MA4447: The Standard Model of Elementary Particle Physics

David Whyte `dawhyte@tcd.ie`

Updated November 20, 2011

1 Natural units

We will use natural units: $\hbar = c = 1$. This means that $[\mathbf{p}] = [m] = [T]^{-1} = [L]^{-1} = [E] =$ MeV. To convert from standard units to natural units, remember that $\hbar c \approx 200$ MeV fm.

For example, to convert from seconds to natural units:

$$\begin{aligned}
\tau = 10^{-23}\ \mathrm{s} &= \frac{10^{-23} c\ \mathrm{s}}{c} \\
&= \frac{3\ \mathrm{fm}}{c} \\
&\approx \frac{3\hbar c}{c}\frac{1}{200\ \mathrm{MeV}} \\
\tau &= 1.5 \times 10^{-2}\ \mathrm{MeV}^{-1}
\end{aligned}$$

If τ is the lifetime of a strongly decaying particle, the decay width $\Gamma = \frac{1}{\tau} \approx 70$ MeV.

2 Klein–Gordon equation

2.1 Derivation

In non-relativistic quantum mechanics, the Schrödinger equation is based on an energy:

$$E = \frac{\mathbf{p}^2}{2m}, \qquad E \to i\frac{\partial}{\partial t}, \quad \mathbf{p} \to -i\boldsymbol{\nabla}$$

We keep this recipe, but use the relativistic energy:

$$E = \pm\sqrt{\mathbf{p}^2 + m^2}$$

We rewrite this to avoid the square root:

$$E^2 = \mathbf{p}^2 + m^2.$$

and note that:

$$E^2 \to -\frac{\partial^2}{\partial t^2}, \quad \mathbf{p}^2 \to -\boldsymbol{\nabla}^2.$$

So we have:

$$-\frac{\partial^2}{\partial t^2}\phi(\mathbf{x},t) = (-\boldsymbol{\nabla}^2 + m^2)\phi(\mathbf{x},t)$$

or, noting that the d'Alembertian operator $\Box = \partial_\mu\partial^\mu = \frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2$,

$$(\Box + m^2)\phi(\mathbf{x},t) = 0,$$

the **Klein–Gordon equation.**

2.2 Example

We will test the equation on the plane wave solution:

$$\phi(\mathbf{x},t) = Ne^{-iEt+i\mathbf{p}\cdot\mathbf{x}} = Ne^{-ipx},$$

since $px \equiv p_\mu x^\mu = p^0x^0 - \mathbf{p}\cdot\mathbf{x} = Et - \mathbf{p}\cdot\mathbf{x}$.

We calculate derivatives:

$$\partial_\mu e^{-ipx} = \partial_\mu(-ipx)e^{-ipx} = -ip_\mu e^{-ipx}.$$

$$\partial^\mu\partial_\mu e^{-ipx} = \Box e^{-ipx} = (-ip^\mu)(-ip_\mu)e^{-ipx} = -p^2e^{-ipx}$$

Note that $p^2 \equiv p_\mu p^\mu = E^2 - \mathbf{p}^2$, so:

$$\begin{aligned}(\Box + m^2)Ne^{-ipx} &= N(-p^2 + m^2)e^{-ipx} \\ &= \underbrace{(-E^2 + \mathbf{p}^2 + m^2)}_{0}Ne^{-ipx} \\ &= 0, \text{ as expected. } \checkmark\end{aligned}$$

3 Probability current

3.1 Derivation

Recall that we used $E = \pm\sqrt{\mathbf{p}^2 + m^2}$, meaning that for any $\mathbf{p}$ there are two solutions, $E < 0$ and $E > 0$.

We use the K–G equation for ϕ, and its complex conjugate ϕ^*:

$$i\phi^*\underbrace{(\partial_\mu\partial^\mu + m^2)\phi}_{0} - i\phi\underbrace{(\partial_\mu\partial^\mu + m^2)\phi^*}_{0} = 0,$$

and rewrite the LHS:

$$\begin{aligned}
0 &= m^2 i \underbrace{(\phi^*\phi - \phi\phi^*)}_{0} + i\phi^*\partial_\mu\partial^\mu\phi - i\phi\partial_\mu\partial^\mu\phi^* \\
&= i\partial_\mu(\phi^*\partial^\mu\phi - \phi\partial^\mu\phi^*) \underbrace{- i(\partial_\mu\phi^*)(\partial^\mu\phi) + i(\partial_\mu\phi)(\partial^\mu\phi^*)}_{0} \\
&= \partial_\mu\left[i(\phi^*\partial^\mu\phi - \phi\partial^\mu\phi^*)\right] \\
&\equiv \partial_\mu j^\mu.
\end{aligned}$$

j^μ can be expressed as $(\rho, \mathbf{j})$, where ρ is probability density and $\mathbf{j}$ is probability current:

$$\begin{aligned}
\rho &= i(\phi^*\frac{\partial\phi}{\partial t} - \phi\frac{\partial\phi^*}{\partial t}) \\
\mathbf{j} &= -i(\phi^*\boldsymbol{\nabla}\phi - \phi\boldsymbol{\nabla}\phi^*)
\end{aligned}$$

$\partial_\mu j^\mu = 0$ can be written in 3D as the continuity equation:

$$\frac{\partial\rho}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{j} = 0.$$

3.2 Example

We evaluate ρ for the plane wave $\phi = Ne^{-ipx}$:

$$\begin{aligned}
\rho &= i|N|^2(e^{ipx}(-iE)e^{-ipx} - e^{-ipx}iEe^{ipx}) \\
&= |N|^2 i(-2iE) \\
&= 2E|N|^2
\end{aligned}$$

If $E < 0$ this gives a negative probability density. This is a problem!

4 Dirac equation

4.1 Derivation

The K–G equation is second-order in space and time (because of $\Box$). Dirac tried to obtain an equation which was first-order. The equation he proposed was:

$$i\frac{\partial}{\partial t}\psi = (-i\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} + \beta m)\psi,$$

where α_1, α_2, α_3 and β are to be determined. This is the **Dirac equation.**

In order to find $\boldsymbol{\alpha}$ and β we require:

- the relativistic energy-momentum equation

- Lorentz covariance of the equation (discussed later)

We also want ϕ to satisfy the K–G equation. We square both sides of the Dirac equation:

$$\left(i\frac{\partial}{\partial t}\right)^2 \psi = (-i\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}+\beta m)(-i\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}+\beta m)\psi.$$

and carefully expand, bearing in mind that α_i and β are not necessarily scalar:

$$-\frac{\partial^2}{\partial t^2}\psi = -\sum_i \alpha_i^2 \frac{\partial^2\psi}{(\partial x^i)^2} - \sum_{i<j}(\alpha_i\alpha_j+\alpha_j\alpha_i)\frac{\partial^2\psi}{\partial x^i\partial x^j} + \beta^2 m^2\psi - im\sum_i(\alpha_i\beta+\beta\alpha_i)\frac{\partial\psi}{\partial x^i},$$

$i, j = 1,2,3.$

In order for this to match the K–G equation, the RHS must equal $-\boldsymbol{\nabla}^2\psi + m^2\psi$. And for this, we have the requirements:

1. $\alpha_i^2 = 1, \quad i = 1,2,3.$
2. $\beta^2 = 1.$
3. $\alpha_i\alpha_j + \alpha_j\alpha_i = 0, \quad i \neq j.$
4. $\alpha_i\beta + \beta\alpha_i = 0, \quad i = 1,2,3.$

4.2 Dirac spinors

Requirements 3 and 4 mean that α_i and β cannot be scalars, so they must be matrices. Furthermore, we require that they are Hermitian in order for the Hamiltonian to be Hermitian.

A tempting choice for α_i is the Pauli spin matrices σ_i, since we know that they satisfy requirements 1 and 3. However no nonzero β exists which anticommutes with σ_i for requirement 4.

Take requirement 4:

$$\begin{aligned}
\alpha_i\beta + \beta\alpha_i &= 0 \\
\Rightarrow \alpha_i &= -\beta\alpha_i\beta \\
\mathrm{tr}(\alpha_i) &= -\mathrm{tr}(\beta\alpha_i\beta) \\
&= -\mathrm{tr}(\beta^2\alpha_i) \\
&= -\mathrm{tr}(\alpha_i),
\end{aligned}$$

so $\mathrm{tr}(\alpha_i) = 0$. By a similar argument, $\mathrm{tr}(\beta) = 0$.

So α_i and β are Hermitian matrices with square 1 and trace 0. This tells us that they have:

- real eigenvalues
- eigenvalues $= \pm 1$

- equal number of $+1$ and -1 eigenvalues $\Rightarrow$ even dimension.

There is no set of 2×2 matrices which satisfy these conditions, so we try 4×4. One possible choice is:

$$\alpha_i = \left(\begin{array}{c|c} 0 & \sigma_i \\ \hline \sigma_i & 0 \end{array}\right), \quad \beta = \left(\begin{array}{c|c} \mathbb{1}_2 & 0 \\ \hline 0 & -\mathbb{1}_2 \end{array}\right),$$

where σ_i are the Pauli spin matrices and $\mathbb{1}_2$ is the 2×2 identity matrix.

This choice is *not* unique. We can construct $\alpha'_i = U\alpha_i U^{-1}$ and $\beta' = U\beta U^{-1}$ such that:

$$\begin{aligned}(\alpha'_i)^2 &= U\alpha U^{-1}U\alpha U^{-1} \\ &= U\alpha^2 U^{-1} \\ &= UU^{-1} \\ &= \mathbb{1}_2\end{aligned}$$

and similarly for β'. We want $(\alpha'_i)^\dagger = \alpha'_i$, so:

$$\begin{aligned}(U^{-1})^\dagger \alpha_i^\dagger U^\dagger &= \alpha'_i \\ (U^{-1})^\dagger \alpha_i U^\dagger &= \alpha'_i,\end{aligned}$$

so $U^\dagger = U^{-1}$, i.e. U is unitary.

Hence ψ is a column vector with four components, called a **Dirac spinor:**

$$\psi = \begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\\ \psi_4\end{pmatrix}, \quad \psi^\dagger = (\psi_1^*\ \psi_2^*\ \psi_3^*\ \psi_4^*).$$

4.3 Probability current

Consider again the Dirac equation:

$$i\partial_t\psi = (-i\boldsymbol{\alpha}\cdot\nabla + \beta m)\psi \tag{1}$$

Pre-multiply (1) by $\psi^\dagger$:

$$\psi^\dagger\, i\partial_t\psi = \psi^\dagger(-i\boldsymbol{\alpha}\cdot\nabla + \beta m)\psi. \tag{2}$$

Take the complex conjugate of (1):

$$\begin{aligned}-i\partial_t\psi^\dagger &= \psi^\dagger(i\boldsymbol{\alpha}\cdot\overleftarrow{\nabla} + \beta m) \\ \Rightarrow -i(\partial_t\psi^\dagger)\psi &= \psi^\dagger(i\boldsymbol{\alpha}\cdot\overleftarrow{\nabla} + \beta m)\psi,\end{aligned} \tag{3}$$

where the $\leftarrow$ means that the ∇ operates to the left, i.e. on $\psi^\dagger$ here.

Subtracting (3) from (2) gives:

$$i\partial_t(\psi^\dagger\psi) = -i\psi^\dagger(\boldsymbol{\alpha}\cdot\nabla + \boldsymbol{\alpha}\cdot\overleftarrow{\nabla})\psi \tag{4}$$

Define:

$$\rho = \psi^\dagger\psi = \sum_{\sigma=1}^{4} \psi_\sigma^*\psi_\sigma = \sum |\psi_\sigma|^2 \geqslant 0.$$

This means no more negative probability densities!

Defining $j^k = \psi^\dagger \alpha_k \psi$ allows (4) to be written:

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{j} = 0,$$

or in four-vector form, $\partial_\mu j^\mu = 0$, where $j^\mu = (\rho, \mathbf{j})$ as before.

4.4 Non-relativistic correspondence

Take an electron at rest: $\mathbf{p} = 0$, so:

$$i\frac{\partial\psi}{\partial t} = \beta m\psi.$$

This has four solutions:

$$\psi^1 = e^{imt}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \quad \psi^2 = e^{-imt}\begin{pmatrix}0\\1\\0\\0\end{pmatrix} \qquad \text{(positive energy)}$$

$$\psi^3 = e^{imt}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \quad \psi^4 = e^{-imt}\begin{pmatrix}0\\0\\0\\1\end{pmatrix} \qquad \text{(negative energy)}$$

We will focus on the positive energy solutions, and work out the connections to Pauli theory in non-relativistic QM by introducing an EM potential:

$$p^\mu \to p^\mu - eA^\mu, \qquad A^\mu = (\phi, \mathbf{A})$$

Substitute this into the Dirac equation:

$$(i\partial_t - e\phi)\psi = (\boldsymbol{\alpha}\cdot(-i\nabla - e\mathbf{A}) + \beta m)\psi. \tag{5}$$

Introduce the kinetic momentum operator $\boldsymbol{\pi} = -i\nabla - e\mathbf{A}$, and let:

$$\psi = \begin{pmatrix}\tilde\varphi\\ \tilde\chi\end{pmatrix}$$

where $\tilde\varphi$ and $\tilde\chi$ are 2-component objects. Now rewrite (5) using $\boldsymbol{\alpha} = \left(\begin{smallmatrix}0 & \boldsymbol{\sigma}\\ \boldsymbol{\sigma} & 0\end{smallmatrix}\right)$, $\beta = \left(\begin{smallmatrix}\mathbb{1} & 0\\ 0 & -\mathbb{1}\end{smallmatrix}\right)$ and $\boldsymbol{\pi} = -i\nabla - e\mathbf{A}$:

$$i\partial_t\begin{pmatrix}\tilde\varphi\\ \tilde\chi\end{pmatrix} = \begin{pmatrix}\boldsymbol{\sigma}\cdot\boldsymbol{\pi}\tilde\chi\\ \boldsymbol{\sigma}\cdot\boldsymbol{\pi}\tilde\varphi\end{pmatrix} + e\phi\begin{pmatrix}\tilde\varphi\\ \tilde\chi\end{pmatrix} + m\begin{pmatrix}\tilde\varphi\\ -\tilde\chi\end{pmatrix} \tag{6}$$

Remember we are looking at the nonrelativistic limit, so all energies $\ll m$. We are considering only the positive energy solutions, so we can write:

$$\begin{pmatrix}\tilde{\varphi}\\ \tilde{\chi}\end{pmatrix} = e^{-imt}\begin{pmatrix}\varphi\\ \chi\end{pmatrix}$$

By assumption, φ and χ are slowly oscillating functions, so (6) gives:

$$i\partial_t\begin{pmatrix}\varphi\\ \chi\end{pmatrix} = \boldsymbol{\sigma}\cdot\boldsymbol{\pi}\begin{pmatrix}\chi\\ \varphi\end{pmatrix} + e\phi\begin{pmatrix}\varphi\\ \chi\end{pmatrix} - 2m\begin{pmatrix}0\\ \chi\end{pmatrix}. \tag{7}$$

Since $\|\partial_t\chi\|, \|e\phi\chi\| \ll \|2m\chi\|$,

$$\boldsymbol{\sigma}\cdot\boldsymbol{\pi} - 2m\chi \approx 0 \Longleftrightarrow \chi = \underbrace{\frac{\boldsymbol{\sigma}\cdot\boldsymbol{\pi}}{2m}}_{O(v/c)}\varphi$$

so χ are the small components of the 4-spinor relative to φ. Insert this expression for χ into (7) and the top line gives:

$$i\partial_t\varphi = \left(\frac{(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})^2}{2m} + e\phi\right)\varphi. \tag{8}$$

We want to simplify $(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})^2$.

$$\begin{aligned}(\boldsymbol{\sigma}\cdot\mathbf{a})(\boldsymbol{\sigma}\cdot\mathbf{b}) &= \sigma_i a_i \sigma_k b_k\\ &= a_i b_k \sigma_i\sigma_k\\ &= a_i b_k(\delta_{ik}\mathbb{1}_2 + i\epsilon_{ikl}\sigma_l)\\ &= \mathbf{a}\cdot\mathbf{b}\,\mathbb{1}_2 + i\sigma_l\epsilon_{lik}a_i b_k\\ &= \mathbf{a}\cdot\mathbf{b}\,\mathbb{1}_2 + i\boldsymbol{\sigma}\cdot(\mathbf{a}\times\mathbf{b})\end{aligned}$$

So $(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})^2 = (\boldsymbol{\sigma}\cdot\boldsymbol{\pi})\cdot(\boldsymbol{\sigma}\cdot\boldsymbol{\pi}) = \boldsymbol{\pi}^2 + i\boldsymbol{\sigma}\cdot(\boldsymbol{\pi}\times\boldsymbol{\pi})$. Note that $\boldsymbol{\pi}\times\boldsymbol{\pi} \neq 0$, since $\boldsymbol{\pi}$ is an operator.

$$\begin{aligned}(\boldsymbol{\pi}\times\boldsymbol{\pi})_i &= \left[(-i\boldsymbol{\nabla} - e\mathbf{A})\times(-i\boldsymbol{\nabla} - e\mathbf{A})\right]_i\\ &= \epsilon_{ijk}(-i\boldsymbol{\nabla}_j - e\mathbf{A}_j)(-i\boldsymbol{\nabla}_k - e\mathbf{A}_k)\\ &= \epsilon_{ijkl}(-\boldsymbol{\nabla}_j\boldsymbol{\nabla}_k + ie\mathbf{A}_j\boldsymbol{\nabla}_k + ie\boldsymbol{\nabla}_j\mathbf{A}_k + ie\mathbf{A}_k\boldsymbol{\nabla}_j + e^2\mathbf{A}_j\mathbf{A}_k)\end{aligned}$$

Since $-\boldsymbol{\nabla}_j\boldsymbol{\nabla}_k$, $e^2\mathbf{A}_j\mathbf{A}_k$ and $(ie\mathbf{A}_j\boldsymbol{\nabla}_k + ie\mathbf{A}_k\boldsymbol{\nabla}_j)$ are symmetric in $j \leftrightarrow k$ they vanish when summed over j and k due to the antisymmetric ϵ_{ijkl}, leaving:

$$\begin{aligned}(\boldsymbol{\pi}\times\boldsymbol{\pi})_i &= ie\epsilon_{ijk}(\boldsymbol{\nabla}_j\mathbf{A}_k)\\ &= ie(\boldsymbol{\nabla}\times\mathbf{A})_i\\ &= ie\mathbf{B}_i\\ \Rightarrow (\boldsymbol{\sigma}\cdot\boldsymbol{\pi})^2 &= \boldsymbol{\pi}^2 - e\boldsymbol{\sigma}\cdot\mathbf{B}.\end{aligned}$$

Using this expression in (8) gives the **non-relativistic Pauli equation:**

$$i\partial_t\varphi = \left(\frac{(-i\boldsymbol{\nabla} - e\mathbf{A})^2}{2m} - \frac{e}{2m}\boldsymbol{\sigma}\cdot\mathbf{B} + e\phi\right)\varphi$$

The equation contains a term which reflects the electron magnetic moment $\boldsymbol{\mu}$:

$$\boldsymbol{\mu} = -\frac{e}{2m}\boldsymbol{\sigma} = -g\frac{e}{2m}\mathbf{S}.$$

The electron g-factor has been found experimentally to be ~ 2.0023. This deviation from 2 can be explained by QFT.

5 Covariance of Dirac equation

5.1 Lorentz group

Lorentz transformations are denoted by $x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu$. The Lorentz group consists of those Λ under which the 4D scalar product,

$$x \cdot y = x_\mu y^\mu = x^0 y^0 - \mathbf{x}\cdot\mathbf{y} = x^\mu g_{\mu\nu} y^\nu,$$

is invariant. We will use $g = \mathrm{diag}(1, -1, -1, -1)$.

What is the implication of invariance?

$$\begin{aligned} x'^\mu y'_\mu &= \Lambda^\mu{}_\nu x^\nu g_{\mu\rho} y'^\rho \\ &= \Lambda^\mu{}_\nu x^\nu g_{\mu\rho} \Lambda^\rho{}_\sigma y^\sigma \\ &= x^\nu y^\sigma \left[\Lambda^\mu{}_\nu g_{\mu\rho} \Lambda^\rho{}_\sigma\right] \end{aligned}$$

In order for this to equal $x^\mu y_\mu$, we require $\left[\Lambda^\mu{}_\nu g_{\mu\rho} \Lambda^\rho{}_\sigma\right] = g_{\nu\sigma}$. This can be written:

$$(\Lambda^{\mathrm{T}})^\nu{}_\mu\, g_{\mu\rho} \Lambda^\rho{}_\sigma = g_{\nu\sigma},$$

(where Λ^{T} is the transpose of Λ), or, in matrix form,

$$\Lambda^{\mathrm{T}} g \Lambda = g. \tag{9}$$

The Lorentz group is hence made up of all Λ which satisfy (9).

5.2 Observations

- Look at $\det\Lambda$:

  $$\begin{aligned} \det(\Lambda^{\mathrm{T}} g \Lambda) &= \det g \\ \det\Lambda^{\mathrm{T}} \det\Lambda &= 1 \\ \Rightarrow \det\Lambda &= \pm 1, \end{aligned}$$

  since Λ has only real entries.

- Look at $g^0{}_0$:

$$\begin{aligned}(\Lambda^{\mathrm{T}} g \Lambda)^0{}_0 &= g^0{}_0 = 1\\ \Rightarrow \Lambda^\mu{}_0 g_{\mu\rho} \Lambda^\rho{}_0 &= 1\end{aligned}$$

  Expand the sum, noting that $g_{\alpha\beta} = 0$ for $\alpha \neq \beta$:

$$\begin{aligned}\Lambda^0{}_0 g_{00} \Lambda^0{}_0 + \Lambda^1{}_0 g_{11} \Lambda^1{}_0 + \Lambda^2{}_0 g_{22} \Lambda^2{}_0 + \Lambda^3{}_0 g_{33} \Lambda^3{}_0 = 1\\ \Longrightarrow (\Lambda^0{}_0)^2 - (\Lambda^1{}_0)^2 - \Lambda^2{}_0 - \Lambda^3{}_0 = 1.\end{aligned}$$

  This means that $(\Lambda^0{}_0)^2 \geqslant 1$, i.e. either $\Lambda^0{}_0 \geqslant 1$ or $\Lambda^0{}_0 \leqslant -1$.

The two properties $\det \Lambda = \pm 1$ and $\Lambda^0{}_0 \geqslant 1$ or $\leqslant -1$ define four disconnected components of the Lorentz group:

	$\det \Lambda$	$\operatorname{sign} \Lambda^0{}_0$	contains
$L_+^\uparrow$	$+1$	$+1$	$\Lambda = \mathbb{1}$
$L_-^\uparrow$	-1	$+1$	$\Lambda = I_s$
$L_+^\downarrow$	$+1$	-1	$\Lambda = I_{st}$
$L_-^\downarrow$	-1	-1	$\Lambda = I_t$

I_s is the parity transformation $(x^0, \mathbf{x}) \to (x^0, -\mathbf{x})$, I_t is the time reversal transformation $(x^0, \mathbf{x}) \to (-x^0, \mathbf{x})$ and $I_{st} = I_s \circ I_t$.

$L_\pm^\uparrow$ are called 'orthochronous' transformations since they do not change the sign of x^0, i.e. the direction of time. Usually only $L_+^\uparrow$ is discussed, since the others can be obtained using I_s and I_t.

5.3 Transforming wavefunctions

We want to find the connection between a wavefunction as seen by observers O and O', where $x'^\mu = \Lambda^\mu{}_\nu x^\nu$. How is an equation for ψ' related to an equation for ψ?

Consider the Klein–Gordon equation:

$$\begin{aligned}O: &\quad (\partial_\mu \partial^\mu + m^2)\phi(x) = 0.\\ O': &\quad (\partial'_\mu \partial'^\mu + m^2)\phi'(x') = 0.\end{aligned}$$

Note that m^2 is Lorentz invariant.

We use the Ansatz:

$$\begin{aligned}\phi'(x') &= f(\Lambda)\phi(x)\\ &= f(\Lambda)\phi(\Lambda^{-1}x')\end{aligned}$$

It turns out that $\partial'_\mu \partial'^\mu = \partial_\mu \partial^\mu$, so we have:

$$(\partial_\mu \partial^\mu + m^2)\phi'(x') = (\partial_\mu \partial^\mu + m^2) f(\Lambda)\phi(\Lambda^{-1}x').$$

$f(\Lambda) = 1$ is a solution, so ϕ is a scalar field.

6 Gamma matrices

6.1 Definition

Define $\gamma^0 = \beta$ and $\gamma^k = \beta\alpha_k$, $k = 1,2,3$. This gives the property:

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu \equiv \{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}. \tag{10}$$

Verify:

- $\mu = \nu = 0$:
$$\gamma^0\gamma^0 + \gamma^0\gamma^0 = 2\beta^2 = 2 = 2g^{00} \quad \checkmark$$

- $\mu = 0, \nu = k$:
$$\begin{aligned}\gamma^0\gamma^k + \gamma^k\gamma^0 &= \beta\beta\alpha_k + \beta\alpha_k\beta \\ &= \alpha_k + \beta(\beta\alpha_k) \\ &= \alpha_k - \alpha_k = 0 = 2g^{0k} \quad \checkmark\end{aligned}$$

- $\mu = \nu = k$:
$$\begin{aligned}\gamma^k\gamma^k + \gamma^k\gamma^k &= 2\beta\alpha_k\beta\alpha_k \\ &= 2(-\alpha_k\beta)\beta\alpha_k \\ &= -2\alpha_k\beta^2\alpha_k \\ &= -2\alpha_k^2 = -2 = -2g^{kk} \quad \checkmark\end{aligned}$$

- $\mu = k$, $\nu = k \neq l$:
$$\begin{aligned}\gamma^k\gamma^l + \gamma^l\gamma^k &= \beta\alpha_k\beta\alpha_l + \beta\alpha_l\beta\alpha_k \\ &= -\beta^2\alpha_k\alpha_l - \beta^2\alpha_l\alpha_k \\ &= -(\alpha_k\alpha_l + \alpha_l\alpha_k) = 0 = 2g^{kl} \quad \checkmark\end{aligned}$$

6.2 Properties

α_k and β are Hermitian. $\gamma^0 = \beta$ so obviously $(\gamma^0)^\dagger = \gamma^0$, and $(\gamma^0)^2 = \beta^2 = 1$.

Since $\gamma^k = \beta\alpha^k$,

$$(\gamma^k)^\dagger = \alpha_k^\dagger\beta^\dagger = \alpha_k^\beta = -\beta\alpha_k = -\gamma^k,$$

i.e. γ^k are anti-Hermitian. $(\gamma^k)^2 = \beta\alpha_k\beta\alpha_k = -\beta^2\alpha_k^2 = -1$.

A useful form is:

$$(\gamma^\mu)^\dagger = \left\{\begin{array}{cc} \gamma^\mu & \text{if } \mu = 0 \\ -\gamma^\mu & \text{if } \mu = k \end{array}\right\} = \gamma^0\gamma^\mu\gamma^0. \tag{11}$$

6.3 Explicit representation

Using our earlier choice of α_k and β gives:

$$\gamma_0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \quad \gamma_k = \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix}$$

7 Back to the Dirac equation

7.1 Alternate expression

We want to rewrite the Dirac equation in terms of γ matrices.

$$\begin{aligned}
\left(i\frac{\partial}{\partial t} + i\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} - \beta m\right)\psi &= 0 \\
\left(i\beta\frac{\partial}{\partial t} + i\beta\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} - m\right)\psi &= 0 \\
\left(i\gamma^0\partial_0 + i\boldsymbol{\gamma}\cdot\boldsymbol{\nabla} - m\right)\psi &= 0 \\
(i\gamma^\mu\partial_\mu - m)\psi &= 0.
\end{aligned}$$

This can be written in **Feynman slash notation** as $(i\not{\partial} - m)\psi = 0$ where $\not{B} \equiv \gamma^\mu B_\mu$.

7.2 Adjoint Dirac equation

$$\begin{aligned}
-i\frac{\partial\psi^\dagger}{\partial t} &= \psi^\dagger(i\boldsymbol{\alpha}\cdot\overleftarrow{\boldsymbol{\nabla}} + \beta m) \\
\Rightarrow \psi^\dagger\left(i\frac{\overleftarrow{\partial}}{\partial t} + i\boldsymbol{\alpha}\cdot\overleftarrow{\boldsymbol{\nabla}} + \beta m\right) &= 0 \\
\psi^\dagger\beta^2\left(i\frac{\overleftarrow{\partial}}{\partial t} + i\boldsymbol{\alpha}\cdot\overleftarrow{\boldsymbol{\nabla}} + \beta m\right) &= 0
\end{aligned}$$

Now define $\overline{\psi} = \psi^\dagger\beta = \psi^\dagger\gamma^0$. So:

$$\begin{aligned}
\overline{\psi}\left(\beta i\frac{\overleftarrow{\partial}}{\partial t} + i\beta\boldsymbol{\alpha}\cdot\overleftarrow{\boldsymbol{\nabla}} + \beta^2 m\right) &= 0 \\
\overline{\psi}\left(\gamma^0 i\overleftarrow{\partial}_0 + i\boldsymbol{\gamma}\cdot\overleftarrow{\boldsymbol{\nabla}} + m\right) &= 0 \\
\overline{\psi}(i\gamma^\mu\overleftarrow{\partial}_\mu + m) &= 0 \\
\overline{\psi}(i\overleftarrow{\not{\partial}} + m) &= 0,
\end{aligned}$$

the **adjoint Dirac equation.**

8 Lorentz transformations

8.1 Formulation

Again we have observers O and O' where $x'^\mu = \Lambda^\mu{}_\nu x^\nu$. The Dirac equations for the observers read:

$$(i\gamma^\mu \partial'_\mu - m)\psi'(x') = 0 \quad \text{and} \quad (i\gamma^\mu \partial_\mu - m)\psi(x) = 0.$$

Use the Ansatz:

$$\psi'(x') = S(\Lambda)\psi(x) = S(\Lambda)\psi(\Lambda^{-1}x')$$

$S(\Lambda)$ must have an inverse to get from O' back to O, so:

$$\begin{aligned}
\psi(x) &= S(\Lambda)^{-1}\psi'(x') = S(\Lambda)^{-1}\psi'(\Lambda x) \\
\psi(x) &= S(\Lambda^{-1})\psi'(x').
\end{aligned}$$

This means that $S(\Lambda)^{-1} = S(\Lambda^{-1})$. Also $S(\Lambda_1\Lambda_2) = S(\Lambda_1)S(\Lambda_2)$, so $S(\Lambda)$ defines a representation of the Lorentz group.

$$\begin{aligned}
(i\not{\partial} - m)\psi(x) &= 0 \\
S(\Lambda)(i\not{\partial} - m)S(\Lambda^{-1})\underbrace{S(\Lambda)\psi(x)}_{\psi'(x')} &= 0 \\
\Rightarrow \big(iS(\Lambda)\gamma^\mu S(\Lambda^{-1})\partial_\mu - m\big)\psi'(x') &= 0
\end{aligned} \tag{12}$$

Note that:

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \underbrace{\frac{\partial x'^\nu}{\partial x^\mu}}_{\Lambda^\nu{}_\mu} \frac{\partial}{\partial x'^\nu}$$

$$\Rightarrow \partial_\mu = \Lambda^\nu{}_\mu \partial'_\nu,$$

allowing (12) to be written:

$$\big(iS(\Lambda)\gamma^\mu S(\Lambda)^{-1}\Lambda^\nu{}_\mu \partial'_\nu - m\big)\psi'(x') = 0.$$

In order for this to match the Dirac equation, we need:

$$\begin{aligned}
S(\Lambda)\gamma^\mu S(\Lambda)^{-1}\Lambda^\nu{}_\mu &= \gamma^\nu, \\
\text{i.e.} \quad \Lambda^\nu{}_\mu \gamma^\mu &= S(\Lambda)^{-1}\gamma^\nu S(\Lambda),
\end{aligned} \tag{13}$$

this is the condition on S which gives a covariant Dirac equation.

We find $S(\Lambda)$ first for an infinitesimal Lorentz transformation $\Lambda \in L_+^\uparrow$. Write:

$$\Lambda^\nu{}_\mu = g^\nu{}_\mu + \Delta\omega^\nu{}_\mu \equiv (1 + \Delta\omega)^\nu{}_\mu.$$

(9) gives (dropping terms of $O(\Delta\omega^2)$ and greater):

$$
\begin{aligned}
\Lambda^{\mathrm{T}} g \Lambda &= g \\
\Rightarrow (1+\Delta\omega)^{\mathrm{T}} g (1+\Delta\omega) &= g \\
g + \Delta\omega^{\mathrm{T}} g + g\Delta\omega + O(\Delta\omega^2) &= g \\
\Delta\omega^{\mathrm{T}} g + g\Delta\omega &= 0 \\
\Delta\omega^{\mu}{}_{\nu} g_{\mu\rho} + g_{\nu\mu}\Delta\omega^{\mu}{}_{\rho} &= 0 \\
\Delta\omega_{\rho\nu} + \Delta\omega_{\nu\rho} &= 0 \\
\Delta\omega_{\rho\nu} &= -\Delta\omega_{\nu\rho},
\end{aligned}
$$

i.e. $\Delta\omega_{\mu\nu}$ is antisymmetric under $\mu \leftrightarrow \nu$. Since it is a 4×4 matrix, this means there are 6 independent parameters. These represent the 6 generators of a Lorentz transformation: 3 rotations and 3 Lorentz boosts.

Now write:

$$
\begin{aligned}
S(\Lambda) &= \mathbb{1} - \tfrac{i}{4}\sigma_{\mu\nu}\Delta\omega^{\mu\nu} + O(\Delta\omega^2) \\
S(\Lambda)^{-1} &= \mathbb{1} + \tfrac{i}{4}\sigma_{\mu\nu}\Delta\omega^{\mu\nu} + O(\Delta\omega^2),
\end{aligned}
\tag{14}
$$

$\sigma_{\mu\nu}$ is an unknown coefficient matrix. We can assume $\sigma_{\mu\nu} = -\sigma_{\nu\mu}$ since any symmetric part would not contribute when contracted with $\Delta\omega^{\mu\nu}$. Insert these into (13):

$$
\begin{aligned}
(\mathbb{1}+\Delta\omega)^{\nu}{}_{\mu}\gamma^{\mu} &= \left(\mathbb{1} + \tfrac{i}{4}\sigma_{\mu\nu}\right)\gamma^{\nu}\left(\mathbb{1} + \tfrac{i}{4}\sigma_{\mu\nu}\Delta\omega^{\mu\nu}\right) \\
\gamma^{\nu} + \Delta\omega^{\nu}{}_{\mu}\gamma^{\mu} &= \gamma^{\nu} + \tfrac{i}{4}\Delta\omega^{\rho\sigma}(\sigma_{\rho\sigma}\gamma^{\nu} - \gamma^{\nu}\sigma_{\rho\sigma}) \\
\Delta\omega^{\nu}{}_{\mu}\gamma^{\mu} &= \tfrac{i}{4}\Delta\omega^{\rho\sigma}\left[\sigma_{\rho\sigma}, \gamma^{\nu}\right].
\end{aligned}
$$

Post-multiply by γ_{ν} to give:

$$
\Delta\omega^{\nu}{}_{\mu}\gamma^{\mu}\gamma_{\nu} = \tfrac{i}{4}\Delta\omega^{\rho\sigma}\left(\sigma_{\rho\sigma}\gamma^{\nu}\gamma_{\nu} - \gamma^{\nu}\sigma_{\rho\sigma}\gamma_{\nu}\right) \tag{15}
$$

What is $\gamma^{\nu}\gamma_{\nu}$?

$$
\begin{aligned}
\gamma^{\nu}\gamma_{\nu} = g_{\nu\mu}\gamma^{\nu}\gamma^{\mu} &= g_{\nu\mu}(2g^{\nu\mu} - \gamma^{\mu}\gamma^{\nu}) \\
&= 2g_{\nu\mu}g^{\nu\mu} - g_{\nu\mu}\gamma^{\mu}\gamma^{\nu} \\
\gamma^{\nu}\gamma_{\nu} &= 2g_{\nu}{}^{\nu} - \gamma^{\nu}\gamma_{\nu} \\
\gamma^{\nu}\gamma_{\nu} &= g_{\nu}{}^{\nu} = 4.
\end{aligned}
\tag{16}
$$

Claim: $\sigma_{\rho\sigma} = \frac{i}{2}\left[\gamma_{\rho}, \gamma_{\sigma}\right]$.

This means that $\gamma^{\nu}\sigma_{\rho\sigma}\gamma_{\nu}$ in (15) will contain a term $\gamma^{\nu}\gamma_{\rho}\gamma_{\sigma}\gamma_{\nu}$.

$$
\begin{aligned}
\gamma^{\nu}\gamma_{\rho}\gamma_{\sigma}\gamma_{\nu} &= \gamma^{\nu}\gamma_{\rho}(2g_{\rho\nu} - \gamma_{\nu}\gamma_{\sigma}) \quad \text{by (10)} \\
&= 2\gamma^{\nu}\gamma_{\rho}g_{\rho\nu} - \gamma^{\nu}\gamma_{\rho}\gamma_{\nu}\gamma_{\sigma} \\
&= 2\gamma_{\sigma}\gamma_{\rho} - \gamma^{\nu}\left(2g_{\rho\nu} - \gamma_{\nu}\gamma_{\rho}\right)\gamma_{\sigma} \\
&= 2\gamma_{\sigma}\gamma_{\rho} - 2\gamma^{\nu}g_{\rho\nu}\gamma_{\sigma} + \gamma^{\nu}\gamma_{\nu}\gamma_{\rho}\gamma_{\sigma} \\
&= 2\gamma_{\sigma}\gamma_{\rho} - 2\gamma_{\rho}\gamma_{\sigma} + 4\gamma_{\rho}\gamma_{\sigma} \\
&= 2\gamma_{\sigma}\gamma_{\rho} + 2\gamma_{\rho}\gamma_{\sigma} \\
\gamma^{\nu}\gamma_{\rho}\gamma_{\sigma}\gamma_{\nu} &= 4g_{\rho\sigma}.
\end{aligned}
$$

$\gamma^\nu\gamma_\rho\gamma_\sigma\gamma_\nu$ is hence symmetric in $\rho \leftrightarrow \sigma$. The term $\gamma^\nu\sigma_{\rho\sigma}\gamma_\nu$ must vanish because of our definition of $\sigma_{\rho\sigma}$. Combining this fact with (16) allows (15) to be written:

$$\Delta\omega^{\nu\mu}\gamma_\mu\gamma_\nu = \tfrac{i}{4}\Delta\omega^{\rho\sigma}4\sigma_{\rho\sigma}.$$

Note that $\gamma_\mu\gamma_\nu = \frac{1}{2}[\gamma_\mu,\gamma_\nu] + \frac{1}{2}\{\gamma_\mu,\gamma_\nu\}$ and that the latter term does not contribute when contracted with the antisymmetric $\Delta\omega^{\rho\sigma}$, so:

$$\begin{aligned}\Delta\omega^{\nu\mu}\tfrac{1}{2}[\gamma_\mu,\gamma_\nu] &= -\tfrac{1}{2}\Delta\omega^{\rho\sigma}[\gamma_\rho,\gamma_\sigma]\\ \Rightarrow -\tfrac{1}{2}\Delta\omega^{\mu\nu}[\gamma_\mu,\gamma_\nu] &= -\tfrac{1}{2}\Delta\omega^{\rho\sigma}[\gamma_\rho,\gamma_\sigma].\end{aligned}$$

Hence our claim is justified.

8.2 Lorentz boost

For an infinitesimal Lorentz transformation we can write:

$$\Delta\omega^\nu{}_\mu = \sum_{n=1}^{6}\Delta\omega^{(n)}(I_n)^\nu{}_\mu$$

where I_n are the generators of the Lorentz transformation. For example:

- boost in x-direction: $I^\nu{}_\mu = \left(\begin{smallmatrix}0 & -1 & 0 & 0\\ -1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{smallmatrix}\right)$
- rotation around z-axis: $I^\nu{}_\mu = \left(\begin{smallmatrix}0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & 0 & 0\end{smallmatrix}\right)$

For a boost in the x-direction, write $\Delta\omega^\nu{}_\mu = \Delta\omega I^\nu{}_\mu$ and set $\Delta\omega = \frac{\omega}{N}$. This gives:

$$\begin{aligned}x'^\nu &= \Lambda^\nu{}_\mu x^\mu\\ &= \lim_{N\to\infty}\left(g + \tfrac{\omega}{N}I\right)^\nu{}_{\alpha_1}\left(g + \tfrac{\omega}{N}I\right)^{\alpha_1}{}_{\alpha_2}\cdots\left(g + \tfrac{\omega}{N}I\right)^{\alpha_{N+1}}{}_{\alpha_N}x^{\alpha_N}\\ &= \underbrace{\left(e^{\omega I}\right)^\nu{}_\mu}_{\Lambda^\nu{}_\mu}x^\mu\end{aligned}$$

Since $I^3 = I$, note that

$$I^n = \begin{cases} I & \text{if } n \text{ is odd,}\\ I^2 & \text{if } n \text{ is even.}\end{cases}$$

Using this in the expansion for $e^{\omega I}$ gives:

$$\begin{aligned}e^{\omega I} &= \mathbb{1} + \sum_{n=1}^{\infty}\frac{\omega^n}{n!}I^n\\ &= \mathbb{1} + I^2\sum_{n=1}^{\infty}\frac{\omega^{2n}}{(2n)!} + I\sum_{n=1}^{\infty}\frac{\omega^{2n-1}}{(2n-1)!}\\ &= \mathbb{1} + I^2(\cosh\omega - 1) + I\sinh\omega.\end{aligned}$$

Writing this as a matrix:

$$e^{\omega I} = \Lambda^{\nu}{}_{\mu} = \begin{pmatrix} \cosh\omega & -\sinh\omega & 0 & 0 \\ -\sinh\omega & \cosh\omega & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

ω is the rapidity. $\cosh\omega = \gamma = \dfrac{1}{\sqrt{1-\beta^2}}$, $\sinh\omega = \gamma\beta$ and $\tanh\omega = \beta$. So:

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}$$

We want to find S for general ω, based on (14) which applies for an infinitesimal $\Delta\omega$.

$$\begin{aligned} S &= \lim_{N\to\infty} \left(\mathbb{1} - \frac{i}{4}\frac{\omega}{N}\sigma_{\mu\nu} I^{\mu\nu} \right)^N \\ &= \exp\left(-\frac{i}{4}\omega\sigma_{\mu\nu} I^{\mu\nu} \right), \end{aligned} \tag{17}$$

where $I^{\mu\nu} = g^{\nu\rho} I^{\mu}{}_{\rho}$ as usual.

$$\begin{aligned} S &= \exp\left(-\tfrac{i}{4}\omega(\sigma_{01} I^{01} + \sigma_{10} I^{10})\right) \\ &= \exp\left(-\tfrac{i}{4}\omega(\sigma_{01} - \sigma_{10}\right) \\ &= \exp\left(-\tfrac{i}{2}\omega\sigma_{01}\right) \end{aligned}$$

Since $(i\sigma_{01})^2 = \mathbb{1}$, we can write:

$$(i\sigma_{01})^n = \begin{cases} i\sigma_{01} & \text{if } n \text{ is odd,} \\ \mathbb{1} & \text{if } n \text{ is even.} \end{cases}$$

So, finally:

$$\begin{aligned} \psi'(x') &= \exp\left(-\frac{\omega}{2} i\sigma_{01}\right)\psi(x) \\ &= \left(\cosh\frac{\omega}{2} - i\sigma_{01}\sinh\frac{\omega}{2}\right)\psi(x) \end{aligned}$$

σ_{01} can be expanded:

$$\begin{aligned} -i\sigma_{01} &= -i\tfrac{1}{2}\left(\gamma_0\gamma_1 - \gamma_1\gamma_0\right) \\ &= \tfrac{1}{2}\left(2\gamma_0\gamma_1\right) = \gamma_0\gamma_1. \end{aligned}$$

This gives:

$$\psi'(x') = \left(\cosh\frac{\omega}{2} + \gamma_0\gamma_1 \sinh\frac{\omega}{2}\right)\psi(x).$$

8.3 Rotation

For a rotation around the z-axis, $I_z^{12} = -1$, $I_z^{21} = 1$, and all other entries are 0. Using a rotation angle of α we obtain an analogous equation to (17):

$$\begin{aligned} S &= \exp\left(\tfrac{i}{4}\alpha\sigma_{\mu\nu}I_z^{\mu\nu}\right) \\ &= \exp\left(-\tfrac{1}{2}\alpha\sigma_{12}\right) \\ &= \cos\tfrac{\alpha}{2} - i\sigma_{12}\sin\tfrac{\alpha}{2} \end{aligned}$$

Since $\sigma_{12} = \frac{i}{2}[\gamma_1,\gamma_2] = i\gamma_1\gamma_2 = \left(\begin{smallmatrix}\sigma_3 & 0 \\ 0 & \sigma_3\end{smallmatrix}\right)$,

$$\phi'(x') = e^{\frac{i}{2}\alpha\sigma_3}\phi(x).$$

So a 4π rotation is needed to get back to ϕ. This is an important property of spinors.

So for a rotation,

$$S(\Lambda) = e^{-\frac{i}{4}\sigma_{\mu\nu}\omega^{\mu\nu}},$$

where $\omega^{\mu\nu}$ is real with $\omega^{\mu\nu} = -\omega^{\nu\mu}$.

$$\sigma_{\mu\nu}^\dagger = \left(\tfrac{i}{2}[\gamma_\mu,\gamma_\nu]\right)^\dagger = -\tfrac{i}{2}[\gamma_\nu^\dagger,\gamma_\mu^\dagger].$$

Recall (11): $(\gamma^\mu)^\dagger = \gamma^0\gamma^\mu\gamma^0$. This allows us to write:

$$\sigma_{\mu\nu}^\dagger = \gamma_0\sigma_{\mu\nu}\gamma_0.$$

In other words,

$$\sigma_{\mu\nu}^\dagger = \begin{cases} \sigma_{\mu\nu} & \text{if } \mu,\nu = 1,2,3 \\ -\sigma_{\mu\nu} & \text{if } \mu = 0 \text{ or } \nu = 0 \end{cases}$$

This means:

$$S^{-1}(\Lambda) = \gamma_0 S(\Lambda)^\dagger\gamma_0, \quad \text{and} \quad S(\Lambda)^\dagger = \gamma_0 S^{-1}(\Lambda)\gamma^0,$$

since for any A:

$$\begin{aligned} e^{\gamma_0 A\gamma_0} &= \sum_n \frac{(\gamma_0 A\gamma_0)^n}{n!}, \text{ but} \\ (\gamma_0 A\gamma_0)^n &= \gamma_0 A\underbrace{\gamma_0\gamma_0}_{=1}A\gamma_0\ldots = \gamma_0 A^n\gamma_0, \\ \Rightarrow e^{\gamma_0 A\gamma_0} &= \gamma_0 e^A\gamma_0. \end{aligned}$$

8.4 Probability current

Recall the probability current:

$$\begin{aligned} j^\mu(x) &= (\psi^\dagger\psi, \psi^\dagger\alpha_k\psi) \\ &= (\psi^\dagger\psi, \psi^\dagger\gamma^0\gamma^k\psi) \\ &= (\overline{\psi}\gamma^0\psi, \overline{\psi}\gamma^k\psi) \\ &= \overline{\psi}\gamma^\mu\psi, \end{aligned}$$

remember $\gamma^k = \beta\alpha_k$, $\beta = \gamma^0$, $\overline{\psi} = \psi^\dagger\gamma_0$. So:

$$\begin{aligned} j'^\mu(x') &= \overline{\psi}(x')\gamma^\mu\psi'(x') \\ &= \psi^\dagger(x')\gamma^0\gamma^\mu\psi'(x') \\ &= \psi^\dagger(x) \underbrace{S(\Lambda)^\dagger}_{=\gamma_0 S^{-1}(\Lambda)} \gamma^0\gamma^\mu S(\Lambda)\psi(x) \\ &= \underbrace{\psi^\dagger(x)\gamma^0}_{=\overline{\psi}(x)}\underbrace{S^{-1}(\Lambda)\gamma^\mu S(\Lambda)}_{=\Lambda^\mu{}_\nu\gamma^\nu}\psi(x) \end{aligned}$$

So:

$$j'^\mu(x) = \Lambda^\mu{}_\nu j^\nu(x),$$

i.e. j^μ is a Lorentz vector field.

8.5 General bilinear form

What about a general bilinear form $\overline{\psi}(x)A\psi(x)$ with an arbitrary 4×4 matrix A?

Observation:

$$\begin{gathered} \psi'(x') = S(\Lambda)\psi(x) \Rightarrow \overline{\psi}'(x') = \overline{\psi}(x)S^{-1}(\Lambda) \\ \Rightarrow \overline{\psi}'(x')A\psi'(x') \longrightarrow \overline{\psi}(x)S^{-1}(\Lambda)\,A\,S(\Lambda)\psi(x). \end{gathered}$$

Any 4×4 matrix A can be expanded in a basis set of 16 matrices; one particular choice is provided by the γ matrices. Define:

$$\gamma_5 = \gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3.$$

This has properties $\{\gamma^5,\gamma^\mu\} = 0$, $(\gamma^5)^2 = 1$ and $\gamma_5^\dagger = \gamma_5$.

This allows us to define 16 matrices:

$$\begin{aligned} \Gamma^S &= \mathbb{1} && (1) \\ \Gamma_\mu{}^V &= \gamma_\mu && (4) \\ \Gamma_\mu\nu^T &= \sigma_{\mu\nu} = \tfrac{i}{2}[\gamma_\mu,\gamma_\nu] && (6) \\ \Gamma_\mu{}^A &= \gamma_5\gamma_\mu && (4) \\ \Gamma^P &= \gamma_5. && (1) \end{aligned}$$

These have properties:

- $(\Gamma^a)^2 = \pm 1$, for $a = 1,\dots,16$.
- For any $\Gamma^a \neq \Gamma^S$ there is a Γ^b such that $\{\Gamma^a,\Gamma^b\} = 0$.
- $\operatorname{tr}\Gamma^a = 0$ for $\Gamma^a \neq \Gamma^S$, since:

$$\operatorname{tr}[\Gamma^a\Gamma^b\Gamma^b] = \operatorname{tr}[\Gamma^b \underbrace{\Gamma^a\Gamma^b}_{=-\Gamma^b\Gamma^a}] = -\operatorname{tr}[\Gamma^a\Gamma^b\Gamma^b].$$

- For any Γ^a, Γ^b, $a \neq b$, there is a $\Gamma^c \neq \Gamma^S$ such that $\Gamma^a\Gamma^b = \eta\Gamma^c$ with $\eta = \pm 1, \pm i$.

  For example, if $\Gamma^a = \gamma_5\gamma_1 = \Gamma^A_1$, and $\gamma^b = \sigma_{01} = \Gamma^T_{01}$, $\Gamma^a\Gamma^b = i\gamma_5\gamma_0 = i\Gamma^A_0$.

- $\{\Gamma^a\}$ are linearly independent:

$$\sum_{a=1}^{16} \lambda_a\Gamma^a = 0 \Rightarrow \lambda_a = 0 \quad \forall a.$$

  This means that any 4×4 matrix A can be written as a linear combination of Γ^as.

How do $\overline{\psi}(x)\Gamma^a\psi(x)$, $a = 1,\dots,16$ transform?

- Γ^S:

$$\begin{aligned}\overline{\psi}'(x')\Gamma^S\psi'(x') &= \overline{\psi}(x)S^{-1}(\Lambda)S(\Lambda)\psi(x)\\ &= \overline{\psi}(x)\psi(x),\end{aligned}$$

  Γ^S represent a scalar field.

- For $\Gamma_\mu{}^V$, a vector field, see j^μ in Sec. 8.4.
- Γ^P: $\overline{\psi}'(x')\gamma_5\psi'(x') = \overline{\psi}(x)S^{-1}(\Lambda)\gamma_5S(\Lambda)\psi(x)$.

  For a *proper* Lorentz transformation,

$$\begin{aligned}S = e^{-\frac{i}{4}\sigma_{\mu\nu}\omega^{\mu\nu}}, &\quad \gamma_5\sigma_{\mu\nu} = \sigma_{\mu\nu}\gamma_5\\ \Rightarrow \gamma^5S(\Lambda) = S(\Lambda)\gamma^5 &\end{aligned}$$

  So,

$$\begin{aligned}\overline{\psi}'(x')\gamma_5\psi'(x') &= \overline{\psi}(x)S^{-1}(\Lambda)S(\Lambda)\gamma^5\psi'(x')\\ &= \overline{\psi}(x)\gamma^5\psi'(x'),\end{aligned}$$

  just like a scalar field!

  In general, $\overline{\psi}'(x')\gamma_5\psi'(x') = \det\Lambda\overline{\psi}(x)\gamma_5\psi(x)$. Γ^P represents a pseudoscalar field.

- $\overline{\psi}'(x')\gamma^5\gamma^\mu\psi'(x') = (\det\Lambda)\Lambda^\mu{}_\nu\overline{\psi}(x)\gamma^5\gamma^\nu\psi(x)$. Γ^A represents an axial vector field.
- $\overline{\psi}'(x')\sigma^{\mu\nu}\psi'(x') = \Lambda^\mu{}_\rho\Lambda^\nu{}_\sigma\sigma^{\rho\sigma}\psi(x)$. Γ^T represents an antisymmetric tensor field.

8.6 Parity symmetry

We want to find $S(P)$, where $P^\mu{}_\nu = \mathrm{diag}(1,-1,-1,-1)$, the parity transformation. Since $\det P = -1$, P is not a proper Lorentz transformation. We know that:

$$S^{-1}(P)\gamma^\mu S(P) = P^\mu{}_\nu\gamma^\nu.$$

Take $\mu = 0$:

$$S^{-1}(P)\gamma^0S(P) = \gamma^0 \Longleftrightarrow \gamma^0S(P) = S(P)\gamma^0.$$

$\mu = k$:

$$S^{-1}(P)\gamma^k S(P) = -\gamma^k \Longleftrightarrow \gamma^k S(P) = -S(P)\gamma^k.$$

These two conditions give:

$$S(P) = e^{i\varphi}\gamma^0,$$

where φ is an arbitrary phase factor. For convenience we set $\varphi = 0$.

We can now see how γ_5 transforms under P:

$$\begin{aligned}
\overline{\psi}'(x')\gamma_5\psi'(x') &= \overline{\psi}(x)S^{-1}(P)\gamma_5 S(P)\psi(x)\\
&= \overline{\psi}(x)\gamma^0\gamma_5\gamma^0\psi(x)\\
&= \overline{\psi}(x)\gamma^0(-\gamma^0\gamma_5)\psi(x)\\
&= -\overline{\psi}(x)\gamma^0\gamma^0\gamma_5\psi(x)\\
&= -\overline{\psi}(x)\gamma_5\psi(x)
\end{aligned}$$

Here a factor of -1 appears. In general the factor is $\det P$.

9 Lagrangian density for free fields

All field equations (e.g. Maxwell, Klein–Gordon, Dirac, Yang–Mills, etc.) can be obtained from a minimum action principle, where the action is:

$$S = \int d^4x\,\mathscr{L}$$

$\mathscr{L}$ is the Lagrangian density, commonly called simply the Lagrangian.

We make the assumption that $\mathscr{L}$ depends only on $\phi(x)$ and $\partial_\mu\phi(x)$, and not on any higher derivatives. $\phi(x)$ can be any field component, e.g. $A_0(x)$, $\psi_1(x)$, etc.

We vary the field: $\phi(x) \to \phi(x) + \epsilon\,\delta\phi(x)$, and require that $\delta S[\phi] = 0$.

$$\begin{aligned}
\delta S[\phi] &\overset{\text{def}}{=} \frac{d}{d\epsilon} S\big[\phi + \epsilon\,\delta\phi\big]\Big|_{\epsilon=0}\\
&= \frac{d}{d\epsilon}\int d^4x\,\mathscr{L}\big(\phi(x) + \epsilon\,\delta\phi(x), \partial_\mu\phi(x) + \epsilon\,\partial_\mu\delta\phi(x)\big)\\
&= \int d^4x\left[\frac{\partial\mathscr{L}}{\partial\phi(x)}\delta\phi(x) + \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi(x))}\partial_\mu\delta\phi(x)\right]\\
&= \int d^4x\left[\frac{\partial\mathscr{L}}{\partial\phi(x)} - \partial_\mu\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi(x))}\right]\delta\phi(x)
\end{aligned}$$

$$\delta S[\phi] = 0 \Rightarrow \frac{\partial\mathscr{L}}{\partial\phi(x)} - \partial_\mu\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi(x))} = 0,$$

the field equation, or equation of motion for the field ϕ.

9.1 Real scalar field

For a real scalar field $\phi(x)$,

$$\mathscr{L}(x) = \tfrac{1}{2}\big(\partial_\mu\phi(x)\big)\big(\partial^\mu\phi(x)\big) - \tfrac{1}{2}m^2\big(\phi(x)\big)^2$$

$$\frac{\partial\mathscr{L}}{\partial\phi(x)} = -m^2\phi(x), \quad \partial_\mu\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi(x))} = \partial_\mu\partial^\mu\phi(x)$$

So the equation of motion reads: $(-m^2 - \partial_\mu\partial^\mu)\phi(x) = 0$, or $(\Box + m^2)\phi(x) = 0$, i.e. the Klein–Gordon equation.

9.2 Complex scalar field

For a complex scalar field $\Phi(x)$,

$$\mathscr{L}(x) = \big(\partial_\mu\Phi^*(x)\big)\big(\partial^\mu\Phi(x)\big) - m^2\Phi^*(x)\Phi(x)$$

$$\frac{\partial\mathscr{L}}{\partial\Phi(x)} = -m^2\Phi^*(x), \quad \partial_\mu\frac{\partial\mathscr{L}}{\partial(\partial_\mu\Phi(x))} = \partial_\mu\partial^\mu\Phi^*(x) = \Box\Phi^*(x)$$

So the equation of motion reads $(\Box + m^2)\Phi^*(x) = 0$. Similarly, $(\Box + m^2)\Phi(x) = 0$. Since Φ and Φ^* can be treated as independent, one complex scalar field is equivalent to two real scalar fields.

Φ can be defined in terms of real scalar fields φ_1 and φ_2:

$$\begin{aligned}\Phi(x) &= \tfrac{1}{\sqrt{2}}\big(\varphi_1(x) + i\varphi_2(x)\big)\\ \Rightarrow \Phi^*(x) &= \tfrac{1}{\sqrt{2}}\big(\varphi_1(x) - i\varphi_2(x)\big)\end{aligned}$$

This gives the properties:

$$\begin{aligned}(\partial_\mu\Phi^*)(\partial^\mu\Phi) &= \tfrac{1}{2}(\partial_\mu\varphi_1)(\partial^\mu\varphi_1) + \tfrac{1}{2}(\partial_\mu\varphi_2)(\partial^\mu\varphi_2)\\ \Phi^*\Phi &= \tfrac{1}{2}{\varphi_1}^2 + \tfrac{1}{2}{\varphi_2}^2.\end{aligned}$$

9.3 Spin-$\frac{1}{2}$ field

$$\begin{aligned}\mathscr{L}(x) &= \overline{\psi}(x)(i\not{\partial} - m)\psi(x)\\ &= \sum_{\alpha,\beta=1}^{4}\overline{\psi}_\alpha(x)(i\not{\partial}_{\alpha\beta} - m\delta_{\alpha\beta})\psi_\beta(x)\end{aligned}$$

So for each α,d

$$\begin{aligned}\frac{\partial\mathscr{L}}{\partial\overline{\psi}_\alpha(x)} &= \sum_{\beta=1}^{4}(i\not{\partial}_\beta - m\delta_{\alpha\beta})\psi_\beta(x)\\ &= (i\not{\partial} - m)\psi(x).\end{aligned}$$

$$\frac{\partial \mathscr{L}}{\partial(\partial_\mu \overline{\psi}_\alpha(x))} = 0$$

These give the equation of motion:

$$(i\not{\partial} - m)\psi(x) = 0.$$

In addition,

$$\frac{\partial \mathscr{L}}{\partial \psi_\beta(x)} = -m\overline{\psi}_\beta(x),$$

$$\begin{aligned}
\partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \psi_\beta(x))} &= \partial_\mu \sum_{\alpha=1}^{4} \overline{\psi}_\alpha(x) i\gamma^\mu_{\alpha\beta} \\
&= \partial_\mu \left(\overline{\psi}(x) i\gamma^\mu\right)_\beta \\
&= \left(\overline{\psi}(x) i\gamma^\mu\right)_\beta \overleftarrow{\partial}_\mu \\
&= (\overline{\psi}(x) i\overleftarrow{\not{\partial}})_\beta
\end{aligned}$$

These give the equation of motion:

$$\overline{\psi}(x)(i\overleftarrow{\not{\partial}} + m) = 0.$$

9.4 Abelian vector field

$$\mathscr{L} = -\tfrac{1}{4} F_{\mu\nu} F^{\mu\nu} + j_\mu A^\mu,$$

where $F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x)$.

$$\frac{\partial \mathscr{L}}{\partial A_\nu(x)} = j^\nu(x), \quad \partial_\rho \frac{\partial \mathscr{L}}{\partial(\partial_\rho A_\sigma(x))} = -\partial_\rho F^{\rho\sigma}$$

So the equation of motion is

$$\partial_\rho F^{\rho\sigma} = j^\sigma.$$

9.5 Summary

- The Lagrangian density must behave like a scalar field under proper Lorentz transformations; $d^4x' = |\det\Lambda| d^4x = d^4x$; so the action $S = \int d^4x \mathscr{L}(x)$ is invariant under $\Lambda \in L_+$.
- The equations of motion obtained from $\delta S = 0$ are covariant; they take the same form for all observers.
- For the standard model, we need the Lagrangians of spin-0 bosons (scalar fields, real or complex), spin-$\frac{1}{2}$ fermions (Dirac eqn.), spin-1 bosons (abelian vector fields) and non-abelian vector fields ($W^\pm$, Z, gluons)

10 More on the Dirac equation

10.1 Free-particle solutions

$$(i\partial\!\!\!/ - m)\psi(x) = 0.$$

This has free particle solutions given by:

$$\psi(x) = u(p)e^{-ipx}$$

where $u(p)$ is a 4-component fixed spinor. Since $\partial_\mu e^{-ip} = -ip_\mu e^{-ipx}$, we have:

$$\begin{aligned} 0 = (i\partial\!\!\!/ - m)u(p)e^{-ipx} &= (p\!\!\!/ - m)u(p)e^{-ipx} \\ \Rightarrow (p\!\!\!/ - m)u(p) &= 0. \end{aligned}$$

To distinguish positive and negative energy solutions, go back to $H = \boldsymbol{\alpha}\cdot\mathbf{p} + \beta m$, with $\mathbf{p} = -i\boldsymbol{\nabla}$.

$$H\,u(p) = (\boldsymbol{\alpha}\cdot\mathbf{p} + \beta m)u(p) = E\,u(p)$$

where E is the energy eigenvalue.

When $\mathbf{p} = 0$ this gives:

$$H\,u = \beta m u = \begin{pmatrix} m & 0 & 0 & 0 \\ 0 & m & 0 & 0 \\ 0 & 0 & -m & 0 \\ 0 & 0 & 0 & -m \end{pmatrix} u = Eu.$$

This has two solutions with $E = m$ and two with $E = -m$.

When $\mathbf{p} \neq 0$,

$$H\,u = \begin{pmatrix} m & \boldsymbol{\sigma}\cdot\mathbf{p} \\ \boldsymbol{\sigma}\cdot\mathbf{p} & -m \end{pmatrix}\begin{pmatrix} u_A \\ u_B \end{pmatrix} = E\begin{pmatrix} u_A \\ u_B \end{pmatrix}$$

Note that this is a 4×4 matrix multiplying a 4-component column vector; u_A and u_B are 2-component spinors.

$$\begin{aligned} \boldsymbol{\sigma}\cdot\mathbf{p}\,u_B &= (E-m)u_A \\ \boldsymbol{\sigma}\cdot\mathbf{p}\,u_A &= (E+m)u_B \end{aligned}$$

10.2 Constructing a basis

We can construct a basis for solutions with $E > 0$:

$$u_A^{(s)} = \chi^{(s)}; \quad \chi^{(1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \chi^{(2)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$u_B^{(s)} = \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\chi^{(s)}$$

So for $E > 0$, and $s = 1, 2$:

$$u^{(s)} = N \begin{pmatrix} \chi^{(s)} \\ \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\chi^{(s)} \end{pmatrix}$$

For solutions with $E < 0$:

$$u_B^{(s)} = \chi^{(s)}$$

$$\begin{aligned} \Rightarrow u_A^{(s)} &= \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E-m}\chi^{(s)} \\ &= -\frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{|E|+m}\chi^{(s)} \end{aligned}$$

So for $E < 0$, and $s = 1, 2$:

$$u^{(s+2)} = N \begin{pmatrix} \frac{-\boldsymbol{\sigma}\cdot\mathbf{p}}{|E|+m}\chi^{(s)} \\ \chi^{(s)} \end{pmatrix}$$

Solutions are orthogonal: $u^{(r)\dagger} u^{(s)} \propto \delta_{rs}$

10.3 Helicity

Within each pair of solutions, one distinguishes spin (up vs. down) by the helicity operator:

$$\frac{1}{2}\boldsymbol{\Sigma}\cdot\hat{\mathbf{p}} = \frac{1}{2}\begin{pmatrix} \boldsymbol{\sigma}\cdot\hat{\mathbf{p}} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\hat{\mathbf{p}} \end{pmatrix}$$

where $\hat{\mathbf{p}} = \frac{\mathbf{p}}{|\mathbf{p}|}$.

Since $\left[\boldsymbol{\Sigma}\cdot\mathbf{p}, H\right] = 0$, helicity is a good quantum number.

$$\left(\frac{1}{2}\boldsymbol{\Sigma}\cdot\hat{\mathbf{p}}\right)^2 = \frac{1}{4}\begin{pmatrix} (\boldsymbol{\sigma}\cdot\hat{\mathbf{p}})(\boldsymbol{\sigma}\cdot\hat{\mathbf{p}}) & 0 \\ 0 & (\boldsymbol{\sigma}\cdot\hat{\mathbf{p}})(\boldsymbol{\sigma}\cdot\hat{\mathbf{p}}) \end{pmatrix} = \tfrac{1}{4}\mathbb{1}.$$

So $\frac{1}{2}\boldsymbol{\Sigma}\cdot\hat{\mathbf{p}}$ has eigenvalues $\pm\frac{1}{2}$, the possible helicities of the particle.

10.4 Interpretation of negative energy solutions

So far, we have attempted to obtain a relativistic wave equation with 1-particle wavefunctions and a probability interpretation as in non-relatvistic quantum mechanics.

With the Klein–Gordon equation we defined the probability current:

$$j^\mu = i(\phi^*\partial^\mu\phi - \phi\partial^\mu\phi^*) = (\rho, \mathbf{j}),$$

and the free-particle solution $\phi = Ne^{-ipx}$ gave $j^\mu = 2p^\mu|N|^2 \Rightarrow \rho \propto E$. For $E < 0$ this gave $\rho < 0$ and so the Klein–Gordon equation was abandoned.

With the Dirac equation,

$$j^\mu = \overline{\psi}\gamma^\mu\psi = (\rho, \mathbf{j}),$$

so $\rho = \overline{\psi}\gamma^0\psi = \psi^\dagger\gamma^0\gamma^0\psi = \psi^\dagger\psi \geq 0$. There are no more negative probability densities, but there are still negative energy solutions!

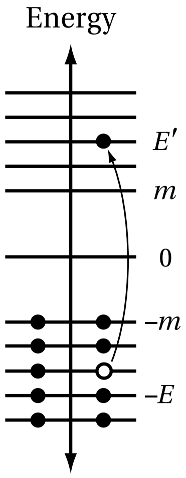


Figure 1: Pair production

Dirac's solution was to use the Pauli exclusion principle and assume that all negative energy states are filled with electrons; the "Dirac sea".

When an electron with negative energy $-E$ is excited to a state with energy $E' > 0$, the result is:

- an electron with charge $-e$, $E' > 0$
- the absence of an electron with charge $-e$ and energy $-E < 0$.

Interpret the latter as the presence of an antiparticle with charge $+e$ and energy $E > 0$.

The net result is hence the creation of a pair $e^-(E') + e^+(E)$ with $E + E' > 2m$. Dirac predicted the existence of the positron in this way in 1932; it is an acceptable (if outdated) theory of spin-$\frac{1}{2}$ particles.

In 1934 Pauli and Weisskopf revived the Klein–Gordon equation by reinterpreting the probability current as charge current density:

$$j^\mu = -ie(\phi^* \partial^\mu \phi - \phi \partial^\mu \phi^*),$$

so having $\rho < 0$ is not a problem.

11 Feynman–Stueckelberg interpretation

11.1 Motivation

Dirac's interpretation of negative energies works only for fermions. What about bosons?

The Feynman–Stueckelberg interpretation treats a negative energy solution propagating backwards in time as identical to a positive energy solution going forwards in time.

11.2 Example

Consider the scattering of a π^+ particle and π^- antiparticle by a potential V:

Both time orderings seen in Fig. 2 must be taken into account. In each case the initial and final situation is identical. The second case, shown in Fig. 2b, can be interpreted as shown in Fig. 3, as the creation of a $\pi^+\pi^-$ pair by one potential, and the absorption of a $\pi^+\pi^-$ pair by the other potential.

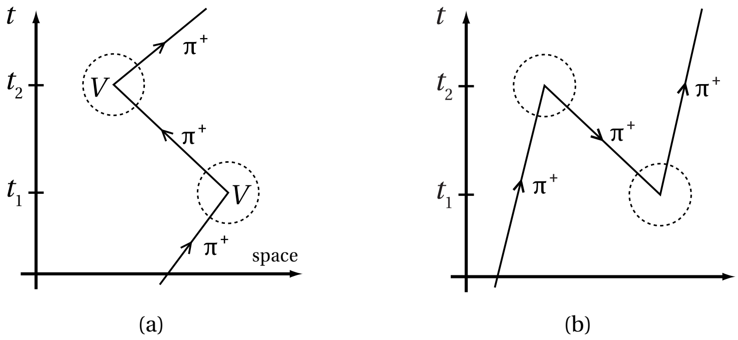


Figure 2: π^+ scattering

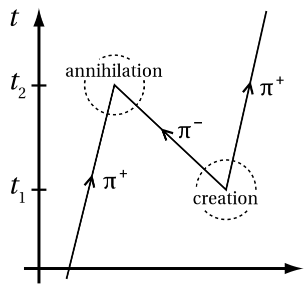


Figure 3: The Feynman–Stueckelberg interpretation of Fig. 2b

11.3 Currents

Consider the electromagnetic current for a positive-energy π^+:

$$\begin{aligned} j^\mu_{\text{em}}(\pi^+) &= +e \times \left(\text{prob. current for } \pi^+\right) \\ &= 2|N|^2 p^\mu \\ &= 2|N|^2 (E, \mathbf{p}) \end{aligned}$$

The current for a π^- with negative energy is:

$$\begin{aligned} j^\mu_{\text{em}}(\pi^-) &= -e \times 2|N|^2 p^\mu \\ &= -2e|N|^2 (E, \mathbf{p}) \\ &= e\,2|N|^2 (-E, -\mathbf{p}) \\ &= j^\mu_{\text{em}}(\pi^+)\Big|_{p^\mu \to -p^\mu} \end{aligned}$$

Consider a system A with p^μ_A and charge q_A and the emission of a π^- with $E > 0$. This can be represented in either of two ways shown in Fig. 4.

p_A and q_A change as:

$$p^\mu_A \to p^\mu_A - p^\mu(\pi^-) = p^\mu_A + (-p^\mu_{\pi^-}).$$

$$q_A \to q_A - (-e) = q_A + e.$$

The emission (absorption) of an antiparticle of 4-momentum p^μ is physically equivalent to the absorption (emission) of a particle with 4-momentum $-p^\mu$.

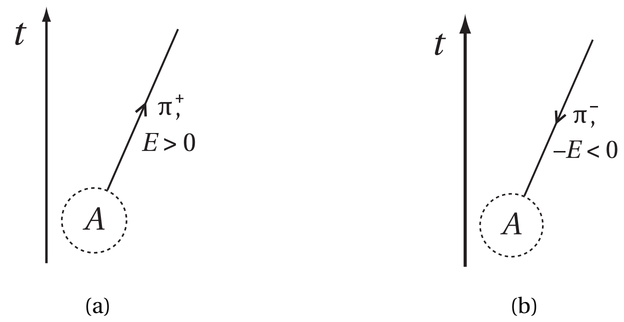


Figure 4: Two representations of the system A

12 Electrodynamics of spin-0 particles

We consider π and K particles as elementary particles and interacting only electromagnetically (although this is highly unrealistic).

Recall time-dependent non-relativistic scattering theory, and the Schrödinger equation for a free particle:

$$H_0\phi_n = E_n\phi_n,$$

where $\{\phi_0, \phi_1, \ldots\}$ form an orthonormal basis:

$$(\phi_n, \phi_m) = \int d^3x\, \phi_n^*(\mathbf{x})\, \phi_m(\mathbf{x}) = \delta_{mn}.$$

Our aim is to solve the Schrödinger equation for a particle in the potential $V(\mathbf{x}, t)$:

$$(H_0 + V(\mathbf{x}, t))\, \psi = i\frac{\partial\psi}{\partial t}. \tag{18}$$

Expand:

$$\psi(\mathbf{x}, t) = \sum_{n=0}^{\infty} a_n(t)\, \phi_n(\mathbf{x}) e^{-iE_n t}$$

and insert into (18):

$$\sum_n a_n(t) V(\mathbf{x}, t)\phi_n(\mathbf{x}) e^{-iE_n t} = i\sum_n \frac{da_n(t)}{dt}\phi_n(\mathbf{x}) e^{-iE_n t}$$

Multiply this by $\phi_f^*(\mathbf{x})$ and integrate.

$$\int d^3x \sum_n a_n(t)\phi_f^*(\mathbf{x}) V(\mathbf{x}, t)\phi_n(\mathbf{x}) e^{-iE_n t} = i\frac{da_f(t)}{dt} e^{-iE_f t}$$

$$\frac{da_f(t)}{dt} = -i\sum_n a_n(t) \int d^3x \phi_f^*(\mathbf{x})\, V(\mathbf{x}, t)\, \phi_n(\mathbf{x})\, e^{-i(E_f - E_n)t}$$

Assume at $t = -\frac{T}{2}$ the system is in state i:

$$a_i\left(-\tfrac{T}{2}\right) = 1, \quad a_n\left(-\tfrac{T}{2}\right) = 0 \quad \forall n \neq i$$

$$\frac{da_f}{dt}\left(-\tfrac{T}{2}\right) = -i\int d^3x\phi_f^*(\mathbf{x})V\phi_i(\mathbf{x})e^{i(E_f-E_i)(T/2)} \tag{19}$$

Assume that the potential is small, so there is no perturbation of the initial state up to a 0^{th}-order approximation.

Integrate (19):

$$a_f(t) = -i\int_{-T/2}^{t} dt'\int d^3x\phi_f^*(\mathbf{x})V\phi_i(\mathbf{x})e^{i(E_f-E_i)t'}$$

At $t = T/2$, after the interaction,

$$\begin{aligned} T_{fi} &= a_f\left(\tfrac{T}{2}\right) \\ &= -i\int_{-T/2}^{T/2} dt\int d^3x\,\underbrace{\left[\phi_f(\mathbf{x})e^{-iE_ft}\right]^*}_{\phi_f^*(x)} V(\mathbf{x},t)\,\underbrace{\left[\phi_i(\mathbf{x})e^{-iE_it}\right]}_{\phi_i(x)}, \end{aligned}$$

where $x \equiv (t,\mathbf{x})$. So a covariant form for T_{fi} is:

$$T_{fi} = -i\int d^4x\phi_f^*(x)V\phi_i(x).$$

12.1 Interpretation of T_{fi}

Consider a time-independent $V(\mathbf{x})$.

$$T_{fi} = -iV_{fi}\int_{-T/2}^{T/2} dt\,e^{-i(E_f-E_i)t},$$

with:

$$V_{fi} = \int d^3x\phi_f^*(\mathbf{x})V(\mathbf{x})\phi_i(\mathbf{x}).$$

Look at the limit as $T \to \infty$:

$$\int_{-T/2}^{T/2} dt\,e^{-i(E_f-E_i)t} \longrightarrow 2\pi\delta(E_f-E_i).$$

So as $T\to\infty$, $|T_{fi}|^2 \propto$ "$\delta^2(E_f-E_i)$". This is not well-defined! ☡ But we can define:

$$\begin{aligned} W &= \lim_{T\to\infty}\tfrac{1}{T}\left|T_{fi}\right|^2 \\ &= \left|V_{fi}\right|^2\lim_{T\to\infty}\left(\frac{1}{T}\int_{-T/2}^{T/2} dt\,e^{-i(E_f-E_i)t}\right)^2 \\ &= \left|V_{fi}\right|^2\lim_{T'\to\infty}\left(\int_{-T'/2}^{T'/2} dt\,e^{-i(E_f-E_i)t}\right)\lim_{T\to\infty}\left(\frac{1}{T}\int_{-T/2}^{T/2} dt\,e^{-i(E_f-E_i)t}\right) \\ &= \left|V_{fi}\right|^2 2\pi\delta(E_f-E_i)\lim_{T\to\infty}\left(\frac{1}{T}\int_{-T/2}^{T/2} dt\,e^{-i(E_f-E_i)t}\right) \end{aligned}$$

Now because of the delta function, we can set $E_f = E_i$ in the integral, since $\delta(E_f - E_i) = 0$ when $E_f \neq E_i$. The integral simply reduces to T, giving:

$$W = \left|V_{fi}\right|^2 2\pi\delta(E_f - E_i).$$

12.2 Fermi's golden rule

We have an initial state specified and a set of final states. Let $\rho(E_f)\,dE_f$ represent the number of final states with energies between E_f and $E_f + dE_f$.

$$\begin{aligned}
W_{fi} &= 2\pi \int dE_f\, \rho(E_f) |V_{fi}|^2 \delta(E_i - E_f) \\
W_{fi} &= 2\pi\rho(E_i)|V_{fi}|^2.
\end{aligned}$$

This is known as **Fermi's golden rule**.

12.3 Determining T_{fi}

Use the Klein–Gordon equation $(\partial_\mu\partial^\mu + m^2)\phi = 0$ and couple to an e.m. potential $A^\mu = (A^0, \mathbf{A})$ through the substitution $p^\mu \to p^\mu + eA^\mu$ (the particle has charge $-e$.) In other words $i\partial^\mu \to i\partial^\mu + eA^\mu$.

So:

$$\left[(\partial_\mu - ieA_\mu)(\partial^\mu - ieA^\mu) + m^2\right]\phi = 0$$

gives:

$$(\partial_\mu\partial^\mu + m^2)\phi = -V\phi,$$

where:

$$V = -ie(\partial_\mu A^\mu + A^\mu\partial_\mu) - e^2 A_\mu A^\mu.$$

Note that V is an operator acting on ϕ, and $\partial_\mu A^\mu\phi = (\partial_\mu A^\mu)\phi + A^\mu(\partial_\mu\phi)$.

Since $\alpha = \frac{e^2}{4\pi} = \frac{1}{137}$, organise interactions in powers of e. Neglect the $e^2 A_\mu A^\mu$ term.

$$\begin{aligned}
T_{fi} &= -i \int \phi_f^*(x) V(x) \phi_i(x)\, d^4x \\
&= -e \int \phi_f^*(x)(A^\mu\partial_\mu + \partial_\mu A^\mu)\phi_i(x)\, d^4x
\end{aligned}$$

This can be integrated by parts:

$$\int \phi_f^* \partial_\mu(A^\mu\phi_i)\, d^4x = -\int (\partial_\mu\phi_f^*) A^\mu\phi_i\, d^4x.$$

12.4 Currents

Now we can define j such that:

$$T_{fi} = -i \int j_\mu^{fi}(x) A^\mu(x)\, d^4x.$$

This requires:

$$j_\mu^{fi}(x) = -ie\left(\phi_f^* \partial_\mu \phi_i - (\partial_\mu \phi_f^*)\phi_i\right),$$

the electromagnetic transition current for $i \longrightarrow f$.

The free particle solutions are:

$$\begin{aligned}\phi_i(x) &= N_i e^{-ip_i x}\\ \phi_f(x) &= N_f e^{-ip_f x},\end{aligned}$$

which gives:

$$j_\mu^{fi}(x) = -eN_i N_f^*(p_i + p_f) e^{i(p_f - p_i)x}$$

To determine A_μ we solve the Maxwell equation: $\partial_\mu F^{\mu\nu}(x) = j^\nu(x)$, where $j^\nu(x)$ is the electromagnetic current produced by a K^-:

$$\begin{aligned}j^\nu(x) &= -ie\left(\phi_4^* \partial^\nu \phi_2 - (\partial^\nu \phi_4^*)\phi_2\right)\\ &= -eN_2 N_4^*(p_2 + p_4)^\nu e^{i(p_4 - p_2)x}\end{aligned}$$

Using the Lorenz gauge condition $\partial_\mu A^\mu = 0$,we can write:

$$\begin{aligned}\partial_\mu F^{\mu\nu} &= \partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu)\\ &= \Box A^\nu - \partial^\nu(\partial_\mu A^\mu)\\ &= \Box A^\nu.\end{aligned}$$

So the Maxwell equation reads simply:

$$\Box A^\nu(x) = j^\nu(x).$$

Using the fact that $\Box e^{ipx} = -p^2 e^{ipx}$, so this is easily solved:

$$\begin{aligned}A^\nu(x) &= -\frac{1}{(p_4 - p_2)^2}(-e)N_2 N_4^*(p_2 + p_4)^\nu e^{i(p_4 - p_2)x}\\ &= -\frac{1}{(p_4 - p_2)^2} j^\nu(x)\\ &= -\frac{1}{q^2} j^\nu(x),\end{aligned}$$

where $q \equiv p_4 - p_2$, the 4-momentum transfer.

Now,

$$T_{fi} = -i \int j_\mu^{fi}(x) \left[-\tfrac{1}{q^2} j^\mu(\mathrm{K}^-; x)\right] d^4x$$

Label $i \to 1$ and $f \to 3$, giving:

$$
\begin{aligned}
T &= -i \int j_{\mu}^{31}(x) \tfrac{-1}{q^2} j_{42}^{\mu}(x)\, d^4x \\
&= N_1 N_3^* N_2 N_4^* \int d^4x\, (-i) e(p_1 + p_3) \tfrac{-1}{q^2} e(p_2 + p_4)^{\mu} e^{i(p_3 - p_1)x} e^{i(p_4 - p_2)x}
\end{aligned}
$$

Recall that $\int d^4x\, e^{ipx} = (2\pi)^4 \delta^{(4)}(p)$. So:

$$
T = (2\pi)^4 \delta^{(4)}(p_1 + p_2 - p_3 - p_4) N_1 N_3^* N_2 N_4^* (-i\mathscr{M}), \tag{20}
$$

where $\mathscr{M}$ is the **invariant amplitude**, defined such that:

$$
-i\mathscr{M} = \left[ie(p_1 + p_3)^{\mu}\right] \left[-i\frac{g_{\mu\nu}}{q^2}\right] \left[ie(p_2 + p_4)^{\nu}\right]
$$

The second term here is the **photon propagator.**

The process can be described using a Feynman diagram as in Fig. 5. The arrows *on* lines follow the flow of electric charge.

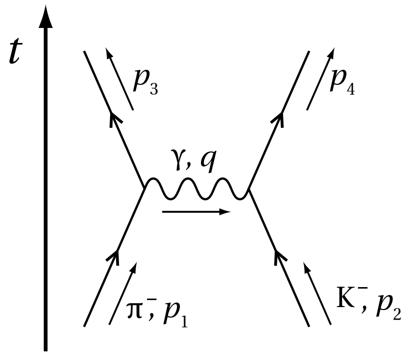


Figure 5: Feynman diagram for π⁻-K⁻ interaction

12.5 Observations

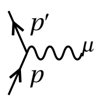
- 4-momentum is conserved at each vertex. At vertex 1,

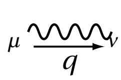
$$
\begin{aligned}
p_1 + (-p_3) + (-q) &= 0 \\
\Rightarrow p_1 - p_3 &= q
\end{aligned}
$$

  At vertex 2,

$$
\begin{aligned}
p_2 + (-p_4) + q &= 0 \\
\Rightarrow p_4 - p_2 &= q
\end{aligned}
$$

  Combining these shows that $p_1 - p_3 = p_4 - p_2$, i.e. $p_1 + p_2 = p_3 + p_4$, which is enforced by the δ in (20).

- Each vertex $\phi\phi\gamma$ gets a factor: 

$= ie(p + p')^{\mu}$, 

$= -\dfrac{g_{\mu\nu}}{q^2}$, the photon propagator in the Lorenz gauge.

13 Scattering

13.1 From scattering amplitude to cross-section

$$T_{12\to34} = (2\pi)^4\delta^{(4)}(p_1+p_2-p_3-p_4)N_1N_2^*N_3N_4^*(-i\mathscr{M})$$

This is the transition rate per unit volume per unit time.

The normalisation condition that we will use for the probability density $\rho = 2E|N|^2$ is:

$$\int_V \rho\, dV = 2E,$$

i.e. there are $2E$ particles per volume V. This gives:

$$N = N^* = \frac{1}{\sqrt{V}},$$

allowing us to write:

$$\begin{aligned} W_{12\to34} &= \frac{|T_{12\to34}|^2}{TV} \\ &= \frac{1}{V^4}(2\pi)^4\delta^{(4)}(p_1+p_2-p_3-p_4)|\mathscr{M}|^2. \end{aligned} \tag{21}$$

The $(\delta^{(4)})^2$ which appears in expanding (21) is dealt with exactly as in section 12.1.

$$\text{Cross-section} = \frac{W_{12\to3,4}}{\text{initial flux}} \times \text{no. of final states,} \tag{22}$$

so we need to find the number of final states and the initial flux.

There are $2E$ particles per volume V; in a finite V the number of states with momentum in $\mathbf{p},\ldots,\mathbf{p}+d^3\mathbf{p}$ is:

$$\frac{V}{(2\pi)^3 2E}d^3p.$$

So the total number of available final states is:

$$\frac{V\,d^3p_3}{(2\pi)^3 2E_3}\frac{V\,d^3p_4}{(2\pi)^3 2E_4}. \tag{23}$$

To find the initial flux we use a frame where particle 2 is at rest. The number of beam particles passing through unit area per unit time is:

$$|\mathbf{V}_1|\cdot\frac{2E_1}{V},$$

where $\mathbf{V}_1$ is the velocity and $\frac{2E_1}{V}$ is the particle density. The number of target particles per unit volume is:

$$\frac{2E_2}{V}.$$

So the initial flux is:

$$\frac{F}{V^2} = |\mathbf{V}_1|\frac{2E_1}{V}\frac{2E_2}{V} \tag{24}$$

So using (23) and (24) in (22) gives:

$$d\sigma = \frac{1}{4E_1E_2\,|\mathbf{V}_1|}(2\pi)^4\delta^{(4)}(p_1+p_2-p_3-p_4)\,|\mathscr{M}|^2\,\frac{d^3p_3}{(2\pi)^3 2E_3}\frac{d^3p_4}{(2\pi)^3 2E_4}. \tag{25}$$

$[d\sigma]$ = area; $d\sigma$ is the effective area over which particles 1 and 2 interact to produce particles 3 and 4, whose momenta lie in $\mathbf{p}_3,\ldots,\mathbf{p}_3+d^3\mathbf{p}_3$ and $\mathbf{p}_4,\ldots,\mathbf{p}_4+d^3\mathbf{p}_4$ respectively.

(25) can be written as:

$$d\sigma = \tfrac{1}{F}\,|\mathscr{M}|^2\;d\text{Lips} \tag{26}$$

where "dLips" is the Lorentz invariant phase space factor:

$$d\text{Lips} = (2\pi)^4\delta^{(4)}(p_1+p_2-p_3-p_4)\frac{d^3p_3}{(2\pi)^3 2E_3}\frac{d^3p_4}{(2\pi)^3 2E_4} \tag{27}$$

F is defined in (24) as $4E_1E_2|\mathbf{V}_1|$. This is not manifestly Lorentz invariant, but by noting that $\mathbf{p}_1 = \gamma_1 m_1\mathbf{V}_1$, and $\gamma_1 = \frac{E_1}{m_1}$, we can write:

$$F = 4\left[(p_1p_2)^2 - m_1^2m_2^2\right]^{1/2} \tag{28}$$

an explicitly Lorentz invariant form.

Are the terms $\frac{d^3\mathbf{p}}{(2\pi)^3 2E}$ in dLips Lorentz invariant? Write:

$$\begin{aligned}&\int\frac{d^4p}{(2\pi)^4}2\pi\delta(p^2-m^2)\theta(p_0)f(p)\\ =&\int\frac{d^3\mathbf{p}}{(2\pi)^3}\int\frac{dp^0}{2\pi}\delta\left((p^0)^2-E^2\right)\theta(p^0)f(p)\end{aligned} \tag{29}$$

where θ is the Heaviside step function. The δ term can be expanded as:

$$\begin{aligned}&\frac{1}{2p_0}\left[\delta(p_0-E)+\delta(p_0+E)\right]\\ =&\frac{1}{2p_0}\left[\delta(p_0-E)\right],\end{aligned}$$

since we define $E = \sqrt{\mathbf{p}^2+m^2} \geqslant 0$. Hence (29) can be written as

$$\int\frac{d^3\mathbf{p}}{(2\pi)^3 2E}f(E,\mathbf{p})\bigg|_{E=\sqrt{\mathbf{p}^2+m^2}}$$

and so the terms are Lorentz invariant.

13.2 Elastic scattering 1,2 → 3,4

We will focus on the centre-of-mass system (CMS); see Fig. 6.

Since we are in the CMS, $\mathbf{p}_1+\mathbf{p}_2 = 0$, so let $\mathbf{p}_1 = -\mathbf{p}_2 = \mathbf{p}$, similarly $\mathbf{p}_3 = -\mathbf{p}_4 = \mathbf{p}'$. So we can write the 4-momenta:

$$p_1^\mu = (E_1,\mathbf{p}),\quad p_2^\mu = (E_2,-\mathbf{p}),\quad p_3^\mu = (E_3,\mathbf{p}'),\quad p_4^\mu = (E_4,-\mathbf{p}').$$

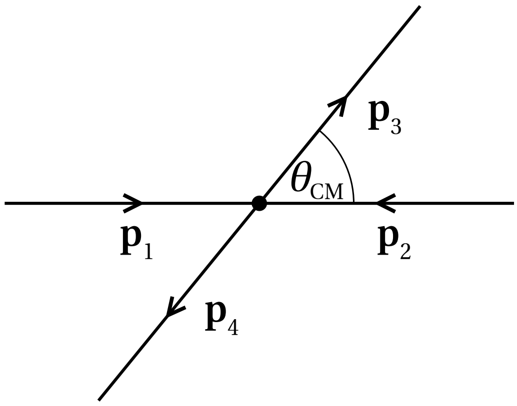


Figure 6: Elastic scattering in CMS

Since we are dealing with elastic scattering, the initial and final states contain the same particles, so $|\mathbf{p}| = |\mathbf{p}'| \equiv p$. Hence:

$$E_1^2 = p^2 + m_1^2, \quad E_2^2 = p^2 + m_2^2, \quad E_3^2 = p^2 + m_3^2 = p^2 + m_1^2, \quad E_4^2 = p^2 + m_4^2 = p^2 + m_2^2.$$

The expression for dLips in (27) contains p_4 explicitly. This is eliminated by integrating so that only the 0-component of the $\delta^{(4)}$ remains:

$$\int \frac{d^3 p_4}{E_4} \delta^{(4)}(p_3 + p_4 - p_1 - p_2) = \frac{1}{E_4} \delta(E_3 + E_4 - E_1 - E_2), \tag{30}$$

where E_4 on the right-hand side is determined by $E_4 = \sqrt{\mathbf{p}_4^2 + m_4^2}$ and $\mathbf{p}_4 = -\mathbf{p}_3 + \mathbf{p}_1 + \mathbf{p}_2$, i.e. it is no longer independent.

We can decompose $d^3 p_3 = |\mathbf{p}_3|^2 \, d|\mathbf{p}_3| \, d\Omega$. Since $E_3^2 = \mathbf{p}_3^2 + m_3^2$, this gives:

$$E_3 \, dE_3 = |\mathbf{p}_3| d|\mathbf{p}_3|.$$

So using this and (30), we get:

$$d\text{Lips} = \frac{1}{(4\pi)^2} d\Omega \frac{|\mathbf{p}_3| E_3 \, dE_3}{E_4} \delta(E_3 + E_4 - E_1 - E_2) \tag{31}$$

where dLips now refers to the original dLips integrated over p_4.

Now, bearing in mind that in the CMS $E_3 \, dE_3 = p \, dp = E_4 \, dE_4$, define:

$$\begin{aligned}
w' &\equiv E_3 + E_4 \\
\Rightarrow dw' &= dE_3 + dE_4 \\
&= \left(\tfrac{1}{E_3} + \tfrac{1}{E_4}\right) p \, dp \\
&= \frac{E_3 + E_4}{E_3 E_4} E_3 \, dE_3 \\
&= \frac{w'}{E_4} dE_3
\end{aligned}$$

This gives, defining $w \equiv E_1 + E_2$:

$$\int p \frac{dE_3}{E_4} \delta\big(E_3 + E_4 - (E_1 + E_2)\big) = \int dw' \frac{E_4}{w'} \frac{p}{E_4} \delta(w' - w)$$
$$= \frac{p}{w}. \tag{32}$$

13.3 Mandelstam variables

At this stage it is convenient to introduce **Mandelstam variables:**

$$\begin{aligned} s &\equiv (p_1 + p_2)^2 \\ t &\equiv (p_1 - p_3)^2 = q^2 \\ u &\equiv (p_1 - p_4)^2 \end{aligned}$$

Note that $s = (p_1 + p_2)^2 = (E_1 + E_2)^2 = w^2$ in the CMS, so $w = \sqrt{s}$.

So by using (32) in (31) we can write the convenient expression:

$$d\text{Lips} = \frac{1}{(4\pi)^2} d\Omega \frac{p}{\sqrt{s}},$$

where again the integration has been absorbed into the definition of dLips.

13.4 Flux factor

In (28) we defined:

$$F = 4\left[(p_1 p_2)^2 - m_1^2 m_2^2\right]^{1/2}.$$

Using the fact that:

$$p_1 p_2 = E_1 E_2 - \mathbf{p}_1 \cdot \mathbf{p}_2 = E_1 E_2 + p^2,$$

we can expand:

$$\begin{aligned} F^2 &= 16\left[E_1^2 E_2^2 + 2p^2 E_1 E_2 + p^4 - m_1^2 m_2^2\right] \\ &= 16\left[(p^2 + m_1^2)(p^2 + m_2^2) + 2p^2 E_1 E_2 - m_1^2 m_2^2 + p^4\right] \\ &= 16\left[2p^4 + p^2(m_1^2 + m_2^2) + 2p^2 E_1 E_2\right] \\ &= 16\big[p^2(\underbrace{p^2 + m_1^2}_{=E_1^2} + \underbrace{p^2 + m_2^2}_{=E_2^2}) + 2p^2 E_1 E_2\big] \\ &= 16p^2(E_1 + E_2)^2 \\ &= 16p^2 s \\ \implies F &= 4p\sqrt{s}. \end{aligned}$$

13.5 Putting things together

Now we can rewrite (26):

$$\begin{aligned} d\sigma &= \frac{1}{F}\,|\mathcal{M}|^2\, d\text{Lips} \\ &= |\mathcal{M}|^2\,\frac{1}{16\pi^2}\,\frac{p}{\sqrt{s}}\,\frac{1}{4p\sqrt{s}}\,d\Omega \end{aligned}$$

This gives:

$$\left.\frac{d\sigma}{d\Omega}\right|_{\text{CM}} = \frac{1}{64\pi^2 s}\,|\mathcal{M}|^2$$

We can use $t = q^2$ as a variable, so that $\frac{d\sigma}{dt}$ is independent of reference frame:

$$\begin{aligned} t = q^2 &= (p_1 - p_3)^2 \\ &= p_1^2 + p_3^2 - (2E_1E_3 - 2\mathbf{p}_1\cdot\mathbf{p}_3) \\ &= 2m_1^2 - 2(p^2+m^2) + 2p^2\cos\theta_{\text{CM}} \\ &= -2p^2(1-\cos\theta_{\text{CM}}) \\ \Rightarrow dt &= 2p^2\,d(\cos\theta_{\text{CM}}) \end{aligned}$$

$d\Omega$ can be expanded:

$$d\Omega = d\phi\, d(cos\theta_{\text{CM}}) = d\phi\frac{dt}{2p^2}.$$

Since these are spinless particles there is no ϕ dependence:

$$d\sigma = \int \frac{1}{64\pi^2 s}\,|\mathcal{M}|^2\, d\phi\frac{dt}{2p^2} = \frac{1}{64\pi s}\,|\mathcal{M}|^2\,\frac{dt}{p^2}$$

Use $F^2 = 16p^2 s$:

$$\frac{d\sigma}{dt} = \frac{1}{64\pi}\,\frac{1}{(p_1p_2)^2 - m_1^2m_2^2}\,|\mathcal{M}|^2,$$

a Lorentz invariant form!

14 External spin-1 particles

14.1 Massive particles

For example, vector bosons, $W^\pm$, Z^0, etc.

For a massive spin-1 particle, $M \neq 0$ so we can go to the rest frame. A spin-1 particle has 3 polarisation states, one particular basis is:

$$\boldsymbol{\epsilon}_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad \boldsymbol{\epsilon}_2 = \begin{pmatrix}0\\1\\0\end{pmatrix}, \quad \boldsymbol{\epsilon}_3 = \begin{pmatrix}0\\0\\1\end{pmatrix}$$

Or, we can single out the z-axis to obtain a more common basis:

$$\boldsymbol{\epsilon}(\lambda=1) = -\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\\ 0\end{pmatrix}, \quad \boldsymbol{\epsilon}(\lambda=-1) = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -i\\ 0\end{pmatrix}, \quad \boldsymbol{\epsilon}(\lambda=0) = \begin{pmatrix}0\\ 0\\ 1\end{pmatrix}.$$

These represent left circular polarisation around z, right circular polarisation and longitudinal polarisation along z respectively.

This basis is orthonormal: $\boldsymbol{\epsilon}(\lambda)\cdot\boldsymbol{\epsilon}(\lambda') = \delta_{\lambda\lambda'}$ for $\lambda = -1, 0, 1$.

To construct the corresponding 4-vectors $\epsilon^\mu(\lambda)$, set $\epsilon^0(\lambda) = 0$ in the rest frame, and define $\epsilon^\mu(\lambda) = \big(\epsilon^0(\lambda), \boldsymbol{\epsilon}(\lambda)\big)$.

In the rest frame, 4-momentum $q^\mu = (M, \mathbf{0})$ so $q_\mu \epsilon^\mu(\lambda) = 0$.

For a boost in the z-direction,

$$\begin{aligned} x'^\mu &= \Lambda^\mu{}_\nu x^\nu \\ x'^0 &= \gamma(x^0 + \beta x^3) \\ x'^3 &= \gamma(x^3 + \beta x^0) \\ x'^1 &= x^1, \quad x'^2 = x^2, \end{aligned}$$

so the ϵ^μ transform as:

$$\begin{aligned} \epsilon'^\mu(\lambda=\pm 1) &= \epsilon^\mu(\lambda = \pm 1) \\ \epsilon'^\mu(\lambda = 0) &= (\gamma\beta, 0, 0, \gamma) = \tfrac{1}{M}(|\mathbf{p}|, 0, 0, E), \end{aligned} \tag{33}$$

using $\gamma = \frac{E}{M}$, $\beta = \frac{|\mathbf{p}|}{E}$, $\gamma\beta = \frac{|\mathbf{p}|}{M}$.

The 4-momentum transforms as:

$$q'^\mu = (\gamma M, 0, 0, \gamma\beta M) = (E, 0, 0, |\mathbf{p}|) \tag{34}$$

(33) and (34) give the relation:

$$q'_\mu \epsilon'^\mu = 0.$$

The completeness relation for $\{\epsilon^\mu\}$ is:

$$\sum_{\lambda=0,\pm1} \epsilon^\mu(p,\lambda)\epsilon^{\nu *}(p,\lambda) = -g^{\mu\nu} + \frac{p^\mu p^\nu}{M^2},$$

and is needed in calculating $|\mathscr{M}|^2$ with external spin-1 particles.

14.2 Massless particles

For example, γ, etc.

The Maxwell equation is:

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

15 Gauge invariance and the gauge principle

15.1 Invariance of Lagrangian

The electron-photon Lagrangian is:

$$\mathscr{L}(x) = \overline{\psi}(x)(i\partial\!\!\!/ + eA\!\!\!/(x) - m)\psi(x) - \tfrac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x)$$

Just like in QM, this $\mathscr{L}$ is invariant under the combined transformations:

$$\begin{aligned}
\psi(x) &\longrightarrow \psi'(x) = e^{i\omega(x)}\psi(x) \\
\overline{\psi}(x) &\longrightarrow \overline{\psi}'(x) = e^{-i\omega(x)}\overline{\psi}(x) \\
A_\mu(x) &\longrightarrow A'_\mu(x) = A_\mu(x) + \tfrac{1}{e}\partial_\mu\omega(x)
\end{aligned}$$

This is easily verified. For the first term:

$$\begin{aligned}
\overline{\psi}'(x)(i\partial\!\!\!/ + eA\!\!\!/')\psi'(x) &= \overline{\psi}(x)e^{-i\omega(x)}\gamma^\mu\left[i\partial_\mu + eA_\mu(x) + \partial_\mu\omega(x)\right]e^{i\omega(x)}\psi(x) \\
&= \overline{\psi}(x)e^{-i\omega(x)}\gamma^\mu\Big[-\partial_\mu\omega(x)e^{i\omega(x)} + e^{i\omega(x)}i\partial_\mu + eA_\mu(x)e^{i\omega(x)} + \\
&\qquad\qquad + \partial_\mu\omega(x)e^{i\omega(x)}\Big]e^{i\omega(x)}\psi(x) \\
&= \overline{\psi}(x)\gamma^\mu(i\partial_\mu + eA_\mu(x))\psi(x). \quad \checkmark
\end{aligned}$$

For the second term:

$$\begin{aligned}
F'_{\mu\nu}(x) &= \partial_\mu A'_\nu(x) - \partial_\nu A'_\mu(x) \\
&= \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) + \tfrac{1}{e}(\partial_\mu\partial_\nu\omega(x) - \partial_\nu\partial_\mu\omega(x)) \\
&= \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) \\
&= F_{\mu\nu}(x). \quad \checkmark
\end{aligned}$$

$e^{i\omega(x)} \in \mathrm{U}(1)$, the unit circle in the complex plane, so this is called a U(1) gauge theory.

15.2 Inverting the logic

Now we start with the free Lagrangian density:

$$\mathscr{L}(x) = \overline{\psi}(x)(i\partial\!\!\!/ - m)\psi(x)$$

which is invariant under $\psi \to e^{i\omega}\psi$, $\overline{\psi} \to e^{-i\omega}\overline{\psi}$ with constant ω – a global U(1) invariance.

Now promote this global invariance to a local one by replacing $\omega \to \omega(x)$. The free $\mathscr{L}$ is *not* invariant:

$$\overline{\psi}(x)(i\partial\!\!\!/ - m)\psi(x) \longrightarrow \overline{\psi}(x)(i\partial\!\!\!/ - \partial\!\!\!/\omega(x) - m)\psi(x).$$

This is because, for $\omega = \omega(x)$, in general:

$$\partial_\mu \psi \not\to e^{i\omega} \partial_\mu \psi$$

So introduce the **covariant derivative:** $D_\mu = \partial_\mu + iqA_\mu(x)$, and demand that:

$$\text{if } \psi(x) \to e^{i\omega(x)}\psi(x), \quad D_\mu\psi(x) \to e^{i\omega(x)} D_\mu \psi(x). \tag{35}$$

What does this imply?

$$\begin{aligned}
(\partial_\mu + iqA_\mu(x))\psi(x) &\to (\partial_\mu + iqA'_\mu(x))\psi'(x) \\
&= e^{i\omega(x)}\big(\partial_\mu + i(\partial_\mu\omega(x)) + iqA'_\mu(x)\big)\psi(x), \\
&\stackrel{!}{=} e^{i\omega(x)}\left(\partial_\mu + iqA_\mu(x)\right)\psi(x). \\
\Rightarrow i\partial_\mu\omega(x) + iqA'_\mu(x) &\stackrel{!}{=} iqA_\mu(x) \\
\iff A'_\mu(x) &= A_\mu(x) - \tfrac{1}{q}\partial_\mu\omega(x).
\end{aligned}$$

So (35) determines how A_μ transforms.

15.3 Conclusion

We may convert the global U(1) symmetry of the free $\mathscr{L}(x)$ to a local one by introducing a gauge field $A_\mu(x)$ with interacts with the fermion. The same approach can be used for complex scalar fields.

15.4 An observation

$$\begin{aligned}
D_\mu D_\nu \psi(x) &= (\partial_\mu + iqA_\mu(x))(\partial_\nu + iqA_\nu(x))\psi(x) \\
&= \left[\partial_\mu\partial_\nu + iq(\partial_\mu A_\nu) + iqA_\nu\partial_\mu + iqA_\mu\partial_\nu - q^2 A_\mu A_\nu\right]\psi(x).
\end{aligned}$$

Note the symmetric terms. Taking the commutator gives:

$$\begin{aligned}
[D_\mu D_\nu]\psi(x) &= iq(\partial_\mu A_\nu - \partial_\mu)\psi(x) \\
&= iqF_{\mu\nu}\psi(x).
\end{aligned}$$

Since there is now no ∂ acting on ψ, this can be taken as a definition of $F_{\mu\nu}$:

$$F_{\mu\nu} = [D_\mu, D_\nu].$$

15.5 Non-abelian symmetries

Suppose ϕ has 2 components , e.g. a proton and a neutron:

$$\psi = \begin{pmatrix} \psi_n \\ \psi_p \end{pmatrix},$$

where ψ_n and ψ_p are Dirac spinors. The free Lagrangian is:

$$\mathscr{L} = \left(\overline{\psi}_n \, \overline{\psi}_p\right) \begin{pmatrix} i\partial\!\!\!/ - m_n & 0 \\ 0 & i\partial\!\!\!/ - m_p \end{pmatrix} \begin{pmatrix} \psi_n \\ \psi_p \end{pmatrix}$$

SU(2) is the non-abelian group of complex 2×2 matrices U with $U^\dagger U = \mathbb{1}$, $\det U = 1$. When $m_n = m_p = m$, $\mathscr{L}(x)$ has a global SU(2) invariance:

$$\begin{aligned} \psi(x) &\longrightarrow U\psi(x) \\ \overline{\psi}(x) &\longrightarrow \overline{\psi}(x)U^\dagger. \end{aligned}$$

$$\mathscr{L}(x) \to \left(\overline{\psi}_n \, \overline{\psi}_p\right) \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix}^\dagger \begin{pmatrix} i\partial\!\!\!/ - m & 0 \\ 0 & i\partial\!\!\!/ - m \end{pmatrix} \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} \begin{pmatrix} \psi_n \\ \psi_p \end{pmatrix}$$

Or, in a more compact notation,

$$\begin{aligned} \mathscr{L}(x) &\to \overline{\psi} U^\dagger (i\partial\!\!\!/ - m) U \psi \\ &= \overline{\psi} U^\dagger U (i\partial\!\!\!/ - m)\psi \\ &= \overline{\psi}(i\partial\!\!\!/ - m)\psi = \mathscr{L}(x). \quad \checkmark \end{aligned}$$

Now we render U local: $U \longrightarrow U(x)$. This gives:

$$\begin{aligned} \partial_\mu \psi(x) &\to \partial_\mu (U(x)\psi(x)) \\ &= \partial_\mu U(x)\psi(x) + U(x)\partial_\mu \psi(x) \end{aligned}$$

So, using the gauge covariant derivative: $\partial_\mu \longrightarrow D_\mu = \partial_\mu + igA_\mu(x)$,

$$\begin{aligned} D_\mu \psi(x) &\to U(x) D_\mu \psi(x) \\ (\partial_\mu + igA_\mu(x))\psi(x) &\to \left[\partial_\mu U(x) + U(x)\partial_\mu + igA'_\mu(x)\right]\psi(x) \\ &\overset{!}{=} U(x)(\partial_\mu + igA_\mu(x))\psi(x) \\ \Longrightarrow A'_\mu(x) &= U(x)A_\mu(x)U^{-1}(x) + \frac{1}{ig}U(x)\partial_\mu U^{-1}(x). \end{aligned}$$

Note: $\partial_\mu(UU^{-1}) = \partial_\mu \mathbb{1} = 0$, so we can use $(\partial_\mu U)U^{-1} + U\partial_\mu U^{-1} = 0$.

Suppose $A_\mu(x) = 0$. Then:

$$A'_\mu(x) = \frac{1}{ig}U(x)\partial_\mu U^{-1}(x). \tag{36}$$

Any $U \in \mathrm{SU}(2)$ can be written:

$$\begin{aligned} U(x) &= \exp\left[i\sum_{a=1}^{3} \alpha^a(x)\frac{\tau^a}{2}\right] \\ &\sim \mathbb{1} + i\sum_{a=1}^{3} \alpha^a(x)\frac{\tau^a}{2} \quad \text{as } \alpha \to 0 \end{aligned}$$

Similarly:

$$U^{-1}(x) \sim \mathbb{1} - i \sum_{a=1}^{3} \alpha^a(x) \frac{\tau^a}{2}.$$

Here $\alpha^a(x)$ are real numbers, and τ^a are the Pauli matrices – a basis for SU(2).

NB:

$$\partial_\mu \exp\left[i \sum_a \alpha^a(x) \frac{\tau^a}{2}\right] \neq \left(i \sum_a \partial_\mu \alpha^a \frac{\tau^a}{2}\right) \exp\left[i \sum_a \alpha^a(x) \frac{\tau^a}{2}\right]$$

because the τ^a do not commute.

But we can use:

$$\partial_\mu e^M = \int_0^1 ds\, e^{sM} \partial_\mu M e^{(1-s)M}$$

Now look at A'_μ in (36).

$$\begin{aligned} A'_\mu(x) &= \frac{1}{ig}\left(1 + i\alpha^a \frac{\tau^a}{2}\right) \partial_\mu \left(1 - i\alpha^b \frac{\tau^b}{2}\right) \\ &= \frac{1}{ig}\left(-i\partial_\mu \alpha^b(x) \frac{\tau^b}{2}\right) \\ &= -\frac{1}{g} \partial_\mu \alpha^a(x) \frac{\tau^a}{2}, \end{aligned}$$

up to $O(\alpha^2)$. So A'_μ is Hermitian and traceless.

In fact, this can be obtained without using an explicit representation for U:

$$\begin{aligned} A &= \frac{1}{ig} e^{-M} \partial_\mu e^M \\ &= \frac{1}{ig} \int_0^1 ds\, e^{(s-1)M} \partial_\mu M e^{(1-s)M} \\ \Rightarrow A^\dagger &= -\frac{1}{ig} \int_0^1 ds\, e^{(1-s)M^\dagger} \partial_\mu M^\dagger e^{(s-1)M^\dagger} \\ &= -\frac{1}{ig} \int_0^1 ds\, e^{(s-1)M} (-\partial_\mu M) e^{(1-s)M} = A. \end{aligned}$$

$$\begin{aligned} \operatorname{tr}(A) &= \frac{1}{ig} \int_0^1 ds\, \operatorname{tr}\left(e^{(s-1)M} \partial_\mu M e^{(1-s)M}\right) \\ &= \frac{1}{ig} \operatorname{tr}(\partial_\mu M) = 0. \end{aligned}$$

An SU(2) gauge field $A_\mu(x)$ is a traceless Hermitian matrix field. This means that for each x, $A_\mu(x) \in \mathfrak{su}(2)$, the **Lie algebra** of SU(2).

A Lie algebra is a linear vector space combined with a **Lie bracket** – a commutator. A basis for this Lie algebra is $T^a = \frac{\tau^a}{2}$, $a = 1, 2, 3$. The scalar product $(T^a, T^b) = 2\,\mathrm{tr}(T^a T^b) = \delta_{ab}$.

The Lie bracket $[T^a, T^b] = i f^{abc} T^c$, where f^{abc} are real constants called the **structure constants**. We know that:

$$\left[\frac{\tau^a}{2}, \frac{\tau^b}{2}\right] = i\epsilon^{abc}\frac{\tau^c}{2},$$

so for $\mathfrak{su}(2)$, $f^{abc} = \epsilon^{abc}$.