Analogies in Physics

Noel Vincze

May 2026

1 Kepler-Coulomb potential

This is perhaps one of the earliest and easiest examples that is demonstrated in physics. This electrogravitational analogy is so fundamental that we are taught about it in high school. Both assume a static solution to the field equations. Maxwell equations (which are the general time-dependent equations of electrodynamics) read

$$\begin{aligned}
\nabla \cdot \mathbf{E} &= \frac{\varrho}{\varepsilon_0} \\
\nabla \times \mathbf{E} + \partial_t \mathbf{B} &= 0 \\
\nabla \cdot \mathbf{B} &= 0 \\
\nabla \times \mathbf{B} &= \mu_0 \mathbf{J} + \varepsilon_0 \mu_0 \partial_t \mathbf{E},
\end{aligned}$$

while the static equations, for the electric field $\mathbf{E}$ (a similar one can be derived for the magnetic field $\mathbf{B}$ with the very same static limit) read

$$\begin{aligned}
\nabla \cdot \mathbf{E} &= \frac{\varrho}{\varepsilon_0} \\
\nabla \times \mathbf{E} &= 0.
\end{aligned}$$

Similarly, the static gravitational equations take the analogous form of

$$\begin{aligned}
\nabla \cdot \mathbf{g} &= 4\pi G \varrho \\
\nabla \times \mathbf{g} &= 0,
\end{aligned}$$

where $\mathbf{g}$ is the gravitational field, and G is the Newton-Cavendish constant. The static equations' solutions are quite straightforward, a Poisson equation really, and have the simple form of

$$\phi(r) = \frac{1}{4\pi\varepsilon_0} \frac{Ze^2}{r}, \tag{1}$$

which is the Coulomb potential, while the Kepler reads

$$V(r) = -G\frac{Mm}{r}. \tag{2}$$

So we can see that these two systems are not only analogous to each other but also effectively identical. What really distinguishes them is that one acts on electric charges (the Coulomb potential) while the other acts on gravitational charges, also called mass (the Kepler potential). It is worthwhile to generalise (Eqs. 1 and 2) to a single form because from that we can draw a conclusion for both equations without repeating them. The generalising form of these potentials reads

$$V(r) = -\frac{k}{r}. \tag{3}$$

Now, generally, we can argue the following about central potentials:

$$V(r) = \begin{cases} \text{If } k > 0 \text{ the system is } \textbf{attractive}. \\ \text{If } k < 0 \text{ the system is } \textbf{repulsive}. \end{cases}$$

Thus, the value of the constants is quite important because it can distinguish the attractive and repulsive potentials. In gravity's case, the constant k is positive (in eq. 3), thus creating an attractive system, while for the Coulomb system to be attractive, we need opposite charges. However, in general, the Coulomb potential has a negative constant k (in eq. 3), thus creating a repulsive system for charges. This is one of the greatest and most fundamental differences between gravity and electricity. While the Coulomb potential can be attractive and repulsive simultaneously because in Nature there exist positive and negative charges, gravity has this symmetry broken by having only positive masses in the Universe. While negative masses or more generally negative energies have been theorised, the cold fact of Nature so far is that gravity will be fundamentally different because it will be able to attract only the semi-positive masses. Nonetheless, if one day we discover negative masses, then in the static case, the Kepler potential would become identical to the Coulomb potential. Finally, the analogy breaks down at relativistic regimes because the two theories are governed by different fundamental theories. namely Quantum Electrodynamics and General Relativity.

Similar results can be derived for static magnetic fields, for example, in the case of cryogenics (there the definition of the magnetic flux $\mathbf{B}$ changes from $\mathbf{B} = \nabla \times \mathbf{A}$ to $\mathbf{B} = -\nabla\phi$) and as well for topological charges such as soft condensed matter. Also we have to mention the string force which is much more alike to electricity than gravity is (because it has both types of colour charges) but it is fundamentally different because it only acts on short ranges $V(r) = -\frac{k}{r}e^{-\lambda r}$ and does not have the inifinite range of the other two interactions

Both theories, that is, gravity and electricity, are invariant under $SO(3)$; thus, they have invariant biaxial rotational vectors like the Laplace-Runge-Lenz (LRL) vector defined as

$$\mathbf{A} = \mathbf{p} \times \mathbf{L} - \frac{k}{r}\hat{\mathbf{r}}.$$

Classically, $\mathbf{A}$ is a constant vector aligned directly along the major axis of the orbit, pointing from the force centre to the perihelion (closest approach). The LRL vector will (Poisson) commute with the Hamiltonian in both the classical and quantum theories. Thus, this allows us to solve complex problems algebraically without ever needing to solve any differential equations like Newton's or Schrödinger's. By crafting new quantities from the LRL vector and the angular momentum defined as $\mathbf{J}_\pm = \frac{1}{2}\left(\mathbf{L} \pm \frac{1}{k}\mathbf{A}\right)$, these new quantities will now form a commuting algebra that under pins these new invariants. This demonstrates that the true spectrum-generating dynamical symmetry group for bound states in both systems is the $SO(4) \cong SU(2) \times SU(2)$. In addition, $SO(4)$ symmetry allows us to solve complex problems algebraically and even extend it to time-independent perturbation theory. Finally, a similar group construction that underpins the Poincaré group, which allows us to find Casimir invariants of particle physics, namely the four-momentum relationship and the Pauli-Lébanski four-vector.

2 Multipole expansion

The multipole expansion is important because the potentials for which you solve the field equations frequently have no physical significance, and it is hard to interpret what they mean. Using a multipole expansion, we reinterpret the potentials with low-energy constants (LECs) such as monopoles, dipoles, quadrupoles, and so on. For the purpose of a general argument, we write the scalar potential as

$$V(x) = k \int \frac{\varrho(y)}{|x - y|}\, d^3y. \tag{4}$$

Using the law of cosines, the numerator can be rewritten. Then, using Legendre polynomial expansion[1], we can write the numerator as an infinite series

$$\frac{1}{|x-y|} = \frac{1}{\sqrt{x^2 + y^2 - 2xy\cos\theta}} = \frac{1}{x}\sum_{k=0}^{\infty}\left[\frac{y}{x}\right]^k P_k(\cos\theta).$$

Thus, the first few terms of the expanded Legendre polynomials of 4 read

[1] We can do this because the law of cosine form of the potential is the generating function for the Legendre polynomials

$$V(x) = \frac{k}{x}\int \varrho(y)\, d^3y + \frac{k}{x^2}\int y\cos\theta\, \varrho(y)\, d^3y + \frac{k}{2x^3}\int y^2[3\cos^2\theta - 1]\, \varrho(y)\, d^3y + \dots. \tag{5}$$

The distinction between electricity and gravity is that there is no gravitational dipole at the centre of mass, which mathematically reads

$$\mathbf{R}_{COM} = \frac{1}{M}\frac{k}{x^2}\int y\cos\theta\, \varrho(y)\, d^3y.$$

When we choose our coordinate origin to sit exactly at the centre of mass, we officially set the dipole moment to equal zero, hence we get

$$M\mathbf{R}_{COM} = \frac{k}{x^2}\int y\cos\theta\, \varrho(y)\, d^3y = 0.$$

In gravity, the mass dipole term can always be mathematically eliminated simply by centring the coordinate origin on the source's centre of mass. In electrostatics, you can only eliminate the electric dipole via an origin shift if the net charge q is non-zero.

The magnetic multipole expansion is different (to Eq. 5) as well because it does not have a monopole or any odd-numbered poles. The exact Legendre polynomial of the multipole expansion reads

$$\mathbf{A}(x) = \frac{k}{x}\oint \mathbf{J}(y)\, d^3y + \frac{k}{x^2}\int y\cos\theta\, \mathbf{J}(y)\, d^3y + \frac{k}{2x^3}\int y^2[3\cos^2\theta - 1]\, \mathbf{J}(y)\, d^3y + \dots. \tag{6}$$

In the magnetic potential case, k reads $k = \mu_0/4\pi$. An important difference between the Legendre expansion for $V(x)$ and $\mathbf{A}(x)$ is that there is no monopole contribution for the magnetic field (Eq. 6) because the loop integral vanishes, i.e. we have

$$\oint \mathbf{J}(y)\, d^3y = 0.$$

In addition, this argument can be extended to an odd number of multipoles as well. On the other hand, from probability calculations, the Legendre expansion serves as a moment series. To better see this, we write the definition of the first four moments. The one variable definitions read

$$\mathrm{E}[\mathrm{X}] = \int xp(x)\,dx \quad \text{1st moment: Mean}$$
$$\mathrm{E}[\mathrm{X}] = \int (x-\mu)^2 p(x)\,dx \quad \text{2nd moment: Variance}$$
$$\mathrm{E}[\mathrm{X}] = \int (x-\mu)^3 p(x)\,dx \quad \text{3rd moment: Skewness}$$
$$\mathrm{E}[\mathrm{X}] = \int (x-\mu)^4 p(x)\,dx \quad \text{4th moment: Kurtosis}$$

where μ is the expectation value. The more general tensor definitions read of these quantities read

$$\mathrm{E}[\mathrm{X}_i] = \int x_i p(x)\,dx \quad \text{1st moment: Mean}$$
$$\mathrm{E}[\mathrm{X}_i\mathrm{X}_j] = \int (x_i-\mu_i)(x_j-\mu_j) p(x)\,dx \quad \text{2nd moment: Variance}$$
$$\mathrm{E}[\mathrm{X}_i\mathrm{X}_j\mathrm{X}_k] = \int (x_i-\mu_i)(x_j-\mu_j)(x_k-\mu_k) p(x)\,dx \quad \text{3rd moment: Skewness}$$
$$\mathrm{E}[\mathrm{X}_i\mathrm{X}_j\mathrm{X}_k\mathrm{X}_l] = \int (x_i-\mu_i)(x_j-\mu_j)(x_k-\mu_k)(x_l-\mu_l) p(x)\,dx \quad \text{4th moment: Kurtosis}$$

Thus, we can see that there is a one-to-one correspondence between probabilistic moments and multipoles. Moments are also important because they are a great tool in effective theories, for example, when we derive the tensor virial theorem due to Chandrasekhar. These types of moment equations naturally occur in statistical mechanics, and we can use them to derive the continuous Navier-Stokes equation, for instance, and it is heavily used in astrophysics. Furthermore, these moments can be found in quantum mechanics as well, for instance, when one calculates the Ehrenfest correspondence principle.

3 Elliptic tensors

The moment of inertia tensor is defined as

$$I_{ij} = \int (x_k^2\delta_{ij} - x_i x_j)\varrho(x)\,d^3x. \tag{7}$$

In matrix form, we have

$$\mathbf{I} = \begin{bmatrix} \int (y^2+z^2)\varrho\,d^3x & -\int xy\varrho\,d^3x & -\int xz\varrho\,d^3x \\ -\int yx\varrho\,d^3x & \int (x^2+z^2)\varrho\,d^3x & -\int yz\varrho\,d^3x \\ -\int zx\varrho\,d^3x & -\int zy\varrho\,d^3x & \int (x^2+z^2)\varrho\,d^3x \end{bmatrix}$$

The mathematical role of the moment of inertia is to find the Principal Axes of Rotation, where cross-products of inertia vanish, stopping a spinning body from wobbling, as well as measuring directional resistance to rotational acceleration around axes. It is used in rigid body dynamics, but when we consider single particles without occupying any physical space, it simplifies to

$$I = \int x_k^2 \varrho(x)\, d^3x = m\, x_k^2.$$

where m is just the mass of the particle. This simplified form can be found in spherical coordinate expansions, and this simplified form essentially tells us that the particle is incapable of internal rotations (in the classical sense, of course). Mathematically, we get the scalar moment of inertia by projecting the moment of inertia tensor to its normal planes by $I = I_{ij}\hat{n}_i\hat{n}_j$, which means that the object is a simple, symmetrical sphere (like a planet or a basketball), and the distribution of mass is completely uniform in all directions. Therefore, the moment of inertia (Eq. 7) is the same regardless of the angle at which you spin it. However, for complex, asymmetrical shapes (like a dumbbell or an aeroplane), spinning the object around the x-axis encounters a different amount of resistance than spinning it around the y-axis. The tensor allows you to calculate the moment of inertia for any 3D angle by defining the angle's direction $\hat{n}_i$ and using the tensor to project the resistance.

From the electromagnetic multipole expansion, say we take the electrostatic quadrupole, and we see that it is similarly defined to the moment of inertia, but there is an important constraint on it. The definition reads

$$Q_{ij} = \int (3x_i x_j - x_k^2 \delta_{ij}) \varrho(x)\, d^3x. \tag{8}$$

The addition of that 3 has an important consequence, namely that it makes the quadrupole tensor Q_{ij} in Eq. 8 traceless (while Eq. 7 is not traceless). Its geometric role is to align the system with the Principal Axes of Elongation, reducing the field to its main prolate or oblate deformation. Furthermore, it measures the trace-free spatial deviation of a charge from spherical symmetry. The Electric Quadrupole Tensor Q is forced by definition to be completely trace-free ($\mathrm{Tr}\, Q_{ij} = 0$). This isolates pure shape asymmetry from the total volumetric charge. Because its trace is zero, its eigenvalues must satisfy $\lambda_1 + \lambda_2 + \lambda_3 = 0$, meaning it can have negative eigenvalues even though it describes a physical distribution.

In continuum mechanics, we have an important quantity called the Cauchy stress tensor. You do not have to be stressed to understand this tensor. There are two important cases which have an analogy to the above-described tensor. Firstly, for a purely mechanical system, it is when, besides the isotropic pressure, we also have shear pressure in the system, which is described by the following stress tensor

$$\sigma_{ij} = p(r)\, \delta^{ij} + s(r) \left[\frac{x_i x_j}{r^2} - \frac{1}{3}\delta^{ij} \right].$$

In the isotropic case, we have

$$\sigma_{ij} = p(r)\,\delta^{ij}. \tag{9}$$

The trace of this tensor is intrinsically unconstrained and usually serves as a definition of (isotropic) pressure, which reads

$$p(x) \equiv \frac{1}{3}\sigma_{kk}.$$

And because of this relation and how the stress-energy tensor (SET) is structured (Eq. 9), we have the famous relativistic equation of state, which reads

$$p = \frac{1}{3}\varrho.$$

This is the reason why it is so hard to define certain properties of particles or systems in field theory, such as mass, energy, and momentum, because they can have contributions from not just energy density or momentum density but from pressures and energy fluxes as well. The geometrical duty of the Cauchy stress tensor is to pinpoint the Principal Planes of Stress, where all shear stresses drop to zero, leaving purely normal tension or compression.

However, if we take non-mechanical definitions of stress tensor, such as magnetic, electrical or gravitational, we find the following form

$$\sigma_{ij} = k[F_i F_j - \frac{1}{2}F_k^2 \delta_{ij}], \tag{10}$$

where k is the appropriate constant for the given interaction, while F_i is the physical field considered from the above fundamental interaction. When we take the trace of the above interactions, especially in the relativistic limit, we find that the trace of the tensor (Eq. 10) vanishes, meaning the photon (or the graviton) is a massless particle, i.e., it is not described by the Procca equation of motion. Furthermore, it has to be clarified that the above statement is true only outside of the superconductivity regime of condensed matter, because there we have the London gauge, which makes the photon massless, with an effective mass. On the other hand, if one calculates the trace of the W or Z boson SET, one will find that the trace is constrained by the mass of that particle.

As mentioned in the multipole expansion, calculating nth moments will generate a rank-n tensor. In the cases of (co)variance, we have

$$\mathrm{E}[\mathrm{X}_i \mathrm{X}_j] = \int (x_i - \mu_i)(x_j - \mu_j) p(x)\, dx.$$

which clearly have the same form as the above-described physical tensors. The variance tensor allows for performing Principal Component Analysis (PCA), rotating the coordinate base to find the axes of completely uncorrelated data variance.

Mathematically, these are real symmetric rank-2 tensors which define an ellipsoid by $T_{ij}x_ix_j = 1$, where T_{ij} can be any of the above tensors. Calculating the principal invariants of this tensor T_{ij} by means of either diagonalising or solving the eigenvalue equation gives us the principal invariants, which, in the case of 3×3 matrices, are the trace, the adjugate and the determinant. The trace is special and, in many cases, can be either constrained or unconstrained. For the moment of inertia, it is unconstrained, but for the quadropole tensor, it is strictly equated to 0. The more interesting case is the stress tensor because the pressure is defined by the trace, thus allowing that even electric and magnetic fields have pressure, also called photon pressure. The other invariants can produce similar physically and mathematically invariant quantities. I would like to emphasise that all of these rank-2 tensors have mathematical similarity to the adjugate structure. Geometrically, they transform under the $SO(3)$ group, and every one of these tensors defines a 3D quadratic form mapping a spatial vector x_i to a constant value $T_{ij}x_ix_j = 1$, producing an ellipsoid. For the moment of inertia, it generates the Poinsot Ellipsoid, for the Cauchy Stress tensor, it generates the Lamé Stress Ellipsoid, while for the Covariance matrix, it generates the Error Contour Ellipsoid of equal probability density.

4 Ladder operators

Ladder operators have become a popular tool in quantum physics because they allow us to solve differential equations essentially algebraically. However, even in the classical Hamilton-Poisson system, we can have and define these ladder operators by means of canonical coordinate transformations. However, it is important to point out that in the classical case, we have Poisson commutativity, and the resulting energy will be different from the quantum case. Thus, since they are more important in quantum mechanics, we will focus on them and point out how they differ from the classical results. The harmonic oscillator in quantum mechanics reads its Hamiltonian as

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2. \tag{11}$$

The ladder operators read for the harmonic oscillator

$$\begin{aligned} a &= \frac{1}{\sqrt{2m\omega\hbar}}p + i\sqrt{\frac{m\omega}{2\hbar}}x \\ a^\dagger &= \frac{1}{\sqrt{2m\omega\hbar}}p - i\sqrt{\frac{m\omega}{2\hbar}}x \end{aligned}$$

In the classical case (Eq. 11), instead of $\hbar\omega$, we would use the action I, which connects to quantum mechanics by the Bohr-Sommerfeld quantisation rule. The Hamiltonian then reads

$$H = \hbar\omega \left(a^\dagger a + \frac{1}{2} \right) \tag{12}$$

The key difference to the classical result is that in the Hamilton-Poisson system we have no $\hbar\omega/2$, which means that the classical system has no zero-point energy (as seen in Eq. 12), thus a lower bound for the system. Furthermore, in quantum mechanics, we can also have ladder operators for radial motion, such as the isotropic oscillator and the hydrogen atom. This is only possible because the physical quantities themselves correspond to differential operators, such as momentum, energy and angular momentum. In the classical case, these radial ladder operators are not possible because the Poisson commutation relations are just imposed upon the classical variables. Regardless, the quantum radial oscillator Hamiltonian reads

$$H = \frac{p_r^2}{2m} + \frac{\hbar\ell(\ell+1)}{2mr} + \frac{1}{2}m\omega^2 r^2.$$

while the radial Kepler-Coulomb problem reads

$$H = \frac{p_r^2}{2m} + \frac{\hbar\ell(\ell+1)}{2mr} - \frac{Ze^2}{r},$$

which is used for the hydrogen, hydrogenlike systems and helium and its like systems with some time-independent perturbation tossed into calculate the energy levels. The ladder operators for radial motion, especially the Hydrogen problem, read

$$\begin{aligned} a &= p_r + \frac{i\hbar(\ell+1)}{r} - \frac{iZme^2}{\hbar(\ell+1)} \\ a^\dagger &= p_r - \frac{i\hbar(\ell+1)}{r} + \frac{iZme^2}{\hbar(\ell+1)} \end{aligned}$$

Then, performing diagonalisation, the Hamiltonian takes the form of

$$H = \frac{a^\dagger a}{2m} + \frac{Z^2 me^4}{2\hbar^2(\ell+1)^2}. \tag{13}$$

Which, as mentioned above, is completely a quantum result. While the commutation results of the radial ladder operators can be reproduced with the Poisson algebra, the Hamiltonian cannot take this simple form that the quantum case does because of how the hyperbolic or elliptic terms mix together, as mentioned above, which is a completely unique result of quantum mechanics (and thus of Eq. 13). Meanwhile, the radial ladder operators for the isotropic oscillator read

$$a = p_r + \frac{i\hbar(\ell+1)}{r} - im\omega r$$
$$a^\dagger = p_r - \frac{i\hbar(\ell+1)}{r} + im\omega r$$

The Hamiltonian reads

$$H = \frac{a^\dagger a}{2m} - \hbar\omega$$

When we compare it to the 1D harmonic oscillator (Eq. 12), the structure is the same in the sense that we have the operator part plus a constant, which is the zero-point energy, but it is different in the dimensions of the radial and 1D ladder operators. Generally, we can say that for any radial or central potential $V(r)$ can be factored into ladder operators, such as the Wood-Saxon potential, Morse potential, Hulthén Potential, Eckart Potential and the Pöschl-Teller potential.

Ladder operators are not only used to diagonalise the Hamiltonian, but they can also be applied to Casimir invariants, such as the squared angular momentum, as well. Here, the notation changes a bit, but the raising and lowering interpretation of the given physical variables' values, in this case, the angular momentum, will not change. In the angular momentum reads

$$L^2 = L_x^2 + L_y^2 + L_z^2.$$

We can diagonalise it as

$$L^2 = L_+ L_- + L_z^2. \tag{14}$$

In spherical coordinates, the angular momentum's ladder operators read

$$L_\pm = \hbar e^{\pm i\phi}(p_\phi \cot\theta \mp ip_\theta).$$

When we impose these operators on the wavefunction, they will naturally generate the spherical harmonics $Y_m^\ell(\theta, \phi)$, which is so much harder to obtain from the second-order radial Bessel equation. Their Cartesian definitions will read

$$L_+ = L_x + iL_y$$
$$L_- = L_x - iL_y$$

Because the spatial components of angular momentum (Eq. 14) do not commute $[L_x, L_y] = i\hbar L_z$, a state with a well-defined projection along the z-axis $L_z = m\hbar$

has completely indeterminate values for L_x and L_y. Classically, the vector $\mathbf{L}$ points along a cone of height $m\hbar$ with a total slant length of $\sqrt{l(l+1)}\hbar$. The ladder operators describe a spatial tilting of angular-momentum eigenvalues. The operators $L_\pm = L_x \pm i L_y$ do not change the length of the vector L^2. Instead, they step the state up or down the cone. The raising operator (L_+) shifts the spatial orientation of the system closer to the positive z-axis, increasing the projection by $1\hbar$ while the lowering operator (L_-) shifts the orientation away from the positive z-axis, decreasing the projection by $1\hbar$.

The final analogy can be seen in the form of the Zeeman effect. The Zeeman effect is the magnetic counterpart of the Stark effect, and its mathematical form is given by

$$H = -\gamma\, \mathbf{S} \cdot \mathbf{B}. \tag{15}$$

This Hamiltonian depends on the spin $\mathbf{S}$ instead of the more traditional angular momentum $\mathbf{L}$, hence it is used to describe the hyperfine splitting of the energy levels as well to describe the anomalous magnetic moment of fermions in the non-relativistic limit. While its original form is just a potential that we add to the Schrödinger equation, the spin-dependent version can be elegantly derived from the Pauli equation, which introduces a Dirac-like operator in the non-relativistic limit, that is, we have $\boldsymbol{\sigma} \cdot (-i\hbar\nabla - e\mathbf{A})$, thus when we calculate the square of this operator, we get back the Zeeman potential (Eq. 15). When we diagonalise it, we have

$$H = -\gamma\, (S_+ B_- + S_- B_+ + S_z B_z). \tag{16}$$

In the classical view of Eq. 16, the fields B_- and B_+ are interpreted as circular polarisation. If the magnetic field is time-dependent (like the driving field $\mathbf{B}_1(t)$ in nuclear magnetic resonance), a static linear transverse field can be mathematically split into two counter-rotating fields. The field B_+ represents a right-handed circularly polarised magnetic field rotating in the xy-plane, while B_- represents a left-handed circularly polarised magnetic field rotating in the xy-plane. Simply said that the circularly polarised waves are chiral. The quantum view of the interaction is that it represents a chiral energy exchange. The cross-terms in the Hamiltonian $B_+ S_-$ and $B_- S_+$ dictate how the field and spin interact with each other. A right-handed field (B_+) is exclusively responsible for driving a spin-lowering transition (S_-). A left-handed field (B_-) exclusively drives a spin-raising transition (S_+). Finally, this maps to the photon helicity, that is, if the magnetic field is fully quantised, $B_\pm$ map directly to the creation and annihilation operators of photons with distinct helicities (spin angular momentum of $\pm 1\hbar$).

5 Electroweak and meson theory

Hideki Yukawa (湯川 秀樹) was the first one to propose particles that intermediate between protons and neutrons. This interaction was later dubbed the strong nuclear interaction, and these intermediating particles were dubbed the mesons. In the SU(2) theory proposed by Yukawa, the pions were to propagate the strong interaction between protons and neutrons. The Lagrangian of this interaction is as follows

$$\mathcal{L} = i\bar{\Psi}\gamma^{\mu}\partial_{\mu}\Psi - \partial^{\mu}\bar{\Phi}\partial_{\mu}\Phi - \bar{\Psi}M_{n,p}\Psi - m_{\pi}^{2}\bar{\Phi}\Phi + i\lambda\bar{\Psi}\gamma_{5}\Phi\Psi,$$

Where the first few terms are the Dirac Lagrangian and Klein-Gordon Lagrangian for the baryons and pions, respectively. The baryon mass matrix is diagona and $M_{n,p} = diag(m_n, m_p)$. The Ψ is a doublet and has the following form

$$\Psi = \begin{bmatrix} n \\ p \end{bmatrix},$$

while the mediator particles the pions (π-meson) Φ are pseudo-scalar particles and take the following forms

$$\Phi = \begin{bmatrix} \pi^0 & \sqrt{2}\pi^- \\ \sqrt{2}\pi^+ & -\pi^0 \end{bmatrix},$$

and they posse all the necessary SU(2) group property. The pseudo-scaler property of the pions can also be seen from $i\gamma_5$ in front of the pion field Φ. Hence, this Yukawa term/ Yukawa Lagrangian is

$$\mathcal{L}_{Yuk} = i\lambda\bar{\Psi}\gamma_5\Phi\Psi. \tag{17}$$

We can also include the Hermitian conjugate of the Yukawa term to complete our Lagrangian. With the added term, we get

$$\begin{aligned} \mathcal{L}_{Yuk} &= i\lambda\bar{\Psi}\gamma_5\Phi\Psi - i\lambda\Psi\gamma_5\bar{\Phi}\bar{\Psi} \\ &= i\lambda\bar{\Psi}\gamma_5\Phi\Psi + \text{h.c.}. \end{aligned}$$

Although typically the complex conjugate term is relegated to the 'h.c.' expression. Furthermore, if we couple the previously presented Lagrangian, then we get the following with the given extra contributions

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + i\bar{\Psi}\gamma^{\mu}D_{\mu}\Psi - D^{\mu}\bar{\Phi}D_{\mu}\Phi - \bar{\Psi}M_{n,p}\Psi - m_{\pi}^{2}\bar{\Phi}\Phi + i\lambda\bar{\Psi}\gamma_5\Phi\Psi. \tag{18}$$

where the D_μ is the covariant derivative, and it is defined as

$$D_\mu = \partial_\mu - ieA_\mu.$$

If we expand the covariant terms, then we are going to get two new terms that correspond to the J^{μ}_{EM} for the electromagnetic current, which describes the flow of positive charges, and J^{μ}_{had} for the hadronic current, which describes the flow of charged intermediate particles. The expanded Lagrangian will be

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + i\bar{\Psi}\gamma^{\mu}\partial_{\mu}\Psi - \partial^{\mu}\bar{\Phi}\partial_{\mu}\Phi - \bar{\Psi}M_{n,p}\Psi - m_{\pi}^{2}\bar{\Phi}\Phi \\ + e\bar{\Psi}\gamma^{\mu}\Psi A_{\mu} + e[\bar{\Phi}\partial^{\mu}\Phi - \partial^{\mu}\bar{\Phi}\Phi - eA^{\mu}\bar{\Phi}\Phi]A_{\mu} + i\lambda\bar{\Psi}\gamma_{5}\Phi\Psi -$$

The two identified currents will take the form of

$$J^{\mu}_{EM} = e\bar{\Psi}\gamma^{\mu}\Psi$$
$$J^{\mu}_{had} = e[\bar{\Phi}\partial^{\mu}\Phi - \partial^{\mu}\bar{\Phi}\Phi - eA^{\mu}\bar{\Phi}\Phi].$$

Overall, Yukawa imagined this interaction (Eq. 17) that we have two massive baryons (the neutron n, and the proton p) and the three massive pseudo-scalar mesons (the pions: π^{+}, π^{-}, π^{0}). The pions transfer the nuclear force between the nuclei. As our accelerators got more powerful, our experimental evidence grew, and new theories arose. Nowadays, pions and other mesons are the only effective way to approximate the QCD interactions into an effective point-wise interaction through the intermediate particles.

However, this was only a linear theory of the mesons (Eq. 18). The fully non-linear theory and its non-linear self-interactions are captured by Chiral Perturbation theory, which introduces a new non-linear field U which depends on Φ or on $\boldsymbol{\pi}$, depending on which notation one uses. The Lagrangian will now read

$$\mathcal{L} = \frac{F_{\pi}^{2}}{4}\,\mathrm{Tr}[(D_{\mu}U)(D^{\mu}U)^{\dagger}] + \frac{F_{\pi}^{2}}{4}\,\mathrm{Tr}[\chi^{\dagger}U + U^{\dagger}\chi], \tag{19}$$

where χ is the chiral matrix, that is, it transforms according to the chiral transformations. In the simple scalar case, it will reduce to the mass matrix ($\chi = 2B_{0}\mathcal{M}$, where $\mathcal{M}$ is a diagonal matrix of quark masses), [2] and in this case, it loses its chiral transformation property. The $U(x)$ field [3] is our main variable and has the following standard form

$$U(x) = \exp\left(\frac{\sqrt{2}i}{F_{\pi}}\pi^{a}T^{a}\right).$$

The simplest group is the $SU(2)$, which only contains pions:

$$\pi^{a} = \begin{bmatrix} \frac{1}{\sqrt{2}}\pi^{0} & \pi^{+} \\ \pi^{-} & -\frac{1}{\sqrt{2}}\pi^{0} \end{bmatrix}. \tag{20}$$

[2] $\mathcal{M} = diag(m_{u}, m_{d}, m_{s}, m_{c})$, including the $SU(4)$ group

[3] It transforms as $U(x) \to g_{L}U(x)g_{R}^{\dagger}$

While for the $SU(3)$ case, we also got the kaons and the etas (η' is not present in this matrix because it is added in differently and its mass is also more unique (understand heavier) than the normal η meson). The matrix takes the form of

$$\pi^a = \begin{bmatrix} \frac{1}{\sqrt{2}}\pi^0 + \frac{1}{\sqrt{6}}\eta & \pi^+ & K^+ \\ \pi^- & -\frac{1}{\sqrt{2}}\pi^0 + \frac{1}{\sqrt{6}}\eta & K^0 \\ K^- & \bar{K}^0 & -\frac{2}{\sqrt{6}}\eta \end{bmatrix},$$

The kinetic part of Eq. 19 reads

$$\mathcal{L} = \frac{F_\pi^2}{4}\,\mathrm{Tr}[(D_\mu U)(D^\mu U)^\dagger],$$

where the covariant derivatives are defined as

$$\begin{aligned} D_\mu U &= \partial_\mu U - iR_\mu U + iUL_\mu \\ D_\mu U^\dagger &= \partial_\mu U^\dagger + iU^\dagger R_\mu - iL_\mu U^\dagger, \end{aligned}$$

where the external fields of R_μ and L_μ denote right- and left-handedly transforming fields. Decomposing Eq. 5 into its three main components, we find the pure kinetic term of the pseudoscalar mesons, the current interaction and finally a Proca mass term, which will be interpreted as the mass of the vector mesons. Thus we have

$$\begin{aligned} \mathcal{L}_\chi &= \frac{F_\pi^2}{4}\,\mathrm{Tr}[(\partial_\mu U + igV_\mu U - igUV_\mu)(\partial^\mu U^\dagger + igV_\mu U^\dagger - igU^\dagger V_\mu)] \\ &= \frac{F_\pi^2}{4}\,\mathrm{Tr}[\partial_\mu U \partial^\mu U^\dagger + 2igV_\mu(U\partial^\mu U^\dagger + U^\dagger \partial^\mu U) + 2g^2 UU^\dagger V_\mu V^\mu]. \end{aligned}$$

Hence, we can identify the meson current as

$$J_\pi^\mu = 2i(U\partial^\mu U^\dagger + U^\dagger \partial^\mu U). \tag{21}$$

and the Procca mass for vector mesons reads

$$\mathcal{L}_{mass} = \frac{2g^2F_\pi^2}{4}\,\mathrm{Tr}[UU^\dagger V_\mu V^\mu] = \frac{2g^2F_\pi^2}{4}\,\mathrm{Tr}[V_\mu V^\mu]. \tag{22}$$

In the last equation, we can identify the famous Kawarabayashi-Suzuki-Riazuddin-Fayyazuddin (KSRF) relation

$$\mathcal{L}_{\mathrm{VMD}} = -\frac{1}{8}\mathrm{Tr}[V^{\mu\nu}V_{\mu\nu}] + \frac{2g^2F_\pi^2}{4}\,\mathrm{Tr}[V_\mu V^\mu]. \tag{23}$$

In an $SU(2)$ set up the vector meson field V_μ in Eq. 23 takes the form of

$$V_\mu = \begin{bmatrix} \frac{1}{\sqrt{2}}(\rho_\mu^0 + \omega_\mu) & \rho_\mu^+ \\ \rho_\mu^- & -\frac{1}{\sqrt{2}}(\rho_\mu^0 + \omega_\mu) \end{bmatrix},$$

thus mimics the pseudoscalar pions (Eq. 20) quite well, while in $SU(3)$ we have

$$V_\mu = \begin{bmatrix} \frac{1}{\sqrt{2}}\rho^0_\mu + \frac{1}{\sqrt{6}}\omega_\mu + \frac{1}{\sqrt{3}}\phi_\mu & \rho^+_\mu & K^{*+}_\mu \\ \rho^-_\mu & -\frac{1}{\sqrt{2}}\rho^0_\mu + \frac{1}{\sqrt{6}}\omega_\mu + \frac{1}{\sqrt{3}}\phi_\mu & K^{*0}_\mu \\ K^{*+}_\mu & K^{*+}_\mu & -\frac{2}{\sqrt{6}}\omega_\mu + \frac{1}{\sqrt{3}}\phi_\mu \end{bmatrix},$$

On the electroweak side, we have the Glashow-Weinberg-Salam model [1, 2, 3], which shares the $SU(2)$ gauge group that underpins both models. Additionally, both models have exponential (pseudo)Goldstone bosons, but in the case of the electroweak theory, it is rotated away with a unitary gauge transformation, leaving us a linear σ model. The other similarity that arises between the models is that when they are gauged either by vector mesons or by W or Z bosons, these vector bosons gain their mass term in the same way. Now we shall turn our attention to the actual GSW model, so we can see the previously made claims. The model's Lagrangian takes the form of

$$\mathcal{L}_{GSW} = -\frac{1}{4}(W^a_{\mu\nu})^2 - \frac{1}{4}(B_{\mu\nu})^2 + (D_\mu H)^\dagger D^\mu H + m^2 H^\dagger H - \lambda(H^\dagger H)^2.$$

The covariant derivative reads

$$D_\mu H = \partial_\mu H - igW^a_\mu T^a H - \frac{1}{2}ig' B_\mu H. \tag{24}$$

The Englert-Higgs mechanism proceeds as we take the derivative of $\mathcal{L}_{GSW}$ with respect to the Higgs field $H^\dagger$ and minimise it. Let us take a detailed look at the covariant derivative of the Higgs field. This term Eq. 24 actually contain the current interaction with the gauge bosons of W^a_μ and B_μ, and hence when we expand them around the minimum, we get the following results

$$|D_\mu H|^2 = \frac{g^2}{4}\begin{bmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{g'}{g}B_\mu + W^3_\mu & W^1_\mu - iW^2_\mu \\ W^1_\mu + iW^2_\mu & \frac{g'}{g}B_\mu + W^3_\mu \end{bmatrix} \begin{bmatrix} \frac{g'}{g}B_\mu + W^3_\mu & W^1_\mu - iW^2_\mu \\ W^1_\mu + iW^2_\mu & \frac{g'}{g}B_\mu - W^3_\mu \end{bmatrix} \begin{bmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{bmatrix} \tag{25}$$

$$= \frac{g^2 v^2}{8}\left[W^+_\mu W^-_\mu + \left(\frac{g'}{g}B_\mu + W^3_\mu\right)^2\right]. \tag{26}$$

As we stated previously, this generates a mass term for the vector boson fields, making them massive, although contrary to the meson theory, to get the actual physical fields, we have to perform a rotation in isospin space to get the massive Z boson and the massless A photon. Thus, we find that the following expression for the gauge term of the GSW Lagrangian will read

$$\begin{aligned}\mathcal{L}_{GSW} = &-\frac{1}{4}(F_{\mu\nu})^2 - \frac{1}{4}(Z_{\mu\nu})^2 + \frac{1}{2}m_Z^2(Z_\mu)^2 - \frac{1}{2}W^+_{\mu\nu}W^-_{\mu\nu} + \frac{1}{2}m_W^2 W^+_\mu W^-_\mu \\ &- \frac{ie\cot\theta}{2}Z_{\mu\nu}W^+_{[\mu}W^-_{\nu]} - \frac{ie}{2}F_{\mu\nu}W^+_{[\mu}W^-_{\nu]} - ie\cot\theta Z_\mu[W^-_\nu\partial_\mu W^-_\nu - W^+_\nu\partial_\mu W^-_\nu] \\ &- ieA_\mu[W^-_\nu\partial_\mu W^-_\nu - W^+_\nu\partial_\mu W^-_\nu] - \frac{ie\cot\theta}{2}[W^-_{\mu\nu}Z_{[\mu}W^+_{\nu]} - W^+_{\mu\nu}Z_{[\mu}W^-_{\nu]}] \\ &- \frac{ie}{2}[W^-_{\mu\nu}A_{[\mu}W^+_{\nu]}W^+_{\mu\nu}A_{[\mu}W^-_{\nu]}] - \frac{e^2}{2\sin^2\theta}(W^+_\mu W^-_\mu)^2 + \frac{e^2}{2\sin^2\theta}(W^+_\mu)^2(W^-_\mu)^2 \\ &- e^2\cot^2\theta Z_\mu W^-_\nu W^+_{[\mu}Z_{\nu]} + e^2 A_\mu W^-_\nu W^+_{[\mu}A_{\nu]} \\ &+ e^2\cot^2\theta A_\mu Z_\nu W^+_{\{\mu}W^-_{\nu\}} - (W^+_\mu W^-_\mu)(Z_\nu A_\nu),\end{aligned}$$

where for sake of convenience we have introduced $\tan\theta = g'/g$ and for the gauge boson masses are given by

$$m_W = \frac{gv}{2} \tag{27}$$

$$m_Z = \frac{m_W}{\cos\theta}. \tag{28}$$

In Eq. 28 we have the definitions for the physical fields of the vector gauge bosons. On the other hand, the Higgs term will read

$$\begin{aligned}\mathcal{L}_{Higgs} = &-\frac{1}{2}h(\Box + m_h^2)h - g\frac{m_h^2}{4m_W}h^3 - g^2\frac{m_h^2}{32m_W^2}h^4 \\ &+ \left(\frac{2}{v}h + h^2\right)\left[m_W^2 W^+_\mu W^-_\mu + \frac{1}{2}m_Z^2(Z_\mu)^2\right].\end{aligned}$$

which is a nonlinear self-interacting Lagrangian similar to the pions, but distinct in the way that the Higgs is a real field while the pions are pseudo-Goldstone bosons. For the pions, this means that self-interaction terms have to contain some derivatives aligning the pions with Goldstone's theorem. However, since they are only pseudo-particles, they will have a mass term that breaks the chiral symmetry of the model. This is not the case for the Higgs. Another difference that emerges is that the Higgs is an Isospin doublet while the pions are an isospin triplet, thus completely altering how Yukawa terms interact with the fermion fields. Although in recent years it has been floated the idea of a Higgs isospin triplet of the form

$$H = \begin{bmatrix}\phi^+ & \phi^{++} \\ \phi^0 & \phi^+\end{bmatrix}$$

which would be closer to how the pions look like (compare to Eq. 20) but still distinct because it has only positively charged Higgs particles and no negative ones. For leptons, we have

$$L_n = \begin{bmatrix} \nu_e \\ e \end{bmatrix}, \begin{bmatrix} \nu_\mu \\ \mu \end{bmatrix}, \begin{bmatrix} \nu_\tau \\ \tau \end{bmatrix}$$

The fermionic sectors Lagrangian takes the form of with the skewed Yukawa coupling as

$$\begin{aligned} \mathcal{L} = \sum_n i\bar{L}_n^L \not{D}_L L_n^L + i\bar{L}_n^R \not{D}_R L_n^R + i\bar{Q}_n^L \not{D}_L Q_n^L \\ + i\bar{Q}_n^R \not{D}_R Q_n^R - \sum_m Y_{nm}\bar{L}_n^L H L_m^R - Y_{nm}\bar{Q}_n^L H Q_m^R. \end{aligned} \tag{29}$$

Where the right-handed fermion fields are all singlets, so that the left-handed isospin doublet could be contracted with the Higgs doublet. This mathematical mismatch is why a Higgs triplet would be nicer for the Yukawa coupling (Eq. 29), and it would not need to introduce a Majorana-transformed Higgs doublet, so we could give mass to the right-handed quarks.

6 Spontanous symmetry breaking

In condensed matter physics, superconductivity requires that certain symmetries be broken. This phenomenon was first described by the London gauge [4] ($\mathbf{J} = -\frac{ne^2}{m_e}\mathbf{A}$), where n is the electron density, which is a phenomenological description of the process. Through this specific choice of gauge, the Ampere-Maxwell equation changes into

$$\nabla \times \mathbf{B} = -\frac{\mu_0 n e^2}{m_e}\mathbf{A} + \varepsilon_0\mu_0\partial_t\mathbf{E}.$$

Taking the curl on both sides, we get the following Procca equation for the magnetic flux

$$\nabla^2\mathbf{B} - \varepsilon_0\mu_0\partial_{tt}\mathbf{B} + \frac{\mu_0 n e^2}{m_e}\mathbf{B} = 0,$$

or written using the d'Alambert operator, we have

$$\Box\mathbf{B} + \frac{\mu_0 n e^2}{m_e}\mathbf{B} = 0. \tag{30}$$

This equation effectively describes that the photon has an effective mass proportional to $m_\gamma^2 = \mu_0 n e^2/m_e$. The more sophisticated description of this procedure was given by Ginzburg and [5], while the procedure has been given the name of spontaneous symmetry breaking. We start with the free energy functional given by

$$F[\psi] = F_0 + \int d^3x\, \frac{1}{2\mu_0}\mathbf{B}^2 + \frac{1}{4m_e}\left|(-i\hbar\nabla - 2e\mathbf{A})\psi\right|^2 + \alpha_0(T-T_c)\psi^\dagger\psi + \frac{\beta}{2}(\psi^\dagger\psi)^2$$

The operator in the brackets $-i\hbar\nabla - 2e\mathbf{A}$ is the covariant derivative in the presence of magnetic fields. The charge and mass are twice that of the usual Hamiltonian because physically, this free energy describes the Cooper pair, which is a boson made out of two electrons. From a particle physics point of view, this would be called a dilepton condensate. The constant α_0 changes sign when the condensation reaches the critical temperature, while $\beta > 0$ is positive and gives stability for the free energy functional. When $T < T_c$, $\alpha_0 < 0$, this is when it changes sign. Minimising the effective potential $V(\psi) = \alpha_0(T - T_c)\psi^\dagger\psi + \frac{\beta}{2}(\psi^\dagger\psi)^2$ we find the condensation as

$$|\psi_0|^2 = -\frac{\alpha_0}{\beta} \equiv v^2.$$

Then, expanding the wavefunction around this new minimum, we have

$$\psi = (v + h(x))e^{i\theta(x)}$$

The covariant derivative gives

$$(-i\hbar\nabla - 2e\mathbf{A})\psi = -i\hbar e^{i\theta(x)}\nabla h(x) + (v+h(x))e^{i\theta(x)}\hbar\nabla\theta(x) - 2e\mathbf{A}(v+h(x))e^{i\theta(x)}.$$

The complex conjugate reads

$$(i\hbar\nabla - 2e\mathbf{A})\psi^\dagger = i\hbar e^{-i\theta(x)}\nabla h(x) + (v+h(x))e^{-i\theta(x)}\hbar\nabla\theta(x) - 2e\mathbf{A}(v+h(x))e^{-i\theta(x)}.$$

Multiplying the two together, we have

$$\left|(-i\hbar\nabla - 2e\mathbf{A})\psi\right|^2 = \hbar^2(\nabla h)^2 + 2e(v+h)^2\left[\mathbf{A} - \frac{\hbar}{2e}\nabla\theta\right].$$

Thus, the free energy takes on the form of

$$F[\psi] = F_0 + \int d^3x\, \frac{1}{2\mu_0}\mathbf{B}^2 + \hbar^2(\nabla h)^2 + 2e(v+h)^2\left[\mathbf{A} - \frac{\hbar}{2e}\nabla\theta\right] + \\ + \alpha_0(T - T_c)(v+h)^2 + \frac{\beta}{2}(v+h)^4$$

expanded out and truncated it reads that

$$F[\psi] = F_0 + \int d^3x\, \frac{1}{2\mu_0}\mathbf{B}^2 + \hbar^2(\nabla h)^2 + 2ev^2\mathbf{A} + \\ + \alpha_0(T - T_c)(v^2 + 2vh + h^2) + \frac{\beta}{2}(v^2 + 2vh + h^2)^2$$

where we have used a gauge transformation to rotate away the Goldstone boson $\theta(x)$. When we evaluate the equation of motion for this new free energy for the gauge field $\mathbf{A}$, we find

$$\nabla^2\mathbf{A} + \frac{\mu_0(2ev)^2}{2m_e}\mathbf{A} = 0.$$

And then, finally, for the photon, we recover the London equations (Eq. 30)

$$\nabla^2 \mathbf{B} + \frac{\mu_0 (2ev)^2}{2m_e} \mathbf{B} = 0. \tag{31}$$

and the massive photon due to the Meißner effect. The magnetic field decays exponentially inside the bulk via the Meißner effect (Eqs. 30 and 31), and the decay has its own penetration strength, which is exactly the London penetration length λ. The important thing is that the system has dynamically generated a mass term for the magnetic field.

Once again, compare the GSW model, whose Lagrangian takes the form of

$$\mathcal{L}_{GSW} = -\frac{1}{4}(W^a_{\mu\nu})^2 - \frac{1}{4}(B_{\mu\nu})^2 + (D_\mu H)^\dagger D^\mu H + m^2 H^\dagger H - \lambda (H^\dagger H)^2.$$

The covariant derivative reads

$$D_\mu H = \partial_\mu H - igW^a_\mu T^a H - \frac{1}{2} i g' B_\mu H.$$

The Englert-Higgs mechanism [6] proceeds as we take the derivative of $\mathcal{L}_{GSW}$ with respect to the Higgs field $H^\dagger$ and minimise it.

before minimalising the potential, we perform a Wick rotation which transforms the Minkowskian metric into an Euclidean one, effectively giving us

$$\mathcal{L}_{GSW} = -\frac{1}{4}(W^a_{\mu\nu})^2 - \frac{1}{4}(B_{\mu\nu})^2 + (D_\mu H)^\dagger D^\mu H - m^2 H^\dagger H - \lambda (H^\dagger H)^2.$$

It is of great importance that the sign for the mass term has now changed. For the Ginzburg-Landau theory, α_0 was defined to be sign changing, but in particle physics, we have to use such tricks or ad hoc assumptions because we have strict restrictions on mass and its semi-positive values. minimising the higgs potential $V(H) = -m^2 H^\dagger H - \lambda (H^\dagger H)^2$ now we find

$$|H_0|^2 = \frac{m}{\lambda} \equiv v^2$$

Now, substituting this expectation value back into the covariant derivative (after the same expansion of the Higgs field around its new minimum as we have done it for superconductivity), we find that

$$\begin{aligned} |D_\mu H|^2 &= \frac{g^2}{4} \begin{bmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{g'}{g} B_\mu + W^3_\mu & W^1_\mu - iW^2_\mu \\ W^1_\mu + iW^2_\mu & \frac{g'}{g} B_\mu + W^3_\mu \end{bmatrix} \begin{bmatrix} \frac{g'}{g} B_\mu + W^3_\mu & W^1_\mu - iW^2_\mu \\ W^1_\mu + iW^2_\mu & \frac{g'}{g} B_\mu - W^3_\mu \end{bmatrix} \begin{bmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{bmatrix} \\ &= \frac{g^2 v^2}{8} \left[W^+_\mu W^-_\mu + \left(\frac{g'}{g} B_\mu + W^3_\mu \right)^2 \right]. \end{aligned}$$

that the gauge bosons become massive. That is why we call the W and Z bosons massive vector bosons, and this is where their masses come from. Furthermore, we say that the gauge bosons have eaten the Goldstone bosons by a unitary gauge transformation. This is the origin of how the massive gauge bosons obtain their mass. Applying this new Higgs field expansion for the rest of the potential, we find

$$\mathcal{L}_{Higgs} = -\frac{1}{2}h(\Box + m_h^2)h - g\frac{m_h^2}{4m_W}h^3 - g^2\frac{m_h^2}{32m_W^2}h^4 \\ + \left(\frac{2}{v}h + h^2\right)\left[m_W^2 W_\mu^+ W_\mu^- + \frac{1}{2}m_Z^2(Z_\mu)^2\right].$$

Here we can see that the Higgs mass now has the correct sign, despite or thanks to the mathematical craftsmanship we did before the minimisation process. These terms are non-linear and introduce a new type of physics called the Higgs physics, which concerns itself with the self-interactions of the Higgs field.

7 Gap equation: BCS vs. NJL

The Bardeen-Cooper-Schrieffer model's Lagrangian density reads

$$\mathcal{L} = \sum_\sigma \psi_\sigma^\dagger \left[i\hbar\partial_t + \frac{\hbar^2}{2m}\nabla^2 + \mu\right]\psi_\sigma + V_0 \psi_\uparrow^\dagger \psi_\downarrow^\dagger \psi_\uparrow \psi_\downarrow.$$

The BCS Hamiltonian will read

$$H_{BCS} = \int d^3x \sum_\sigma \psi_\sigma^\dagger \left[-\frac{\hbar^2}{2m}\nabla^2 - \mu\right]\psi_\sigma + V_0 \int d^3x\, \psi_\uparrow^\dagger \psi_\downarrow^\dagger \psi_\uparrow \psi_\downarrow. \tag{32}$$

When we perform the calculation of matrix elements by using orthogonality relations, or to paraphrase the multipole expansion section, we calculate the first moment of the energy (Eq. 32). Then we have the following discrete Hamiltonian, aligning with what we get using Dirac's procedure in time-dependent perturbation theory. The discrete BCS equation [7] reads

$$H_{BCS} = \sum_{\mathbf{k},\sigma}(\epsilon_\mathbf{k} - \mu)a_{\mathbf{k},\sigma}^\dagger a_{\mathbf{k},\sigma} - \sum_\mathbf{k} \Delta_\mathbf{k} a_{\mathbf{k},\uparrow}^\dagger a_{-\mathbf{k},\downarrow}^\dagger + \Delta_\mathbf{k}^* a_{-\mathbf{k},\downarrow} a_{\mathbf{k},\uparrow}, \tag{33}$$

Where $\epsilon_\mathbf{k} - \mu$ is the electron kinetic energy relative to the Fermi surface. The model describes an attractive, phonon-mediated interaction that binds electrons of opposite momentum and spin $(\mathbf{k}, \uparrow; -\mathbf{k}, \downarrow)$ into Cooper pairs. The continuous gap equation reads

$$\Delta(\mathbf{k}) = -\int \frac{d^3k'}{(2\pi)^3} V(\mathbf{k},\mathbf{k'}) \frac{\Delta(\mathbf{k'})}{2\sqrt{(\epsilon(\mathbf{k})-\mu)-|\Delta(\mathbf{k'})|^2}} \tanh\left[\frac{\sqrt{(\epsilon(\mathbf{k})-\mu)-|\Delta(\mathbf{k'})|^2}}{2k_BT}\right]$$

while analogously, the discrete version reads

$$\Delta_{\mathbf{k}} = -\sum_{\mathbf{k'}} V_{\mathbf{k},\mathbf{k'}} \frac{\Delta_{\mathbf{k'}}}{2\sqrt{(\epsilon_{\mathbf{k}}-\mu)^2+|\Delta_{\mathbf{k'}}|^2}} \tanh\left[\frac{\sqrt{(\epsilon_{\mathbf{k}}-\mu)^2+|\Delta_{\mathbf{k'}}|^2}}{2k_BT}\right], \quad (34)$$

Where $V_{\mathbf{k},\mathbf{k'}}$ is the interaction potential, and $E_{\mathbf{k}} = \sqrt{(\epsilon_{\mathbf{k}}-\mu)^2+|\Delta_{\mathbf{k}}|^2}$ is the total excitation energy of a quasiparticle. The term $|\Delta_{\mathbf{k'}}|$ physically represents the minimum energy required to break a Cooper pair.

At its core, the gap equation (Eq. 34) represents a self-consistency condition. A Cooper pair is loosely bound by an attractive interaction (usually mediated by lattice vibrations, or phonons). However, the strength of this "pairing glue" depends on how many other Cooper pairs have already formed. The gap equation balances these interdependent quantities mathematically.

The introduction of pions, or rather a more precise calculation of their masses compared to Yukawa's model, was first done by Nambu and Jona-Lasinio [8, 9]. The tool they used was lifted from superconductivity in the form of Cooper pairs and gap equations. Although in particle physics we call them chiral condensates of quarks $\langle \bar{q}_i q_j \rangle$ and not Cooper pairs. Although at higher energies we have diquark condensates $\langle q_i q_j \rangle$, which are closer to Cooper pairs. The Nambu-Jona-Lasinio model starts out with a chiral Lagrangian, but it contains quartic terms. namely, we have

$$\mathcal{L} = \sum_n i\bar{\psi}_n \gamma^\mu \partial_\mu \psi_n \,+\, \frac{G}{4}[(\bar{\psi}_n\psi_n)^2 - (\bar{\psi}_n\gamma_5 T^a \psi_n)^2].$$

In terms of left and right-handed spinor fields, we have

$$\mathcal{L} = \sum_n i\bar{\psi}_n^L \gamma^\mu \partial_\mu \psi_n^L + i\bar{\psi}_n^R \gamma^\mu \partial_\mu \psi_n^R \,+\, G(\bar{\psi}_n^L\psi_n^R)(\bar{\psi}_n^R\psi_n^L).$$

The gauged (by gauged I mean we plug in the chiral condensates, namely the σ and the π^a mesons) equation takes the form of

$$\mathcal{L}_{NJL} = \sum_n i\bar{\psi}_n \gamma^\mu \partial_\mu \psi_n \,-\, \bar{\psi}_n(\sigma + i\gamma_5\pi^a T^a)\psi_n \,+\, \frac{G}{4}[(\bar{\psi}_n\psi_n)^2 - (\bar{\psi}_n\gamma_5 T^a\psi_n)^2]. \quad (35)$$

Completing the square will yield

$$\mathcal{L}_{NJL} = \sum_n i\bar{\psi}_n \gamma^\mu \partial_\mu \psi_n - \frac{1}{G}[\sigma^2 + (\pi^a)^2] + \frac{G}{4}\left(\bar{\psi}_n\psi_n - \frac{2}{G}\sigma\right)^2 + \frac{G}{4}\left(i\bar{\psi}_n\gamma_5 T^a\psi_n - \frac{2}{G}\pi^a\right)^2 .$$

Evaluating it in the partition function (with Eq. 35), we get a new functional, the famous trace-log term, which is denoted by $\Gamma[\sigma, \pi^a]$ and once we minimise this new $\Gamma[\sigma, \pi^a]$ functional with respect to σ, we get the famous gap equation

$$\langle\bar{\psi}\psi\rangle = 8GN_c \int \frac{d^4p}{(2\pi)^4}\frac{m+\sigma}{p^2-(m+\sigma)^2}. \tag{36}$$

With current quark masses, it becomes

$$\langle\bar{\psi}\psi\rangle = m + 8GN_c \int \frac{d^4p}{(2\pi)^4}\frac{m+\sigma}{p^2-(m+\sigma)^2}. \tag{37}$$

Physically, what's important in this equation? In the Nambu-Jona-Lasinio model (Eq. 35), the pions are actual Goldstone bosons, that is, they are massless. However, it is to be noted that the sigma particle obtains mass. It obtains mass even if the original quarks are massless, which was the case in our example (Eqs. 36 and 37). However, in some more advanced models, they are considered quarks which have acquired their fundamental mass. The masses for the mesons and baryons are purely created from the dynamical symmetry breaking of the chiral colour symmetry.

Here, the interaction generates a chiral condensate vacuum expectation value, $\langle\bar{\psi}\psi\rangle \neq 0$. The gap parameter here is the constituent mass M itself. Breaking this dynamic "gap" requires providing enough energy to create a heavy quark out of the vacuum. In both cases, the gap equation creates and calculates an energy gap between bound states; however, for the gap equation in the NJL case, it is emphasised that we calculate the masses of mesons rather than just energy bands. The other important difference is that in QCD, we focus on the chiral nature of these condensates. While in low energy the NJL model holds supreme in high energies like the cores of neutron stars, it breaks down partially restoring symmetries, creating a diquark condensate that is a quark Cooper pair and actually producing colour superconductivity, which is the colourful analogue of the electric superconductivity.

References

[1] Sheldon L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579–588, 1961. URL: https://www.sciencedirect.com/science/article/pii/0029558261904692, doi:10.1016/0029-5582(61)90469-2.

[2] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967. URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264, doi:10.1103/PhysRevLett.19.1264.

[3] Abdus Salam. Weak and Electromagnetic Interactions. *Conf. Proc. C*, 680519:367–377, 1968. doi:10.1142/9789812795915_0034.

[4] F. London and H. London. The electromagnetic equations of the supraconductor. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 149(866):71–88, 03 1935. arXiv:https://royalsocietypublishing.org/rspa/article-pdf/149/866/71/33451/rspa.1935.0048.pdf, doi:10.1098/rspa.1935.0048.

[5] V. L. Ginzburg and L. D. Landau. On the Theory of superconductivity. *Zh. Eksp. Teor. Fiz.*, 20:1064–1082, 1950. doi:10.1016/b978-0-08-010586-4.50078-x.

[6] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508, doi:10.1103/PhysRevLett.13.508.

[7] J. Bardeen, L. N. Cooper, and J. R. Schrieffer. Microscopic theory of superconductivity. *Phys. Rev.*, 106:162–164, Apr 1957. URL: https://link.aps.org/doi/10.1103/PhysRev.106.162, doi:10.1103/PhysRev.106.162.

[8] Y. Nambu and G. Jona-Lasinio. Dynamical model of elementary particles based on an analogy with superconductivity. i. *Phys. Rev.*, 122:345–358, 1961. doi:10.1103/PhysRev.122.345.

[9] Y. Nambu and G. Jona-Lasinio. Dynamical model of elementary particles based on an analogy with superconductivity. ii. *Phys. Rev.*, 124:246–254, Oct 1961. URL: https://link.aps.org/doi/10.1103/PhysRev.124.246, doi:10.1103/PhysRev.124.246.