\frac{k^2}{i}\tilde A^\mu(k)
\end{aligned}
$$

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:

$$
\begin{aligned}
\partial^2 &\to -k^2\\
i\not\partial &\to \not k\\
e^{\pm ikx}\phi(x) &\to \tilde\phi(\mp k)\\
\frac{k^2-m^2}{i} &\to \frac{1}{\tilde D(k)}\\
\frac{\pm\not k-m}{i} &\to \frac{1}{\tilde S(\pm k)}
\end{aligned}
$$

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:

$$
\begin{aligned}
\langle f|i\rangle = \langle\Omega|b_{\mathbf{p}_{2'}}b_{\mathbf{p}_{1'}}b^\dagger_{\mathbf{p}_1}b^\dagger_{\mathbf{p}_2}|\Omega\rangle &= \langle\Omega|T\, b_{\mathbf{p}_{2'}}b_{\mathbf{p}_{1'}}b^\dagger_{\mathbf{p}_1}b^\dagger_{\mathbf{p}_2}|\Omega\rangle =\\
&= i^4\int \mathrm{d}^4x_1\mathrm{d}^4x_2\mathrm{d}^4x_{1'}\mathrm{d}^4x_{2'}\\
&\quad e^{-ip_{1'}x_{1'}}\left[\bar u^{s_{1'}}(\mathbf{k}_{1'})\left(-i\not\partial_{1'}+m\right)\right]_{\alpha_{1'}}\\
&\quad e^{-ip_{2'}x_{2'}}\left[\bar u^{s_{2'}}(\mathbf{k}_{2'})\left(-i\not\partial_{2'}+m\right)\right]_{\alpha_{2'}}\\
&\quad \langle\Omega|T\,\psi_{\alpha_{2'}}(x_{2'})\psi_{\alpha_{1'}}(x_{1'})\bar\psi_{\alpha_1}(x_1)\bar\psi_{\alpha_2}(x_2)|\Omega\rangle\\
&\quad \left[\left(i\overleftarrow{\not\partial}_1+m\right)u^s(\mathbf{p}_1)\right]_{\alpha_1}e^{ip_1x_1}\\
&\quad \left[\left(i\overleftarrow{\not\partial}_2+m\right)u^s(\mathbf{p}_2)\right]_{\alpha_2}e^{ip_2x_2}
\end{aligned}
$$

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:

$$\tilde{G}^{(n)}(p_1,\dots,p_n) = \int \prod_{i=1}^{n} \mathrm{d}^4 x_i e^{-ip_i x_i}\, G^{(n)}(x_1,\dots,x_n)$$

then the scattering amplitude becomes (resuming the previous calculation):

$$\begin{gathered}\langle f|i\rangle = \dots = i^4 \\ \left[\bar{u}^{s_{1'}}(\mathbf{k}_{1'})\,(-\not{p}_{1'} + m)\right]_{\alpha_{1'}} \\ \left[\bar{u}^{s_{2'}}(\mathbf{k}_{2'})\,(-\not{p}_{2'} + m)\right]_{\alpha_{2'}} \\ \tilde{G}^{(4)}_{\alpha_{1'}\alpha_{2'}\alpha_1\alpha_2}(p_{1'}, p_{2'}, -p_1, -p_2) \\ \left[(\not{p}_1 + m)\, u^s(\mathbf{p}_1)\right]_{\alpha_1} \\ \left[(\not{p}_2 + m)\, u^s(\mathbf{p}_2)\right]_{\alpha_2}\end{gathered}$$

We can get the same result much faster if we use the momentum space from the beginning:

$$\begin{gathered}\langle f|i\rangle = \langle \Omega| b_{\mathbf{p}_{2'}} b_{\mathbf{p}_{1'}} b^\dagger_{\mathbf{p}_1} b^\dagger_{\mathbf{p}_2} |\Omega\rangle = \langle \Omega| T\, b_{\mathbf{p}_{2'}} b_{\mathbf{p}_{1'}} b^\dagger_{\mathbf{p}_1} b^\dagger_{\mathbf{p}_2} |\Omega\rangle = \\ = \langle \Omega| T\, \bar{u}^s(\mathbf{p}_{2'}) \frac{1}{\tilde{S}(\mathbf{p}_{2'})} \tilde{\psi}(\mathbf{p}_{2'}) \bar{u}^s(\mathbf{p}_{1'}) \frac{1}{\tilde{S}(\mathbf{p}_{1'})} \tilde{\psi}(\mathbf{p}_{1'}) \tilde{\bar{\psi}}(-\mathbf{p}_1) \frac{1}{\tilde{S}(-\mathbf{p}_1)} u^s(\mathbf{p}_1) \tilde{\bar{\psi}}(-\mathbf{p}_2) \frac{1}{\tilde{S}(-\mathbf{p}_2)} u^s(\mathbf{p}_2) |\Omega\rangle = \\ = \left[\bar{u}^s(\mathbf{p}_{2'}) \frac{1}{\tilde{S}(\mathbf{p}_{2'})}\right]_{\alpha_{2'}} \left[\bar{u}^s(\mathbf{p}_{1'}) \frac{1}{\tilde{S}(\mathbf{p}_{1'})}\right]_{\alpha_{1'}} \\ \langle \Omega| T\, \tilde{\psi}_{\alpha_{2'}}(\mathbf{p}_{2'}) \tilde{\psi}_{\alpha_{1'}}(\mathbf{p}_{1'}) \tilde{\bar{\psi}}_{\alpha_1}(-\mathbf{p}_1) \tilde{\bar{\psi}}_{\alpha_2}(-\mathbf{p}_2) |\Omega\rangle \\ \left[\frac{1}{\tilde{S}(-\mathbf{p}_1)} u^s(\mathbf{p}_1)\right]_{\alpha_1} \left[\frac{1}{\tilde{S}(-\mathbf{p}_2)} u^s(\mathbf{p}_2)\right]_{\alpha_1}\end{gathered}$$

This is called Lehmann-Symanzik-Zimmermann (LSZ) reduction formula. One obtains similar expressions for other fields as well (if there were different creation operators been in the initial and final states). We now need to calculate the interacting Green functions.

Evaluation of the Interacting Green Functions

The interacting Green functions can be evaluated using the formula:

$$\begin{gathered}G^{(n)}(x_1,\dots,x_n) = \langle \Omega| T\, \phi_H(x_1)\dots\phi_H(x_n) |\Omega\rangle = \\ = \frac{\langle 0| T\, \phi_I(x_1)\dots\phi_I(x_n) S |0\rangle}{\langle 0|S|0\rangle}\end{gathered}$$

where

$$S = U_I(\infty, -\infty) = T \exp\left(-\frac{i}{\hbar}\int_{-\infty}^{\infty} H_1(t)\mathrm{d}t\right) = T \exp\left(-\frac{i}{\hbar}\int \mathrm{d}^4 x \mathcal{H}_1(x)\right)$$

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 ($\phi(\mathbf{x},t) = e^{iH_0 t}\phi(\mathbf{x},0)e^{-iH_0 t}$), 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:

$$U_I(t_{k-1},t_k)\phi_I(x_k)U_I(t_k,t_{k+1}) = U_I(t_{k-1},0)U_I^\dagger(t_k,0)\phi_I(x_k)U_I(t_k,0)U_I(0,t_{k+1}) = U_I(t_{k-1},0)\phi_H(x_k)U_I(0,t_{k+1})$$

$$U_I(0,-\infty)\,|0\rangle = U_I(0,-\infty)\left[|\Omega\rangle\,\langle\Omega| + \sum_{n\neq 0}|n\rangle\,\langle n|\right]|0\rangle = |\Omega\rangle\,\langle\Omega|0\rangle + \lim_{t\to-\infty}\sum_{n\neq 0} e^{iE_n t}\,|n\rangle\,\langle n|0\rangle = |\Omega\rangle\,\langle\Omega|0\rangle$$

$$\langle\Omega|\Omega\rangle = 1$$

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:

$$S = U(\infty,-\infty) = 1 - \frac{i}{\hbar}\int_{-\infty}^{\infty} H(t')U(t',-\infty)\mathrm{d}t' =$$

$$= 1 + \left(-\frac{i}{\hbar}\right)\int_{-\infty}^{\infty} H(t')U(t',-\infty)\mathrm{d}t' + \left(-\frac{i}{\hbar}\right)^2\int_{-\infty}^{\infty}\int_{-\infty}^{t'} H(t')H(t'')U(t'',-\infty)\mathrm{d}t'\mathrm{d}t'' =$$

$$= \cdots = \sum_{n=0}^{\infty} \left(-\frac{i}{\hbar}\right)^n \frac{1}{n!} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots T\{H(t_1)H(t_2)\cdots\} \mathrm{d}t_1 \mathrm{d}t_2 \cdots =$$

$$= T\exp\left(-\frac{i}{\hbar}\int_{-\infty}^{\infty} H(t)\mathrm{d}t\right) = T\exp\left(-\frac{i}{\hbar}\int_{-\infty}^{\infty} \mathrm{d}^4x\mathcal{H}(x)\right)$$

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 $k_1 + k_2 \to p_1 + p_2$:

$$\langle p_1p_2|S|k_1k_2\rangle = \frac{\int \mathrm{d}^4x_1\mathrm{d}^4x_2 e^{-i(k_1x_1+k_2x_2)} \int \mathrm{d}^4y_1\mathrm{d}^4y_2 e^{i(p_1y_1+p_2y_2)}}{\tilde{D}(k_1)\tilde{D}(k_2)\tilde{D}(p_1)\tilde{D}(p_2)}$$

$$\frac{\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(y_1)\phi_I(y_2)\exp\left(-\frac{i\lambda}{4\hbar}\int\mathrm{d}^4x\phi_I^4(x)\right)\}|0\rangle}{\langle 0|T\exp\left(-\frac{i\lambda}{4\hbar}\int\mathrm{d}^4x\phi_I^4(x)\right)|0\rangle} =$$

$$= \frac{\int \mathrm{d}^4x_1\mathrm{d}^4x_2 e^{-i(k_1x_1+k_2x_2)} \int \mathrm{d}^4y_1\mathrm{d}^4y_2 e^{i(p_1y_1+p_2y_2)}}{\tilde{D}(k_1)\tilde{D}(k_2)\tilde{D}(p_1)\tilde{D}(p_2)}$$

$$\left[\frac{\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(y_1)\phi_I(y_2)\}|0\rangle}{\langle 0|T\exp\left(-\frac{i\lambda}{4\hbar}\int\mathrm{d}^4x\phi_I^4(x)\right)|0\rangle} +\right.$$

$$+\frac{\left(-\frac{i\lambda}{4\hbar}\right)\int\mathrm{d}^4x\,\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(y_1)\phi_I(y_2)\phi_I^4(x)\}|0\rangle}{\langle 0|T\exp\left(-\frac{i\lambda}{4\hbar}\int\mathrm{d}^4x\phi_I^4(x)\right)|0\rangle}+$$

$$\left.+\frac{\left(-\frac{i\lambda}{4\hbar}\right)^2\int\mathrm{d}^4x\,\mathrm{d}^4y\,\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(y_1)\phi_I(y_2)\phi_I^4(x)\phi_I^4(y)\}|0\rangle}{\langle 0|T\exp\left(-\frac{i\lambda}{4\hbar}\int\mathrm{d}^4x\phi_I^4(x)\right)|0\rangle} + \cdots\right] =$$

$$= \frac{1}{\tilde{D}(k_1)\tilde{D}(k_2)\tilde{D}(p_1)\tilde{D}(p_2)}$$

$$\Big[(2\pi)^4\delta^{(4)}(p_1+p_2)(2\pi)^4\delta^{(4)}(k_1+k_2)\tilde{D}(p_1)\tilde{D}(k_1)+$$

$$(-i\lambda)6(2\pi)^4\delta^{(4)}(p_1+p_2-k_1-k_2)\tilde{D}(k_1)\tilde{D}(k_2)\tilde{D}(p_1)\tilde{D}(p_2)+$$

$$(-i\lambda)(\text{disconnected terms with not enough }\tilde{D}(\cdots)\text{s})+(-i\lambda)^2(\cdots)+\cdots\Big]=$$

$$=(2\pi)^4\delta^{(4)}(p_1+p_2-k_1-k_2)\left[6(-i\lambda)+3(-i\lambda)^2\int\frac{\mathrm{d}^4k}{(2\pi)^4}\tilde{D}(k)\tilde{D}(p1+p2-k)+(-i\lambda)^3(\cdots)+\cdots\right]$$

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):

$$\langle 0|T\{\psi_I(x)\bar{\psi}_I(y)\}|0\rangle \equiv S(x-y) = \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \tilde{S}(p) e^{-ip(x-y)}$$

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:

$$\langle 0|T\{A_\mu(x)A_\nu(y)\}|0\rangle \equiv D_{\mu\nu}(x-y) = \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \tilde{D}_{\mu\nu}(p) e^{-ip(x-y)}$$

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:

$$\langle 0|T\{\phi_I(x)\phi_I(y)\}|0\rangle \equiv D(x-y) = \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \tilde{D}(p) e^{-ip(x-y)}$$

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.

$$[c_k, c_l^\dagger] = \delta_{kl}$$
$$[c_k, c_l] = [c_k^\dagger, c_l^\dagger] = 0$$

so the commutation relations for $\hat\psi$ and $\hat\psi^\dagger$ are:

$$[\hat\psi(\mathbf{x}), \hat\psi^\dagger(\mathbf{y})] = \delta^{(3)}(\mathbf{x}-\mathbf{y})$$
$$[\hat\psi(\mathbf{x}), \hat\psi(\mathbf{y})] = [\hat\psi^\dagger(\mathbf{x}), \hat\psi^\dagger(\mathbf{y})] = 0$$

For fermions:

$$\hat\psi(\mathbf{x}) = \sum_k \sum_{s=1}^2 \psi_k^s(\mathbf{x}) c_k$$
$$\hat\psi^\dagger(\mathbf{x}) = \sum_k \sum_{s=1}^2 \psi_k^{s*}(\mathbf{x}) c_k^\dagger$$

where

$$\{c_k, c_l^\dagger\} = \delta_{kl}$$
$$\{c_k, c_l\} = \{c_k^\dagger, c_l^\dagger\} = 0$$

so the commutation relations for $\hat\psi$ and $\hat\psi^\dagger$ are:

$$\{\hat\psi(\mathbf{x}), \hat\psi^\dagger(\mathbf{y})\} = \delta^{(3)}(\mathbf{x}-\mathbf{y})$$
$$\{\hat\psi(\mathbf{x}), \hat\psi(\mathbf{y})\} = \{\hat\psi^\dagger(\mathbf{x}), \hat\psi^\dagger(\mathbf{y})\} = 0$$

The (interacting) Hamiltonian for both bosons and fermions is

$$i\hbar\partial_t\left|\Psi(t)\right\rangle = \hat H\left|\Psi(t)\right\rangle$$
$$\hat H = \hat T + \hat V = \sum_{ij} c_i^\dagger \langle i|T|j\rangle c_j + \tfrac{1}{2}\sum_{ijkl} c_i^\dagger c_j^\dagger \langle ij|V|kl\rangle c_l c_k$$

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 $|\mathbf{p}|/m \ll 1$ (we substitute $\omega = p_0 - m$):

$$\tilde D(p) = \frac{i}{p^2 - m^2 + i\epsilon} = \frac{i}{p_0^2 - \mathbf{p}^2 - m^2 + i\epsilon} = \frac{i}{\left(p_0 - \sqrt{\mathbf{p}^2 + m^2}\right)\left(p_0 + \sqrt{\mathbf{p}^2 + m^2}\right) + i\epsilon} \approx$$
$$\approx \frac{i}{\left(p_0 - m - \frac{\mathbf{p}^2}{2m}\right)\left(p_0 + m + \frac{\mathbf{p}^2}{2m}\right) + i\epsilon} = \frac{i}{\left(\omega - \frac{\mathbf{p}^2}{2m}\right)\left(\omega + 2m + \frac{\mathbf{p}^2}{2m}\right) + i\epsilon}$$

the behavior of the propagator in the vicinity of its positive frequency pole $\omega \approx \frac{\mathbf{p}^2}{2m}$ is (remember $\omega \to 0$ in the nonrelativistic limit):

$$\tilde D(p) \approx \frac{i}{\left(\omega - \frac{\mathbf{p}^2}{2m}\right)\left(\omega + 2m + \frac{\mathbf{p}^2}{2m}\right) + i\epsilon} \approx \frac{i}{\left(\omega - \frac{\mathbf{p}^2}{2m}\right)2m + i\epsilon} = \frac{1}{2m}\frac{i}{\omega - \frac{\mathbf{p}^2}{2m} + i\epsilon'}$$

Similarly for fermions:

$$\tilde S(p) = \frac{i(\not p + m)}{p^2 - m^2 + i\epsilon} = \frac{i(p^0\gamma_0 - p^j\gamma_j + m)}{p^2 - m^2 + i\epsilon} \approx \frac{1}{2m}\frac{i(p^0\gamma_0 - p^j\gamma_j + m)}{\omega - \frac{\mathbf{p}^2}{2m} + i\epsilon'} =$$
$$= \frac{1}{2m}\frac{i((\omega + m)\gamma_0 - p^j\gamma_j + m)}{\omega - \frac{\mathbf{p}^2}{2m} + i\epsilon'} \approx \frac{1}{2m}\frac{i(m\gamma_0 - p^j\gamma_j + m)}{\omega - \frac{\mathbf{p}^2}{2m} + i\epsilon'} =$$

$$= \frac{i\left(\frac{1}{2}(\gamma_0 + 1) - \frac{p^j\gamma_j}{2m}\right)}{\omega - \frac{\mathbf{p}^2}{2m} + i\epsilon'} \tag{4.1}$$

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 (4.1) 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 nonineracting 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:

$$G_0^+(x-y) = G_0^+(\mathbf{x},t_x,\mathbf{y},t_y) = i\int\frac{\mathrm{d}^3p\mathrm{d}E}{(2\pi\hbar)^4} G_0^+(\mathbf{p},E) e^{\frac{i}{\hbar}\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} e^{-\frac{i}{\hbar}E(t_x-t_y)} =$$
$$= i\int\frac{\mathrm{d}^3k\mathrm{d}\omega}{(2\pi)^4} G_0^+(k,\omega) e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})} e^{-i\omega(t_x-t_y)}$$

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:

$$G_0^+(k,t) = \int_{-\infty}^{\infty} \frac{\mathrm{d}\omega}{2\pi} e^{-i\omega t} G_0^+(k,\omega)$$
$$G_0^+(k,\omega) = \int_{-\infty}^{\infty} \mathrm{d}t\, e^{i\omega t} G_0^+(k,t)$$

The relation to the contraction of operators is:

$$G_0^+(\mathbf{k}, t_2 - t_1) = -i\theta_{t_2-t_1} \langle\Psi_0|c_\mathbf{k}(t_2)c_\mathbf{k}^\dagger(t_1)|\Psi_0\rangle$$

where $|\Psi_0\rangle$ is the ground state wavefunction and:

$$c_\mathbf{k}(t) = e^{iH_0t}c_\mathbf{k}e^{-iH_0t}$$

so to understand the meaning of $G_0^+(\mathbf{k}, t_2 - t_1)$, we write it as:

$$G_0^+(\mathbf{k}, t_2 - t_1) = -i\theta_{t_2-t_1} \langle\Psi_0|c_\mathbf{k}(t_2)c_\mathbf{k}^\dagger(t_1)|\Psi_0\rangle = -i\theta_{t_2-t_1} \langle\Psi_0|e^{iH_0t_2}c_\mathbf{k}e^{-iH_0(t_2-t_1)}c_\mathbf{k}^\dagger e^{-iH_0t_1}|\Psi_0\rangle =$$
$$= -i\theta_{t_2-t_1}\left(e^{-iH_0t_2}|\Psi_0\rangle\right)^\dagger\left(c_\mathbf{k}e^{-iH_0(t_2-t_1)}c_\mathbf{k}^\dagger e^{-iH_0t_1}|\Psi_0\rangle\right)$$

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 $e^{-iH_0t_2}|\Psi_0\rangle$).

Feynman Rules

We can deduce a set of rules, so that one doesn't have to repeat the whole calculation each time. For a scalar field we derived the rules above, for fermion and vector boson fields it's more difficult.

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 $H(p) \to Z(k) + Z(l)$:

$$\begin{aligned}
\langle f|S|i\rangle = \langle f|iT|i\rangle = \langle kl|iT|p\rangle = \langle\Omega|a_\mathbf{k}^r a_\mathbf{l}^s a_\mathbf{p}^\dagger|\Omega\rangle = \langle\Omega|T a_\mathbf{k}^r a_\mathbf{l}^s a_\mathbf{p}^\dagger|\Omega\rangle = \\
= \langle\Omega|T\epsilon_\mu^{r*}(\mathbf{k})\frac{k^2}{i}\tilde{A}^\mu(k)\epsilon_\nu^{s*}(\mathbf{l})\frac{l^2}{i}\tilde{A}^\nu(l)\frac{1}{\tilde{D}(p)}\tilde\phi(-p)|\Omega\rangle = \\
= \frac{\epsilon_\mu^{r*}(\mathbf{k})\epsilon_\nu^{s*}(\mathbf{l})}{\frac{i}{k^2}\frac{i}{l^2}\tilde{D}(p)}\langle\Omega|T\tilde{A}^\mu(k)\tilde{A}^\nu(l)\tilde\phi(-p)|\Omega\rangle \\
= \frac{\epsilon_\mu^{r*}(\mathbf{k})\epsilon_\nu^{s*}(\mathbf{l})}{\frac{i}{k^2}\frac{i}{l^2}\tilde{D}(p)} i\lambda(2\pi)^4\delta(k+l-p)\tilde{D}^\mu{}_\alpha(k)\tilde{D}^{\nu\alpha}(l)\tilde{D}(p) \\
= \frac{\epsilon_\mu^{r*}(\mathbf{k})\epsilon_\nu^{s*}(\mathbf{l})}{\frac{i}{k^2}\frac{i}{l^2}\tilde{D}(p)} i\lambda(2\pi)^4\delta(k+l-p)\frac{-ig^\mu{}_\alpha}{k^2}\frac{-ig^{\nu\alpha}}{l^2}\tilde{D}(p) \\
= \epsilon_\mu^{r*}(\mathbf{k})\epsilon_\nu^{s*}(\mathbf{l}) i\lambda(2\pi)^4\delta(k+l-p)g^\mu{}_\alpha g^{\nu\alpha} \\
= \epsilon_\mu^{r*}(\mathbf{k})\epsilon_\nu^{s*}(\mathbf{l}) i\lambda(2\pi)^4\delta(k+l-p)g^{\mu\nu} \\
= i\lambda(2\pi)^4\delta(k+l-p)\epsilon_\mu^{r*}(\mathbf{k})\epsilon^{s\mu*}(\mathbf{l})
\end{aligned}$$

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 $\mathcal{L}_{ee\gamma} = -\lambda\bar{e}\gamma^\mu eA_\mu$ 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 $e^-(p_1) + e^+(p_2) \to \gamma(q) \to e^-(k_1) + e^+(k_2)$:

$$\langle f|S|i\rangle = \langle f|iT|i\rangle = \langle k_1 k_2|iT|p_1 p_2\rangle = \langle\Omega|b^r_{\mathbf{k}_1} d^s_{\mathbf{k}_2} b^{t\dagger}_{\mathbf{p}_1} d^{u\dagger}_{\mathbf{p}_2}|\Omega\rangle =$$

$$= \langle\Omega|T b^r_{\mathbf{k}_1} d^s_{\mathbf{k}_2} b^{t\dagger}_{\mathbf{p}_1} d^{u\dagger}_{\mathbf{p}_2}|\Omega\rangle =$$

$$= \langle\Omega|T\left[\bar{u}^r(\mathbf{k}_1)\frac{1}{\tilde{S}(k_1)}\tilde{\psi}(k_1)\right]\left[-\tilde{\bar{\psi}}(k_2)\frac{1}{\tilde{S}(-k_2)}v^s(\mathbf{k}_2)\right]\left[\tilde{\bar{\psi}}(-p_1)\frac{1}{\tilde{S}(-p_1)}u^t(\mathbf{p}_1)\right]\left[-\bar{v}^u(\mathbf{p}_2)\frac{1}{\tilde{S}(p_2)}\tilde{\psi}(-p_2)\right]|\Omega\rangle =$$

$$= \left[\bar{u}^r(\mathbf{k}_1)\frac{1}{\tilde{S}(k_1)}\right]\left[\bar{v}^u(\mathbf{p}_2)\frac{1}{\tilde{S}(p_2)}\right]$$
$$\langle\Omega|T\tilde{\psi}(k_1)\tilde{\bar{\psi}}(k_2)\tilde{\bar{\psi}}(-p_1)\tilde{\psi}(-p_2)|\Omega\rangle$$
$$\left[\frac{1}{\tilde{S}(-k_2)}v^s(\mathbf{k}_2)\right]\left[\frac{1}{\tilde{S}(-p_1)}u^t(\mathbf{p}_1)\right] =$$

$$= \left[\bar{u}^r(\mathbf{k}_1)\frac{1}{\tilde{S}(k_1)}\right]\left[\bar{v}^u(\mathbf{p}_2)\frac{1}{\tilde{S}(p_2)}\right]$$
$$(-i\lambda)^2(2\pi)^4\delta(k_1+k_2-p_1-p_2)\left[\tilde{S}(k_1)\gamma^\mu\tilde{S}(-k_2)D_{\mu\nu}(k_1+k_2)\tilde{S}(p_2)\gamma^\nu\tilde{S}(-p_1)+\right.$$
$$\left.+\tilde{S}(k_1)\gamma^\mu\tilde{S}(-p_1)D_{\mu\nu}(k_1-p_1)\tilde{S}(p_2)\gamma^\nu\tilde{S}(-k_2)\right]$$
$$\left[\frac{1}{\tilde{S}(-k_2)}v^s(\mathbf{k}_2)\right]\left[\frac{1}{\tilde{S}(-p_1)}u^t(\mathbf{p}_1)\right] =$$

$$= -\lambda^2(2\pi)^4\delta(k_1+k_2-p_1-p_2)\left[\bar{u}^r(\mathbf{k}_1)\gamma^\mu v^s(\mathbf{k}_2)\frac{1}{(k_1+k_2)^2}\bar{v}^u(\mathbf{p}_2)\gamma_\mu u^t(\mathbf{p}_1)+\right.$$
$$\left.+\bar{u}^r(\mathbf{k}_1)\gamma^\mu u^t(\mathbf{p}_1)\frac{1}{(k_1-p_1)^2}\bar{v}^u(\mathbf{p}_2)\gamma_\mu v^s(\mathbf{k}_2)\right]$$

where we used the fact, that the interacting Green function is in the lowest nonzero order equal to:

$$\langle\Omega|T\tilde{\psi}(k_1)\tilde{\bar{\psi}}(k_2)\tilde{\bar{\psi}}(-p_1)\tilde{\psi}(-p_2)|\Omega\rangle =$$
$$= (-i\lambda)^2(2\pi)^4\delta(k_1+k_2-p_1-p_2)\left[\tilde{S}(k_1)\gamma^\mu\tilde{S}(-k_2)D_{\mu\nu}(k_1+k_2)\tilde{S}(p_2)\gamma^\nu\tilde{S}(-p_1)+\right.$$
$$\left.+\tilde{S}(k_1)\gamma^\mu\tilde{S}(-p_1)D_{\mu\nu}(k_1-p_1)\tilde{S}(p_2)\gamma^\nu\tilde{S}(-k_2)\right]$$

4.2.3 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 element can be derived directly from the Lagrangian:

$$\mathcal{L} = -\frac{G_\mu}{\sqrt2}[\bar\psi_{\nu_\mu}\gamma^\mu(1-\gamma_5)\psi_\mu][\bar\psi_e\gamma^\mu(1-\gamma_5)\psi_{\nu_e}]$$

with

$$\frac{G_\mu}{\sqrt2} = \frac{g^2}{8m_W^2}$$

This is the universal V-A theory Lagrangian (after adding the h.c. term).

4.3 Quantum Mechanics

4.3.1 From QED to Quantum Mechanics

The QED Lagrangian density is

$$\mathcal{L} = \bar\psi(i\hbar c\gamma^\mu D_\mu - mc^2)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

where

$$\psi = \begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\\ \psi_4\end{pmatrix}$$

and

$$D_\mu = \partial_\mu + \frac{i}{\hbar}eA_\mu$$

is the gauge covariant derivative and (e is the elementary charge, which is 1 in atomic units, i.e. the electron has a charge $-e$)

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

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.

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

$$j^0 = c\bar\psi\gamma^0\psi = c\psi^\dagger\psi = c\varphi^\dagger\varphi$$

$$j^i = c\bar\psi\gamma^i\psi = c\psi^\dagger\alpha^i\psi = c\varphi^\dagger\alpha^i\varphi$$

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:

$$0 = (-i\hbar c\gamma^\mu D_\mu - mc^2)(i\hbar c\gamma^\nu D_\nu - mc^2)\psi = (\hbar^2c^2\gamma^\mu\gamma^\nu D_\mu D_\nu + m^2c^4)\psi =$$

$$= (\hbar^2c^2g^{\mu\nu}D_\mu D_\nu + m^2c^4)\psi = (\hbar^2c^2D^\mu D_\mu + m^2c^4)\psi$$

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 $D^\mu D_\mu$:

$$D^\mu D_\mu = g^{\mu\nu}D_\mu D_\nu = \partial^\mu\partial_\mu + ie((\partial^\mu A_\mu) + 2A^\mu\partial_\mu) - e^2A^\mu A_\mu =$$

$$= \partial^\mu\partial_\mu + ie((\partial^0A_0) + 2A^0\partial_0 + (\partial^iA_i) + 2A^i\partial_i) - e^2(A^0A_0 + A^iA_i) =$$

$$= \partial^\mu\partial_\mu + i\frac{1}{c^2}\frac{\partial V}{\partial t} + 2i\frac{V}{c^2}\frac{\partial}{\partial t} + ie(\partial^iA_i) + 2ieA^i\partial_i - \frac{V^2}{c^2} - e^2A^iA_i$$

The nonrelativistic limit can also be applied directly to the Klein-Gordon equation:

$$0 = (\hbar^2c^2D^\mu D_\mu + m^2c^4)\psi =$$

$$= \left(\hbar^2c^2\partial^\mu\partial_\mu + i\frac{\partial V}{\partial t} + 2iV\frac{\partial}{\partial t} + i\hbar ec^2(\partial^iA_i) + 2i\hbar ec^2A^i\partial_i - V^2 - e^2c^2A^iA_i + m^2c^4\right)e^{-\frac{i}{\hbar}mc^2t}\varphi =$$

$$= \left(\hbar^2\frac{\partial^2}{\partial t^2} - c^2\hbar^2\nabla^2 + 2iV\frac{\partial}{\partial t} + i\frac{\partial V}{\partial t} + i\hbar ec^2(\partial^iA_i) + 2i\hbar ec^2A^i\partial_i - V^2 - e^2c^2A^iA_i + m^2c^4\right)e^{-\frac{i}{\hbar}mc^2t}\varphi =$$

$$= e^{-\frac{i}{\hbar}mc^2t}\left(\hbar^2(-\frac{i}{\hbar}mc^2 + \frac{\partial}{\partial t})^2 - \hbar^2c^2\nabla^2 + 2iV(-\frac{i}{\hbar}mc^2 + \frac{\partial}{\partial t}) + i\frac{\partial V}{\partial t} + i\hbar ec^2(\partial^iA_i) + 2i\hbar ec^2A^i\partial_i - V^2 + \right.$$

$$-e^2c^2A^iA_i + m^2c^4\big)\varphi =$$

$$= e^{-\frac{i}{\hbar}mc^2t}\left(-2i\hbar mc^2\frac{\partial}{\partial t} + \hbar^2\frac{\partial^2}{\partial t^2} - c^2\hbar^2\nabla^2 + 2Vm\frac{c^2}{\hbar} + 2iV\frac{\partial}{\partial t} + i\frac{\partial V}{\partial t} + i\hbar ec^2(\partial^iA_i) + 2i\hbar ec^2A^i\partial_i - V^2 + \right.$$

$$\left. -e^2c^2A^iA_i\right)\varphi =$$

$$= -2mc^2e^{-\frac{i}{\hbar}mc^2t}\left(i\hbar\frac{\partial}{\partial t} + \hbar^2\frac{\nabla^2}{2m} - V - \frac{1}{2mc^2}\frac{\partial^2}{\partial t^2} - \frac{i}{2mc^2}\frac{\partial V}{\partial t} + \frac{V^2}{2mc^2} - \frac{iV}{mc^2}\frac{\partial}{\partial t} + \right.$$

$$\left. -\frac{i\hbar e}{2m}\partial^iA_i - \frac{i\hbar e}{m}A^i\partial_i + \frac{e^2}{2m}A^iA_i\right)\varphi$$

Taking the limit $c \to \infty$ we again recover the Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}\varphi = \left(-\hbar^2\frac{\nabla^2}{2m} + V + \frac{i\hbar e}{2m}\partial^iA_i + \frac{i\hbar e}{m}A^i\partial_i - \frac{e^2}{2m}A^iA_i\right)\varphi\,,$$

we rewrite the right hand side a little bit:

$$i\hbar\frac{\partial}{\partial t}\varphi = \left(\frac{\hbar^2}{2m}(\partial^i\partial_i + \frac{i}{\hbar}e\partial^iA_i + 2\frac{i}{\hbar}eA^i\partial_i - \frac{e^2}{\hbar^2}A^iA_i) + V\right)\varphi\,,$$

$$i\hbar\frac{\partial}{\partial t}\varphi = \left(\frac{\hbar^2}{2m}(\partial^i + \frac{i}{\hbar}eA^i)(\partial_i + \frac{i}{\hbar}eA_i) + V\right)\varphi\,,$$

$$i\hbar\frac{\partial}{\partial t}\varphi = \left(\frac{1}{2m}\hbar^2D^iD_i + V\right)\varphi\,,$$

Using (see the appendix for details):

$$\hbar^2D^iD_i = -\hbar^2\delta_{ij}D^iD^j = -\hbar^2\left(\frac{i}{\hbar}(\mathbf{p} - e\mathbf{A})\right)^2 = (\mathbf{p} - e\mathbf{A})^2$$

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:

$$0 = (\hbar^2c^2D^\mu D_\nu + m^2c^4)\psi =$$

$$= (\hbar^2c^2D^0D_0 + \hbar^2c^2D^iD_i + m^2c^4)\psi =$$

$$= 2mc^2\left(\frac{\hbar^2}{2m}D^0D_0 + \frac{\hbar^2}{2m}D^iD_i + \tfrac{1}{2}mc^2\right)\psi =$$

$$= 2mc^2\left(\frac{\hbar^2}{2m}\left(\partial^0 + \frac{i}{\hbar}eA^0\right)\left(\partial_0 + \frac{i}{\hbar}eA_0\right) + \tfrac{1}{2}mc^2 + \frac{\hbar^2}{2m}D^iD_i\right)e^{-\frac{i}{\hbar}mc^2t}\varphi =$$

$$= 2mc^2 \left(\frac{\hbar^2}{2m}\left(\partial^0 + \frac{i}{\hbar}eA^0\right) e^{-\frac{i}{\hbar}mc^2 t}\left(\partial_0 - \frac{i}{\hbar}mc + \frac{i}{\hbar}eA_0\right) + \tfrac{1}{2}mc^2 + \frac{\hbar^2}{2m}D^i D_i\right)\varphi =$$

$$= 2mc^2 e^{-\frac{i}{\hbar}mc^2 t}\left(\frac{\hbar^2}{2m}\left(\partial^0 - \frac{i}{\hbar}mc + \frac{i}{\hbar}eA^0\right)\left(\partial_0 - \frac{i}{\hbar}mc + \frac{i}{\hbar}eA_0\right) + \tfrac{1}{2}mc^2 + \frac{\hbar^2}{2m}D^i D_i\right)\varphi =$$

$$= 2mc^2 e^{-\frac{i}{\hbar}mc^2 t}\left(\frac{\hbar^2}{2m}\partial^0\partial_0 - \tfrac{1}{2}mc^2 - \frac{e^2 A^0 A_0}{2m} + ceA^0 + \frac{\hbar^2}{m}\frac{i}{\hbar}e(\partial^0 A^0 + A^0\partial^0) - i\hbar c\partial_0 + \tfrac{1}{2}mc^2 + \frac{\hbar^2}{2m}D^i D_i\right)\varphi =$$

$$= 2mc^2 e^{-\frac{i}{\hbar}mc^2 t}\left(-i\hbar\frac{\partial}{\partial t} + \frac{\hbar^2}{2m}D^i D_i + ceA^0 + \frac{\hbar^2}{2mc^2}\frac{\partial^2}{\partial t^2} - \frac{e^2\phi^2}{2mc^2} + \frac{ie\hbar}{mc^2}(\frac{\partial}{\partial t}\phi + \phi\frac{\partial}{\partial t})\right)\varphi =$$

$$= 2mc^2 e^{-\frac{i}{\hbar}mc^2 t}\left(-i\hbar\frac{\partial}{\partial t} + \frac{(\mathbf{p} - e\mathbf{A})^2}{2m} + e\phi + \frac{\hbar^2}{2mc^2}\frac{\partial^2}{\partial t^2} - \frac{e^2\phi^2}{2mc^2} + \frac{ie\hbar}{mc^2}(\frac{\partial}{\partial t}\phi + \phi\frac{\partial}{\partial t})\right)\varphi$$

and letting $c \to \infty$ we get the Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}\varphi = \left(\frac{(\mathbf{p} - e\mathbf{A})^2}{2m} + e\phi\right)\varphi$$

4.3.2 Perturbation Theory

We want to solve the equation:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left|\psi(t)\right\rangle = H(t)\left|\psi(t)\right\rangle \tag{4.2}$$

with $H(t) = H^0 + H^1(t)$, where H^0 is time-independent part whose eigenvalue problem has been solved:

$$H^0\left|n^0\right\rangle = E_n^0\left|n^0\right\rangle$$

and $H^1(t)$ is a small time-dependent perturbation. $\left|n^0\right\rangle$ form a complete basis, so we can express $\left|\psi(t)\right\rangle$ in this basis:

$$\left|\psi(t)\right\rangle = \sum_n d_n(t)e^{-\frac{i}{\hbar}E_n^0 t}\left|n^0\right\rangle \tag{4.3}$$

Substituting this into (4.2), we get:

$$\sum_n \left(i\hbar\frac{\mathrm{d}}{\mathrm{d}t}d_n(t) + E_n^0 d_n(t)\right)e^{-\frac{i}{\hbar}E_n^0 t}\left|n^0\right\rangle = \sum_n \left(E_n^0 d_n(t) + H^1 d_n(t)\right)e^{-\frac{i}{\hbar}E_n^0 t}\left|n^0\right\rangle$$

so:

$$\sum_n i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left(d_n(t)\right)e^{-\frac{i}{\hbar}E_n^0 t}\left|n^0\right\rangle = \sum_n d_n(t)e^{-\frac{i}{\hbar}E_n^0 t}H^1\left|n^0\right\rangle$$

Choosing some particular state $\left|f^0\right\rangle$ of the H^0 Hamiltonian, we multiply the equation from the left by $\left\langle f^0\right| e^{\frac{i}{\hbar}E_f^0 t}$:

$$\sum_n i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left(d_n(t)\right)e^{iw_{fn}t}\left\langle f^0|n^0\right\rangle = \sum_n d_n(t)e^{iw_{fn}t}\left\langle f^0|H^1|n^0\right\rangle$$

where $w_{fn} = \frac{E^0_f - E^0_n}{\hbar}$. 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:

$$i\hbar \left((d_f(t) - d_f(t_1)) \right) = \sum_n \int_{t_1}^{t} d_n(t') e^{iw_{fn}t'} \langle f^0 | H^1(t') | n^0 \rangle \, \mathrm{d}t'$$

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:

$$d_f(t) = \delta_{fi} - \frac{i}{\hbar} \sum_n \int_{t_1}^{t} d_n(t') e^{iw_{fn}t'} \langle f^0 | H^1(t') | n^0 \rangle \, \mathrm{d}t' \tag{4.4}$$

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 (4.4):

$$d_f(t) = \delta_{fi} - \frac{i}{\hbar} \int_{t_1}^{t} e^{iw_{fi}t'} \langle f^0 | H^1(t') | i^0 \rangle \, \mathrm{d}t'$$

In the second order, we take the last solution, substitute into the right hand side of (4.4) 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 (4.3) 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:

$$E_i^0 = \langle i^0|H^0|i^0\rangle$$

$$E_i^1 = \langle i^0|H^1|i^0\rangle$$

$$E_i^2 = -\sum_{n\neq i}\frac{|\langle n^0|H^1|i^0\rangle|^2}{E_n^0 - E_i^0}$$

4.3.3 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:

$$G_+(\mathbf{r},\mathbf{r}') = \langle\mathbf{r}|G_+|\mathbf{r}'\rangle = \langle\mathbf{r}|\,\frac{1}{E_k - H_0 + i\epsilon}\,|\mathbf{r}'\rangle =$$

$$= \int \mathrm{d}^3k' \frac{\langle \mathbf{r}|\mathbf{k}'\rangle \langle \mathbf{k}'|\mathbf{r}'\rangle}{E_k - E_{k'} + i\epsilon} = \int \mathrm{d}^3k' \frac{e^{i\mathbf{k}'\cdot(\mathbf{r}-\mathbf{r}')}}{E_k - E_{k'} + i\epsilon} = \frac{2m}{\hbar^2}\int \mathrm{d}^3k' \frac{e^{i\mathbf{k}'\cdot(\mathbf{r}-\mathbf{r}')}}{k^2 - k'^2 + i\epsilon} =$$

$$= \frac{4\pi m}{\hbar^2 i|\mathbf{r}-\mathbf{r}'|}\int_{-\infty}^{\infty} \mathrm{d}^3k' k' \frac{e^{ik'|\mathbf{r}-\mathbf{r}'|}}{k^2 - k'^2 + i\epsilon} = \frac{4\pi m}{\hbar^2 i|\mathbf{r}-\mathbf{r}'|}(2\pi i)k\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{2k} =$$

$$= \frac{4\pi^2 m e^{ik|\mathbf{r}-\mathbf{r}'|}}{\hbar^2|\mathbf{r}-\mathbf{r}'|}$$

$$G_-(\mathbf{r},\mathbf{r}') = \langle \mathbf{r}|G_-|\mathbf{r}'\rangle = \langle \mathbf{r}| \frac{1}{E_k - H_0 - i\epsilon}|\mathbf{r}'\rangle = \cdots = \frac{4\pi^2 m e^{-ik|\mathbf{r}-\mathbf{r}'|}}{\hbar^2|\mathbf{r}-\mathbf{r}'|}$$

Assuming $|\mathbf{r}'| \ll |\mathbf{r}|$, we can taylor expand $|\mathbf{r}-\mathbf{r}'|$:

$$|\mathbf{r}-\mathbf{r}'| = e^{-\mathbf{r}'\cdot\nabla}|\mathbf{r}| = \left(1 - \mathbf{r}'\cdot\nabla + (-\mathbf{r}'\cdot\nabla)^2 + O\left(r'^3\right)\right)|\mathbf{r}| = |\mathbf{r}| - \mathbf{r}'\cdot\nabla|\mathbf{r}| + O\left(r'^2\right) =$$

$$= r - \mathbf{r}'\cdot\hat{\mathbf{r}} + O\left(r'^2\right)$$

and simplify the result even further:

$$G_+(\mathbf{r},\mathbf{r}') = \frac{4\pi^2 m}{\hbar^2}\frac{e^{ikr}}{r}e^{-ik\mathbf{r}'\cdot\hat{\mathbf{r}}}$$

$$G_-(\mathbf{r},\mathbf{r}') = \frac{4\pi^2 m}{\hbar^2}\frac{e^{-ikr}}{r}e^{ik\mathbf{r}'\cdot\hat{\mathbf{r}}}$$

Note: both functions may be divided by the factor $(2\pi)^3$ due to the momentum integration.

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:

$$T = \frac{V}{1 - G_+V} = V(1 + G_+V + (G_+V)^2 + \cdots) =$$

$$= V + VG_+V + VG_+VG_+V + \cdots =$$

$$= V + V\frac{1}{E_k - H_0 + i\epsilon}V + V\frac{1}{E_k - H_0 + i\epsilon}V\frac{1}{E_k - H_0 + i\epsilon}V + \cdots$$

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:

$$\psi(\mathbf{r}) = \langle\mathbf{r}|\psi\rangle = \langle\mathbf{r}|\mathbf{k}\rangle + \langle\mathbf{r}|G_+T|\mathbf{k}\rangle = \langle\mathbf{r}|\mathbf{k}\rangle + \int \mathrm{d}^3r'\,\langle\mathbf{r}|G_+|\mathbf{r}'\rangle\,\langle\mathbf{r}'|T|\mathbf{k}\rangle =$$

$$= \langle\mathbf{r}|\mathbf{k}\rangle + \int \mathrm{d}^3r'\mathrm{d}^3k'\,\langle\mathbf{r}|G_+|\mathbf{r}'\rangle\,\langle\mathbf{r}'|\mathbf{k}'\rangle\,\langle\mathbf{k}'|T|\mathbf{k}\rangle =$$

$$= e^{i\mathbf{k}\cdot\mathbf{r}} + \int \mathrm{d}^3r'\mathrm{d}^3k' G_+(\mathbf{r},\mathbf{r}')e^{i\mathbf{k}'\cdot\mathbf{r}'}\,\langle\mathbf{k}'|T|\mathbf{k}\rangle$$

Plugging the representation of the Green function for $|\mathbf{r}'| \ll |\mathbf{r}|$ in:

$$\psi(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} + \frac{4\pi^2 m}{\hbar^2}\frac{e^{ikr}}{r}\int \mathrm{d}^3r'\mathrm{d}^3k' e^{-ik\mathbf{r}'\cdot\hat{\mathbf{r}}}e^{i\mathbf{k}'\cdot\mathbf{r}'}\,\langle\mathbf{k}'|T|\mathbf{k}\rangle =$$

$$= e^{i\mathbf{k}\cdot\mathbf{r}} + \frac{4\pi^2 m}{\hbar^2}\frac{e^{ikr}}{r}\int \mathrm{d}^3r'\mathrm{d}^3k' e^{i\mathbf{r}'\cdot(\mathbf{k}'-k\hat{\mathbf{r}})}\,\langle\mathbf{k}'|T|\mathbf{k}\rangle =$$

$$= e^{i\mathbf{k}\cdot\mathbf{r}} + \frac{4\pi^2 m}{\hbar^2}\frac{e^{ikr}}{r}\int \mathrm{d}^3k'\delta(\mathbf{k}'-k\hat{\mathbf{r}})\,\langle\mathbf{k}'|T|\mathbf{k}\rangle =$$

$$= e^{i\mathbf{k}\cdot\mathbf{r}} + \frac{4\pi^2 m}{\hbar^2}\frac{e^{ikr}}{r}\,\langle k\hat{\mathbf{r}}|T|\mathbf{k}\rangle =$$

$$= e^{i\mathbf{k}\cdot\mathbf{r}} + f(\theta,\phi)\,\frac{e^{ikr}}{r}$$

where the scattering amplitude $f(\theta,\phi)$ is:

$$f(\theta,\phi) = \frac{4\pi^2 m}{\hbar^2}\,\langle k\hat{\mathbf{r}}|T|\mathbf{k}\rangle = \frac{4\pi^2 m}{\hbar^2}\,\langle\mathbf{k}'|T|\mathbf{k}\rangle$$

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

$$\mathbf{j}_o\cdot\hat{\mathbf{r}} = \frac{\hbar}{2mi}\left(\psi^*\nabla\psi - \psi\nabla\psi^*\right)\cdot\hat{\mathbf{r}} = \frac{\hbar}{2mi}\left(\psi^*\frac{\partial}{\partial r}\psi - \psi\frac{\partial}{\partial r}\psi^*\right) =$$

$$= \frac{\hbar}{2mi}\left(f^*(\theta,\phi)\frac{e^{-ikr}}{r}\frac{\partial}{\partial r}\left(f(\theta,\phi)\frac{e^{ikr}}{r}\right) - f(\theta,\phi)\frac{e^{ikr}}{r}\frac{\partial}{\partial r}\left(f^*(\theta,\phi)\frac{e^{-ikr}}{r}\right)\right) =$$

$$= \frac{\hbar k}{m}\left(\frac{1}{r^2}+\frac{i}{kr^3}\right)|f(\theta,\phi)|^2$$

Let's write the explicit formula for the transition matrix:

$$\langle\mathbf{k}'|T|\mathbf{k}\rangle = \int \mathrm{d}^3 r \,\langle\mathbf{k}'|\mathbf{r}\rangle\langle\mathbf{r}|V|\mathbf{k}\rangle + \int \mathrm{d}^3 r \mathrm{d}^3 r' \,\langle\mathbf{k}'|\mathbf{r}\rangle\langle\mathbf{r}|VG_+|\mathbf{r}'\rangle\langle\mathbf{r}'|V|\mathbf{k}\rangle + \cdots =$$

$$= \int \mathrm{d}^3 r e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}} V(\mathbf{r}) + \int \mathrm{d}^3 r \mathrm{d}^3 r' e^{-i\mathbf{k}'\cdot\mathbf{r}} V(\mathbf{r}) \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} V(\mathbf{r}') e^{i\mathbf{k}\cdot\mathbf{r}'} + \cdots =$$

The Born approximation is just the first term:

$$\langle\mathbf{k}'|T|\mathbf{k}\rangle \approx \int \mathrm{d}^3 r e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}} V(\mathbf{r}) = \int \mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\phi\, e^{iqr\cos\theta} V(r) r^2 \sin\theta =$$

$$= 4\pi \int_0^\infty rV(r)\sin(qr)\,\mathrm{d}r$$

4.4 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:

$$\begin{gathered} H|\Omega\rangle = (H_0 + H_1)|\Omega\rangle = E|\Omega\rangle \\ \langle 0|H_0 + H_1|\Omega\rangle = E\langle 0|\Omega\rangle \\ E_0\langle 0|\Omega\rangle + \langle 0|H_1|\Omega\rangle = E\langle 0|\Omega\rangle \\ E = E_0 + \frac{\langle 0|H_1|\Omega\rangle}{\langle 0|\Omega\rangle} \end{gathered}$$

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:

$$\Delta E = E - E_0 = \frac{\langle 0|H_1|\Omega\rangle}{\langle 0|\Omega\rangle} = \frac{\langle 0|H_1 U(0,-\infty)|0\rangle}{\langle 0|U(0,-\infty)|0\rangle} =$$
$$= \lim_{t\to 0}\frac{\langle 0|H_1 U(t,-\infty)|0\rangle}{\langle 0|U(t,-\infty)|0\rangle} = \lim_{t\to 0}\frac{\langle 0|i\partial_t U(t,-\infty)|0\rangle}{\langle 0|U(t,-\infty)|0\rangle} = \lim_{t\to 0}\frac{i\partial_t \langle 0|U(t,-\infty)|0\rangle}{\langle 0|U(t,-\infty)|0\rangle} =$$
$$= \lim_{t\to 0} i\partial_t \log\langle 0|U(t,-\infty)|0\rangle \equiv \lim_{t\to\infty(1-i\epsilon)} i\frac{\mathrm{d}}{\mathrm{d}t}\log\langle 0|U(t,-\infty)|0\rangle$$

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:

$$G(\mathbf{k}_1,\mathbf{k}_2,t_2-t_1) = i\langle\Omega|Tc_{\mathbf{k}_2}(t_2)c^\dagger_{\mathbf{k}_1}(t_1)|\Omega\rangle = i\frac{\langle 0|Tc_{\mathbf{k}_2}(t_2)c^\dagger_{\mathbf{k}_1}(t_1)U(\infty,-\infty)|0\rangle}{\langle 0|U(\infty,-\infty)|0\rangle}$$

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:

$$|i\rangle = a^\dagger_{\mathbf{k}} b^{s\dagger}_{\mathbf{l}}|0\rangle$$
$$|f\rangle = a^\dagger_{\mathbf{p}} a^{r\dagger}_{\mathbf{q}}|0\rangle$$

This is just an example, the $|i\rangle$ and $|f\rangle$ states can contain any number of (arbitrary) particles.

4.5 Appendix

4.5.1 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:

$$[\phi] = \mathrm{V} = \frac{\mathrm{kg\,m^2}}{\mathrm{A\,s^3}} = \mathrm{M}$$

$$[A_\mu] = \frac{[\phi]}{[c]} = \frac{\mathrm{V\,s}}{\mathrm{m}} = \frac{\mathrm{kg\,m}}{\mathrm{A\,s^2}} = \mathrm{M}$$

$$[c] = \frac{\mathrm{m}}{\mathrm{s}} = 1$$

$$[e] = \mathrm{C} = \mathrm{A\,s} = 1$$

$$[\hbar] = \mathrm{J\,s} = \frac{\mathrm{m^2\,kg}}{\mathrm{s}} = 1$$

$$[\partial_\mu] = \frac{1}{\mathrm{m}} = \mathrm{M}$$

$$[F_{\mu\nu}] = [\partial_\mu A_\nu] = \frac{\mathrm{kg}}{\mathrm{A\,s^2}} = \mathrm{M^2}$$

$$[\mathcal{L}] = [F_{\mu\nu}]^2 = \frac{\mathrm{kg^2}}{\mathrm{A^2\,s^4}} = \mathrm{M^4}$$

$$[\psi] = \frac{\mathrm{kg^{\frac{1}{2}}}}{\mathrm{A\,m\,s}} = \mathrm{M^{\frac{3}{2}}}$$

The SI units are useful for checking that the c, e and $\hbar$ constants are at correct places in the expression.

4.5.2 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:

$$v^2 = v_\mu v^\mu = g_{\mu\nu} v^\mu v^\nu = -\gamma^2 c^2 + \mathbf{v}^2 = -\frac{c^2}{1 - \frac{\mathbf{v}^2}{c^2}} + \mathbf{v}^2 = c^2$$

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:

$$p^2 = p_\mu p^\mu = (p^0)^2 - \mathbf{p}^2 = (p_0)^2 - \mathbf{p}^2 = \frac{E^2}{c^2} - \mathbf{p}^2$$

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:

$$\left(\frac{\hbar^2}{c^2}\frac{\partial^2}{\partial t^2} - \hbar^2\nabla^2 + m^2c^2\right)\psi = 0$$

can be for signature -2 written as:

$$(+\hbar^2\partial^2 + m^2c^2)\psi = (-p^2 + m^2c^2)\psi = 0$$

and for $+2$ as:

$$(-\hbar^2\partial^2 + m^2c^2)\psi = (p^2 + m^2c^2)\psi = 0$$

Now if the particle is not moving ($\mathbf{p} = 0$), then it's energy $E = mc^2$, where m is the rest mass, so $|p| = \frac{E}{c} = mc$, from which:

$$E^2 = p^2c^2 + \mathbf{p}^2c^2 = m^2c^4 + \mathbf{p}^2c^2$$
$$E = \sqrt{m^2c^4 + \mathbf{p}^2c^2} = mc^2\sqrt{1 + \frac{\mathbf{p}^2}{m^2c^2}} = mc^2\left(1 + \frac{\mathbf{p}^2}{2m^2c^2} + O\left(\frac{p^4}{m^4c^4}\right)\right) =$$
$$= mc^2 + \frac{\mathbf{p}^2}{2m} + O\left(\frac{p^4}{m^3c^2}\right)$$

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 $D_\mu = \partial_\mu + \frac{i}{\hbar}eA_\mu$ we get:

$$D^0 = \partial^0 + \frac{i}{\hbar}eA^0$$

$$D^j = \partial^j + \frac{i}{\hbar}eA^j = -\frac{i}{\hbar}(i\hbar\partial^j - eA^j) = -\frac{i}{\hbar}(\mathbf{p} - e\mathbf{A})$$

and for the lower indices:

$$D_0 = \partial_0 + \frac{i}{\hbar}eA_0$$

$$D_j = \partial_j + \frac{i}{\hbar}eA_j = -\frac{i}{\hbar}(i\hbar\partial_j - eA_j) = \frac{i}{\hbar}(i\hbar\partial^j - eA^j) = \frac{i}{\hbar}(\mathbf{p} - e\mathbf{A})$$

4.5.3 Multipole expansion

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|} = \frac{1}{\sqrt{(\mathbf{r}-\mathbf{r}')^2}} = \frac{1}{\sqrt{r^2 - 2\mathbf{r}\cdot\mathbf{r}' + r'^2}} = \frac{1}{r\sqrt{1 - 2\left(\frac{r'}{r}\right)\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}' + \left(\frac{r'}{r}\right)^2}} =$$

$$= \frac{1}{r}\sum_{l=0}^{\infty}\left(\frac{r'}{r}\right)^l P_l(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}') =$$

$$= \frac{1}{r}\left(P_0(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}') + P_1(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}')\frac{r'}{r} + P_2(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}')\left(\frac{r'}{r}\right)^2 + O\left(\frac{r'^3}{r^3}\right)\right) =$$

$$= \frac{1}{r}\left(1 + \hat{\mathbf{r}}\cdot\hat{\mathbf{r}}'\frac{r'}{r} + \tfrac{1}{2}\left(3(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}')^2 - 1\right)\left(\frac{r'}{r}\right)^2 + O\left(\frac{r'^3}{r^3}\right)\right) =$$

$$= \frac{1}{r} + \frac{\mathbf{r}\cdot\mathbf{r}'}{r^3} + \frac{3(\mathbf{r}\cdot\mathbf{r}')^2 - r^2r'^2}{2r^5} + O\left(\frac{r'^3}{r^4}\right)$$

We can also use the formula:

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

and rewrite the expansion using spherical harmonics:

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|} = \frac{1}{r}\sum_{l=0}^{\infty}\left(\frac{r'}{r}\right)^l P_l(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}') =$$

$$= \frac{1}{r}\sum_{l,m}\left(\frac{r'}{r}\right)^l \frac{2l+1}{4\pi}\langle\hat{\mathbf{r}}|lm\rangle\langle lm|\hat{\mathbf{r}}'\rangle = \frac{1}{r}\sum_{l,m}\left(\frac{r'}{r}\right)^l \frac{2l+1}{4\pi}Y_{lm}(\hat{\mathbf{r}})Y^*_{lm}(\hat{\mathbf{r}}')$$

4.6 Examples

4.6.1 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

$$\left(-\tfrac{1}{2}\frac{\partial^2}{\partial u^2} - \tfrac{1}{2}\frac{\partial^2}{\partial v^2} + \frac{1}{\sqrt{2}|u|} + \tfrac{1}{2}\omega^2 u^2 + \tfrac{1}{2}\omega^2 v^2\right)\Psi(u, v) = E\Psi(u, v)$$

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:

$$\Psi(u, v) = f(u)g(v)$$
$$\left(-\tfrac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}u^2} + \frac{1}{\sqrt{2}|u|} + \tfrac{1}{2}\omega^2 u^2\right) f_k(u) = \epsilon_k f_k(u)$$
$$\left(-\tfrac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}v^2} + \tfrac{1}{2}\omega^2 v^2\right) g_l(v) = \epsilon_l g_l(v)$$
$$E_{kl} = \epsilon_k + \epsilon_l$$

the solution of the second equation is:

$$g_l(v) = \frac{1}{\sqrt{2^l l!}}\left(\frac{\omega}{\pi}\right)^{\frac{1}{4}} e^{-\frac{\omega v^2}{2}} H_l(\sqrt{\omega}v)$$
$$\epsilon_l = \omega(l + \tfrac{1}{2}) \quad \text{for } l = 0, 1, 2, \dots$$

where $H_n(x)$ are the Hermite polynomials:

$$H_n(x) = (-1)^n e^{x^2}\frac{\mathrm{d}^n}{\mathrm{d}x^n}e^{-x^2}$$

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 $u > 0$):

$$V(u) = \frac{1}{\sqrt{2}|u|} + \tfrac{1}{2}\omega^2 u^2 = \left(2^{-\frac{1}{3}} + 2^{-\frac{4}{3}}\right)\omega^{\frac{2}{3}} + \frac{3}{2}\omega^2 (u-u_0)^2 + O\left((u-u_0)^3\right)$$

$$u_0 = 2^{-\frac{1}{6}}\omega^{-\frac{2}{3}}$$

So the first few states can be approximated by the harmonic oscillator solution with frequency $\sqrt{3}\omega$:

$$f_k(u) = \frac{1}{\sqrt{2^k k!}}\left(\frac{\sqrt{3}\omega}{\pi}\right)^{\frac{1}{4}} e^{-\frac{\sqrt{3}\omega(u-u_0)^2}{2}} H_k(3^{\frac{1}{4}}\sqrt{\omega}(u-u_0))$$

$$\epsilon_k = \left(2^{-\frac{1}{3}} + 2^{-\frac{4}{3}}\right)\omega^{\frac{2}{3}} + \sqrt{3}\omega(k+\tfrac{1}{2}) \quad \text{for } k = 0, 1, 2, \dots$$

The final solution is then:

$$\Psi_{kl}(u,v) = f_k(u) g_l(v) =$$

$$= \frac{1}{\sqrt{2^k k!}}\left(\frac{\sqrt{3}\omega}{\pi}\right)^{\frac{1}{4}} e^{-\frac{\sqrt{3}\omega(u-u_0)^2}{2}} H_k(3^{\frac{1}{4}}\sqrt{\omega}(u-u_0)) \frac{1}{\sqrt{2^l l!}}\left(\frac{\omega}{\pi}\right)^{\frac{1}{4}} e^{-\frac{\omega v^2}{2}} H_l(\sqrt{\omega} v)$$

$$E_{kl} = \epsilon_k + \epsilon_l = \left(2^{-\frac{1}{3}} + 2^{-\frac{4}{3}}\right)\omega^{\frac{2}{3}} + \sqrt{3}\omega(k+\tfrac{1}{2}) + \omega(l+\tfrac{1}{2})$$

4.6.2 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 Schröodinger 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

$$A_{ij} = \int_a^b \frac{1}{2}\phi_i'(x)\phi_j'(x) + V(x)\phi_i(x)\phi_j(x)\,\mathrm{d}x$$

$$B_{ij} = \int_a^b \phi_i(x)\phi_j(x)\,\mathrm{d}x$$

4.6.3 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:

$$-\frac{1}{2}\nabla^2(RY) + VRY = ERY$$

$$-\frac{1}{2}R''Y - \frac{1}{\rho}R'Y + \frac{L^2RY}{2\rho^2} + VRY = ERY$$

Using the fact that $L^2Y = l(l+1)Y$ 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:

$$\int_\Omega \frac{1}{2}\nabla\psi(\mathbf{x})\nabla v(\mathbf{x}) + V(r)\psi(\mathbf{x})v(\mathbf{x})\,\mathrm{d}^3x = E\int_\Omega \psi(\mathbf{x})v(\mathbf{x})\,\mathrm{d}^3x$$

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:

$$\left(\int_0^a \frac{1}{2}r^2\phi_{il}'(r)\phi_{jl}'(r) + \frac{1}{2}X + \frac{l(l+1)}{2}\phi_{il}(r)\phi_{jl}(r) + r^2V(r)\phi_{il}(r)\phi_{jl}(r)\,\mathrm{d}r\right)y_{jlm} =$$

$$= E\int_0^a r^2\phi_{il}(r)\phi_{jl}(r)\,\mathrm{d}r\,y_{jlm}$$

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:

$$-\frac{1}{2}R''(r) - \frac{1}{r}R'(r) + \frac{l(l+1)}{2r^2}R(r) = ER(r)$$

The general solution is a linear combination of the spherical Bessel functions $j_l(kr)$ and $n_l(kr)$, whose 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:

$$R_l(kr) = A_l\frac{1}{kr}\sin\left(kr - \frac{l\pi}{2}\right) + B_l\frac{1}{kr}\cos\left(kr - \frac{l\pi}{2}\right) =$$
$$= \sqrt{A_l^2 + B_l^2}\frac{1}{kr}\sin\left(kr - \frac{l\pi}{2} + \delta_l\right) = C_l\frac{1}{kr}\sin\left(kr - \frac{l\pi}{2} + \delta_l\right)$$

where

$$\delta_l = \mathrm{atan2}(B_l, A_l)$$
$$C_l = \sqrt{A_l^2 + B_l^2}$$

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

$$-\frac{1}{2}R''(r) - \frac{1}{r}R'(r) + \left(V + \frac{l(l+1)}{2r^2}\right)R(r) = ER(r)$$

and then fit it to the above asymptotic solution for V=0. We require that the value and the slope must be continuous. In particular, we take the logarithmic derivative ($(\log u)' = \frac{u'}{u}$) at the point $r = a$:

$$\gamma_l \equiv \left.\frac{\mathrm{d}}{\mathrm{d}r}\log u\right|_{r=a} = \left.\frac{\mathrm{d}}{\mathrm{d}r}\log R_l(kr)\right|_{r=a}$$

expressing $R_l(kr)$ using δ_l and solving for it we get:

$$\tan\delta_l = \frac{kj_l'(ka) - \gamma_l j_l(ka)}{kn_l'(ka) - \gamma_l n_l(ka)}$$

Now we can use these δ_l in the formula for the total cross section.

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.

CHAPTER

FIVE

INDICES AND TABLES

- *Index*
- *Search Page*

INDEX

L

M

N

O

P

Q

R

S

T

V

W

Y