Advanced Field Theory

ETH Zurich and University of Zurich

FS 2010

Prof. Dr. Babis Anastasiou [ETH Zurich]
Prof. Dr. Thomas Gehrmann [University of Zurich]
Typeset by Peter Reiter

June 28, 2010

Please feel free to send feedback or error reports to rimathia@ethz.ch

Contents

Chapter 1

Dimension Regularization and the Axial Anomaly

From quantum field theory we know that field theories contain loop diagrams with ultraviolet ($k^2 \to \infty$) or the infrared ($k^2 \to 0$) divergences.

One example of a UV-divergent integral is,

$$\int \frac{\mathrm{d}^4k}{(2\pi)^4} \frac{1}{(k^2-m^2)[(k-q)^2-m^2]} \xrightarrow{k\to\infty} \int \frac{\mathrm{d}^4k}{(2\pi)^4} \frac{1}{k^4}. \tag{1.1}$$

Regulating this integral will yield a logarithmic divergence which can be removed using renormalization techniques.

In the low energy limit of a massless theory, consider for example the vertex diagram corresponding to the integral

$$\int \frac{\mathrm{d}^4k}{(2\pi)^4} \frac{1}{k^2(k+p_1)^2(k-p_2)^2}, \tag{1.2}$$

where we identify in the vertex diagram $k \to p_1, p_2, p_1^2 = p_2^2 = 0, k = p_1 + p_2$. We observe that due to the missing mass in the propagator, we have a divergence for $k \to 0$. In the IR limit we can distinguish between two limits:

1. The soft limit is defined as the limit of $k \to \lambda k, \lambda \to 0$:

$$\int \frac{\mathrm{d}^4k}{(2\pi)^4} \frac{1}{k^2(2k\cdot p_1)(-2k\cdot p_1)} \xrightarrow{k\to 0} \int \frac{d^4k}{(2\pi)^4} \frac{1}{k^4}, \tag{1.3}$$

which is obviously logarithmically divergent.

2. In the collinear limit we have $k^2 = 0$ and $k \parallel p_1, k \parallel p_2$, use the Sudakov parametrization (also known as light cone parametrization) by choosing a frame where p_1 and p_2 move in opposite directions, i.e.

$$e_+^\mu = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\0\\1\end{pmatrix}, \quad e_-^\mu = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\0\\-1\end{pmatrix}, \quad p_1^\mu = p_1 e_+^\mu. \tag{1.4}$$

We can thus decompose

$$k^\mu = k_+ e_+^\mu + k_- e_-^\mu + \mathbf{k}_\perp = \frac{1}{\sqrt{2}}\begin{pmatrix}k_+ + k_-\\ \mathbf{k}_\perp \\ k_+ - k_-\end{pmatrix}, \tag{1.5}$$

and then using $\mathbf{k}_\perp^2 = k_1^2 + k_2^2$ continue with

$$k^2 = k_0^2 - k_1^2 - k_2^2 - k_3^2 = (k_0 - k_3)(k_0 + k_3) - \mathbf{k}_\perp^2 = 2k_+k_- - \mathbf{k}_\perp^2 \tag{1.6}$$

to change our measure,

$$\mathrm{d}^4k = 2\mathrm{d}k_+\mathrm{d}k_-\mathrm{d}^2k_T = \mathrm{d}k_+\mathrm{d}k_- - \mathrm{d}k^2. \tag{1.7}$$

We can rewrite our loop integral, and take, as usual, the infrared limit,

$$\int \frac{\mathrm{d}k^2}{k^2}\frac{\mathrm{d}k_+}{(k^2+4p_1k_-)}\frac{\mathrm{d}k_-}{(\dots)} \xrightarrow{k\to 0} \int \frac{\mathrm{d}k^2}{k^2}\frac{\mathrm{d}k_+}{4p_1k_-}\frac{\mathrm{d}k_-}{\dots}, \tag{1.8}$$

which again is logarithmically divergent.

1.1 Basics of Dimensional Regularization

Let d be a complex number. We wish to define an operation that we may regard as integration over a d-dimensional space:

$$\int \mathrm{d}^d\mathbf{p} f(\mathbf{p}). \tag{1.9}$$

Here $f(\mathbf{p})$ is any given function of a vector $\mathbf{p}$, which is in the d-dimensional space. We will suppose that the space is Euclidean. (Minkowski space is regarded as a one-dimensional time together witha $(d-1)$-dimensional Euclidian space.)

What properties must we impose on a functional of f in order to regard it as d-dimensional integration? The following properties or axioms are natural and are necessary in applications to Feynman graphs:

1. Linearity: For any $a, b \in \mathbb{C}$:
$$\int \mathrm{d}^d\mathbf{p}\ (af(\mathbf{p}) + bg(\mathbf{p})) = a \int \mathrm{d}^d\mathbf{p}\ f(\mathbf{p}) + b \int \mathrm{d}^d\mathbf{p}\ g(\mathbf{p}). \tag{1.10}$$

2. Scaling: For any number s
$$\int \mathrm{d}^d\mathbf{p}\ f(s\mathbf{p}) = s^{-d} \int \mathrm{d}^d\mathbf{p}\ f(\mathbf{p}). \tag{1.11}$$

3. Translation invariance: For any vector $\mathbf{q}$:
$$\int \mathrm{d}^d\mathbf{p}\ f(\mathbf{p}+\mathbf{q}) = \int \mathrm{d}^d\mathbf{p}\ f(\mathbf{p}). \tag{1.12}$$

We will also require rotational covariance of our results.

Linearity is true of any integration, while translation and rotation invariance are basic properties of Euclidian space, and the scaling property embodies the d-dimensionality.

Not only are the above three axioms necessary, but they also ensure that integration is unique, aside from an overall normalization. In fact, they determine the usual integration measure in an integer-dimensional space (again up to normalization).

1.2 Clifford Algebra

The Dirac matrices satisfy the following properties:

1. Anticommutation relation:
$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}\mathbb{1} \tag{1.13}$$

2. Hermiticity:
$$\gamma^{\mu\dagger} = \gamma_\mu = \begin{cases} \gamma^\mu & \text{if } \mu = 0, \\ -\gamma^\mu & \text{if } \mu \geq 1. \end{cases} \tag{1.14}$$

When we use dimensional regularization, the Lorentz indices range over an infinite set of values, so we need infinite-dimensional matrices to represent the algebra. We will also need a trace operation:
$$\mathrm{tr}\ 1 = f(d),$$

so that the representation behaves as if its dimension were $f(d)$. We must require $f(d_0)$ to be the usual value at the physical space-time dimension, $d = d_0$. Usually this means $f(4) = 4$.

The trace is a linear operation on the matrices which we will define later. In an even integer dimension $d = 2\omega$, the standard representation of the γ^μ's has dimension 2^ω. However, it is not necessary to choose $f(d) = 2^{d/2}$. The variation $f(d) - f(d_0)$ is only relevant for a divergent graph. It is usually convenient to set $f(d) = f(d_0)$ for all d.

In four dimensions, $\gamma_5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3$ and $\varepsilon_{\kappa\lambda\mu\nu}$ is a totally antisymmetric Lorentz-invariant tensor with $\varepsilon_{0123} = 1$. We need γ_5, for example, to define the axial current $\bar{\psi}\gamma^\mu\gamma_5\psi$. The ε-tensor comes in because $\gamma_5 = i\varepsilon_{\kappa\lambda\mu\nu}\gamma^\kappa\gamma^\lambda\gamma^\mu\gamma^\nu/4!$, and we have the trace formula:

$$\text{tr}\ \gamma^5\gamma^\kappa\gamma^\lambda\gamma^\mu\gamma^\nu = i\varepsilon^{\kappa\lambda\mu\nu}\text{tr}1 = -i\varepsilon_{\kappa\lambda\mu\nu}\text{tr}1.$$

The appropriate definition changes when we go to two dimensions: Instead of γ_5 we have $\hat{\gamma}_{(1)} = \gamma^0\gamma^1$, and instead of $\varepsilon_{\kappa\lambda\mu\nu}$ we have $\varepsilon_{\mu\nu}$, for which $\varepsilon_{01} = -\varepsilon_{10} = 1, \varepsilon_{00} = \varepsilon_{11} = 0$.

To continue dimensionally, we might expect γ_5 to satisfy

$$\{\gamma_5, \gamma^\mu\} = 0,$$

just as in four dimensions. But then, the only consistent result for γ_5 is that it has zero trace when multiplied by any string of γ^μ's. Thus we do not have a regularization involving the usual γ_5.

A consistent definition is obtained by writing

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = i\gamma^\kappa\gamma^\lambda\gamma^\mu\gamma^\nu\varepsilon_{\kappa\lambda\mu\nu}/4! \tag{1.15}$$

$$\varepsilon_{\kappa\lambda\mu\nu} = \sigma(\kappa\lambda\mu\nu), \tag{1.16}$$

where $\sigma(\kappa\lambda\mu\nu)$ gives the sign of the permutation of $\kappa\lambda\mu\nu$ with respect to (0123). This definition is not Lorentz invariant on the full space, but only on the first four dimensions. We have

$$\{\gamma_5, \gamma^\mu\} = 0, \quad \text{if } \mu = 0, 1, 2, 3, \tag{1.17}$$

$$[\gamma_5, \gamma^\mu] = 0, \quad \text{otherwise}, \tag{1.18}$$

$$(\gamma_5)^2 = 1, \tag{1.19}$$

$$\gamma_5^\dagger = \gamma_5. \tag{1.20}$$

The lack of full Lorentz invariance is a nuisance, but it does give the correct axial anomaly.

Lorentz and Dirac-Algebra We require $g^\mu_\mu = d$, so a vector A^μ has $d - 2$ physical degrees of freedom. We don't have to generalize the gamma matrices themselves, just the traces. So we have to discuss the Lorentz indices and the $\mathbb{1}$-matrix in $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}\mathbb{1}$.

In dimensional regularization we only care about $\lim_{d\to 4} \mathrm{tr}\mathbb{1} = 4$. The consequence is that if we have spinors, then they do not have additional degrees of freedom than in 4 dimension, $\psi :$ 2 degrees of freedom.

So we are only left with $\gamma_5, \varepsilon_{\mu\nu\rho\sigma}$.

- introduce $(d-4)$-dimensional metric $\hat{g}_{\mu\nu} = \hat{g}_{\nu\mu}$, which acts as projector on generalized gamma matrices, with:
  - $\hat{g}_{\mu\nu} g^{\mu}{}_{\rho} = \hat{g}_{\rho\nu} = \hat{g}_{\mu\nu}\hat{g}^{\mu}{}_{\rho}$ (projects on $(d-4)$ subspaces)
  - $\hat{g}^{\mu}{}_{\mu} = (d-4)$
  - $\hat{g}^{\mu}{}_{\nu} p^{\nu} = \hat{p}^{\mu}$, $\hat{g}^{\mu}{}_{\nu}\gamma^{\nu} = \hat{\gamma}^{\mu}$: $(d-4)$-dim components of p^{μ}, γ^{μ}
  - $\{\gamma_{\mu}, \hat{\gamma}_{\nu}\} = \{\hat{\gamma}_{\mu}, \hat{\gamma}_{\nu}\} = 2\hat{g}_{\mu\nu}\mathbb{1}$
- define product of ε-tensors (4 dim!)

$$\varepsilon_{\mu_1\mu_2\mu_3\mu_4}\ \varepsilon_{\nu_1\nu_2\nu_3\nu_4} = -\sum_{\pi\in S_4} \mathrm{sgn}(\pi) \prod_{i=1}^{4} (g_{\mu_i\nu_{\pi(i)}} - \hat{g}_{\mu_i\nu_{\pi(i)}}) \tag{1.21}$$

  such that $\varepsilon_{\mu_1\mu_2\mu_3\mu_4} = \mathrm{sgn}(\pi)\varepsilon_{\mu_{\pi(1)}}\varepsilon_{\mu_{\pi(2)}}\varepsilon_{\mu_{\pi(3)}}\varepsilon_{\mu_{\pi(4)}}$ antisymmetric and $\varepsilon_{\mu\nu\rho\sigma}\hat{g}^{\mu\alpha} = 0$

  we can think of γ^{μ} in d-dimensions, but ε in 4 dimensions. this is just a prescription with a definition of γ_5 such that it does not spoil symmetries.

1.3 Chiral Symmetry

In massless QED (also QCD), the left-handed and right-handed fermions decouple in the Lagrangian:

$$\mathcal{L} = \bar{\psi}(i\not{D})\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu},\ \not{\partial} = \gamma^{\mu}\partial_{\mu},\ D_{\mu} = \partial_{\mu} + ieA_{\mu}$$

which we can rewrite it using

$$\bar{\psi}\gamma^{\mu}\psi = \bar{\psi}_L\gamma^{\mu}\psi_L + \bar{\psi}_R\gamma_{\mu}\psi_R$$

with $\psi_{R,L} = \frac{1\pm\gamma_5}{2}\psi$.

Question How is $m\bar{\psi}\psi =$?

Noether currents and conserved charges

- vector current $J^\mu = \bar{\psi}\gamma^\mu\psi$ is conserved. $\partial_\mu J^\mu = 0$ (even for $m \neq 0$). This implies a conserved number $N = \int \mathrm{d}^3x J^0(x)$, which is constant $\frac{\mathrm{d}N}{\mathrm{d}t} = 0$. (Fermion number conservation).

- Axial vector current ${J_5}^\mu = \bar{\psi}\gamma^\mu\gamma_5\psi = \bar{\psi}_R\gamma^\mu\psi_R - \bar{\psi}_L\gamma^\mu\psi_L$ is conserved, $\partial_\mu {J_5}^\mu = 0$ for $m = 0$.

  Question What is the symmetry for ${J_5}^\mu$?

  This means that the difference $\Delta N = N_R - N_L$ is conserved as well.

  $\Rightarrow N_R, N_L$ both conserved independently.

1.4 QED Axial Anomaly in Two Dimensions

Eventually, we will want to analyze the current conservation equation for the axial current in massless QCD. However, this discussion will involve some technical complication, so we will first study the physics that violates axial current conservation in a context in which the calculations are relatively simple. A particularly simple model problem is that of two-dimensional massless QED.

The Lagrangian of the massless two-dimensional QED is

$$\mathcal{L} = \bar{\psi}(i\not{D})\psi - \frac{1}{4}(F_{\mu\nu})^2, \tag{1.22}$$

with $\mu, \nu = 0, 1$ and $D_\mu \equiv \partial_\mu + ieA_\mu$. The Dirac matrices must be chosen to satisfy the Dirac algebra

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}. \tag{1.23}$$

In two dimensions, this set of relations can be represented by 2×2 matrices; we choose

$$\gamma^0 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \gamma^1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}. \tag{1.24}$$

The Dirac spinors will be two-component fields.

The product of the Dirac matrices, which anticommutes with each of the γ^μ, is

$$\gamma^5 = \gamma^0\gamma^1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.25}$$

Then, just as in four dimensions, there are two possible currents,

$$j^\mu = \bar{\psi}\gamma^\mu\psi, \quad j^{\mu 5} = \bar{\psi}\gamma^\mu\gamma^5\psi, \tag{1.26}$$

and both are conserved if there is no mass term in the Lagrangian.

To make the conservation laws quite explicit, we label the components of the fermion field ψ in this spinor basis as

$$\psi = \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix}. \tag{1.27}$$

The subscript indicates the γ^5 eigenvalue. Then, using the explicit representations of γ^0, γ^1, we can rewrite the fermionic part as

$$\mathcal{L} = \psi_+^\dagger i(D_0 + D_1)\psi_+ + \psi_-^\dagger i(D_0 - D_1)\psi_-. \tag{1.28}$$

In the free theory, the field equation of ψ_+, ψ_- would be

$$i(\partial_0 + \partial_1)\psi_+ = i(\partial_0 - \partial_1)\psi_- = 0; \tag{1.29}$$

the solutions to this equation are waves that move to the right in the one dimensional space at the speed of light. We will thus refer to the particles associated with ψ_+ as right-moving fermions. The quanta associated with ψ_- are, similarly, left-moving. This distinction is analogous to the distinction between left- and right-handed particles which gives the physical interpretation of γ^5 in four dimensions. Since the Lagrangian contains no terms that mix left- and right-moving fields, it seems obvious that the number currents for these fields are seperately conserved. Thus,

$$\partial_\mu\left(\bar{\psi}\gamma^\mu\left(\frac{1-\gamma^5}{2}\right)\psi\right) = 0, \quad \partial_\mu\left(\bar{\psi}\gamma^\mu\left(\frac{1+\gamma^5}{2}\right)\psi\right) = 0. \tag{1.30}$$

It is a curious property of two-dimensional spacetime that the vector and axial vector fermionic currents are not independent of each other. Let $\varepsilon^{\mu\nu}$ be the totally antisymmetric symbol in two dimensions, with $\varepsilon^{01} = +1$. Then the two-dimensional Dirac matrices obey the identity

$$\gamma^\mu\gamma^5 = -\varepsilon^{\mu\nu}\gamma_\nu. \tag{1.31}$$

The currents $j^{\mu 5}, j^\mu$ have the same relation. Thus we can study the properties of the axial vector current by using results that we have already defined for the vector current.

Once we have an explicit expression for the vacuum polarization, we can find the expectation value of the current induced by a background electromagnetic field. This quantity is generated by the diagram of Fig. 1.1, which gives

$$\int \mathrm{d}^2x e^{iq\cdot x}\langle j^\mu(x)\rangle = \frac{i}{e}(i\Pi^{\mu\nu}(x))A_\nu(q) = -\left(g^{\mu\nu} - \frac{q^\mu q^\nu}{q^2}\right)\cdot\frac{e}{\pi}A_\nu(q), \tag{1.32}$$

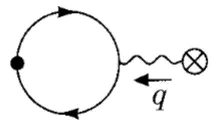

Figure 1.1: Computation of $\langle j^\mu \rangle$ in a background electromagnetic field.

where $A_\nu(q)$ is the Fourier transform of the background field. This quantity manifestly satiesfies the current conservation relation $q_\mu \langle j^\mu(q) \rangle = 0$.

The identity $\gamma^\mu \gamma^5 = -\varepsilon^{\mu\nu}\gamma_\nu$ between the vector and axial vector currents allows us to derive from the preceding integral the corresponding expectation value of $j^{\mu 5}$. We find

$$\langle j^{\mu 5}(q) \rangle = -\varepsilon^{\mu\nu} \langle j_\nu(q) \rangle \tag{1.33}$$

$$= \varepsilon^{\mu\nu} \frac{e}{\pi} \left(A_\nu(q) - \frac{q_\nu q^\lambda}{q^2} A_\lambda(q) \right). \tag{1.34}$$

If the axial vector current were conserved, this object would satisfy the Ward identity. Instead, we find

$$q_\mu \langle j^{\mu 5}(q) \rangle = \frac{e}{\pi} \varepsilon^{\mu\nu} q_\mu A_\nu(q). \tag{1.35}$$

This is the Fourier transform of the field equation

$$\partial_\mu j^{\mu 5} = \frac{e}{2\pi} \varepsilon^{\mu\nu} F_{\mu\nu}. \tag{1.36}$$

Apparently, the axial vector current is not conserved in the presence of electromagnetic fields, as the result of an anomalous behaviour of its vacuum polarization diagram. Thus we observe that the seperate conservation of left and right moving particles is violated by quantum interactions with the background field. The symmetries of the classical theory are broken by quantum effects.

To complete our discussion of the two-dimensional axial vector current, we will show that the nonconservation equation also has global aspect. In free fermion theory, the integral of the axial current conservation law gives

$$\int \mathrm{d}^2x \; \partial_\mu j^\mu_5(x) = \int \mathrm{d}\tau \frac{\mathrm{d}}{\mathrm{d}\tau}(N_R - N_L) = (N_R - N_L)|_{t=+\infty} - (N_R - N_L)|_{t=-\infty}. \tag{1.37}$$

This relation implies the difference in the number of right-moving and left-moving fermions cannot be changed in any possible process. Combining this with the conservation law for the vector current, we conclude that the number of each type of fermion is seperately conserved. We might conclude that these seperate conservation laws hold also in two-dimensional QED. However, we have already found that we must be careful in making statements about the axial current.

In two-dimensional QED, the conservation equation for the axial current is replaced by the anomalous nonconservation equation. If the right-hand side of this equation were the total derivative of a quantity falling off sufficiently rapidly at infinity, its integral would vanish and we would still retain the global conservation law. In fact, $\varepsilon^{\mu\nu}F_{\mu\nu}$ is a total derivative:

$$\varepsilon^{\mu\nu}F_{\mu\nu} = 2\partial_\mu(\varepsilon^{\mu\nu}A_\nu). \tag{1.38}$$

However, it is easy to imagine examples where the integral of this quantity does not vanish, for example, a world with a constant background electric field. In such a world, the conservation law must be violated. But how can this happen?

Let us analyse this problem by thinking about fermions in one space dimension in a background A^1 field that is constant in space and has a very slow time dependence. We will assume that the system has a finite length L, with periodic boundary conditions. Notice that the constant A^1 field cannot be removed by a gauge transformation that satisfies the periodic boundary conditions.

Following the derivation of the three-dimensional Hamiltonian, we find that the Hamiltonian of this one-dimensional system is

$$H = \int \mathrm{d}x \psi^\dagger(-i\gamma^0\gamma^1 D_1)\psi \tag{1.39}$$
$$= \int \mathrm{d}x \left\{-i\psi_+^\dagger(\partial_1 - ieA^1)\psi + i\psi_-^\dagger(\partial_1 - ieA^1)\psi_-\right\}. \tag{1.40}$$

For a constant A^1 field, it is easy to diagonalize this Hamiltonian. The eigenstates of the covariant derivatives are wavefunctions

$$e^{ik_n x}, \quad \text{with } k_n = \frac{2\pi n}{L}, n \in \mathbb{Z}, \tag{1.41}$$

to satisfy the periodic boundary conditions. Then the single-particle eigenstates of H have energies

$$\psi_+ : \quad E_n = +(k_n - eA^1), \tag{1.42}$$
$$\psi_- : \quad E_n = -(k_n - eA^1). \tag{1.43}$$

Each type of fermion has an infinite tower of equally spaced levels. To find the ground state of H, we fill the negative energy levels and interpret holes created among these filled states as antiparticles.

Now, adiabatically change the value of A^1,

$$\Delta A^1 = \frac{2\pi}{eL}. \tag{1.44}$$

The fermion energy levels slowly shift in accord with the previously stated relation back to its original value, the spectrum of H returns to its original form. In this process, each

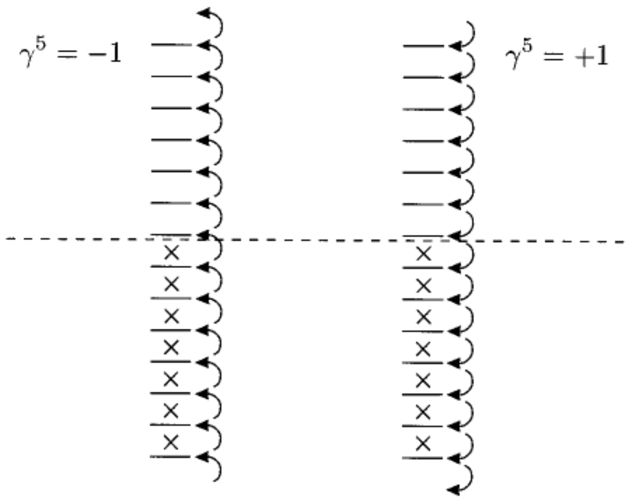

Figure 1.2: Effect on the vacuum state of the Hamiltonian H of one-dimensional QEd due to an adiabatic change in the background A^1 field.

level of ψ_+ moves down to the next position, and each level of ψ_- moves up to the next position, as shown in Fig. 1.2. The occupation numbers of levels should be maintained in this adiabatic process. Thus, remarkably, one right-moving fermion disappears from the vacuum and one extra left-moving fermion appears. At the same time,

$$\int \mathrm{d}^2x \left(\frac{e}{\pi}\varepsilon^{\mu\nu}F_{\mu\nu}\right) = \int \mathrm{d}t\mathrm{d}x\frac{e}{\pi}\partial_0 A_1 \tag{1.45}$$

$$= \frac{e}{\pi}L(-\Delta A^1) \tag{1.46}$$

$$= -2. \tag{1.47}$$

Thus the integrated form of the anomalous nonconservation equation is indeed satisfied:

$$N_R - N_L = \int \mathrm{d}^2x \left(\frac{e}{2\pi}\varepsilon^{\mu\nu}F_{\mu\nu}\right). \tag{1.48}$$

Even in this simple example, we see that it is not possible to escape the question of ultraviolet regularization in analysing the chiral conservation law. Right-moving fermions are lost and left-moving fermions appear from the depths of the fermionic spectrum, $E \to -\infty$. In computing the changes in the seperate fermion numbers, we have assumed that the vacuum cannot change the charge it contains at large negative energies. This prescription is gauge invariant, but it leads to the nonconservation of the axial vector current.

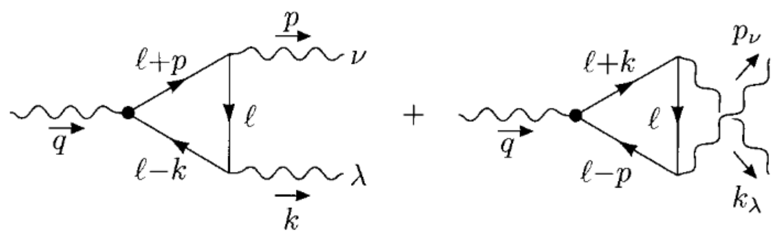

Figure 1.3: Diagrams contributing to the two-photon matrix element of the divergence of the axial vector current.

1.5 QED Axial Anomaly in Four Dimensions

We can confirm the Adler-Bell-Jackiw relation by checking, in standard perturbation theory, that the divergence of the axial vector current has a nonzero matrix element to create two photons. To do this, we must analyse the matrix element

$$\int \mathrm{d}^4 x e^{-iq\cdot x}\langle p,k|j^{\mu 5}(x)|0\rangle = (2\pi)^4\delta^{(4)}(p+k-q)\varepsilon^*_\nu(p)\varepsilon^*_\lambda(k)\mathcal{M}^{\mu\nu\lambda}(p,k). \tag{1.49}$$

The leading-order diagram contributing to $\mathcal{M}^{\mu\nu\lambda}$ are shown in Fig. 1.3. The first diagram gives the contribution

$$= (-1)(-ie)^2\int\frac{\mathrm{d}^4 l}{(2\pi)^4}\mathrm{tr}\left[\gamma^\mu\gamma^5\frac{i(\not{l}-\not{k})}{(l-k)^2}\gamma^\lambda\frac{i\not{l}}{l^2}\gamma^\nu\frac{i(\not{l}+\not{p})}{(l+p)^2}\right], \tag{1.50}$$

and the second diagram gives an identical contribution with (p,ν) and (k,λ) interchanged.

It is easy to give a formal argument that the matrix element of the divergence of the axial vector current vanishes at this order. Taking the divergence of the axial current is equivalent to dotting this quantity with iq_μ. Now we operate on the right-hand side as we do to prove a Ward identity. Replace

$$q_\mu\gamma^\mu\gamma^5 = (\not{l}+\not{p}-\not{l}+\not{k})\gamma^5 = (\not{l}+\not{p})\gamma^5 + \gamma^5(\not{l}-\not{k}). \tag{1.51}$$

Each momentum factor combines with the numerator adjacent to it to cancel the corresponding denominator. Thus we get

$$iq_\mu\cdot[triangle] = e^2\int\frac{\mathrm{d}^4 l}{(2\pi)^4}\mathrm{tr}\left[\gamma^5\frac{(\not{l}-\not{k})}{(l-k)^2}\gamma^\lambda\frac{\not{l}}{l^2}\gamma^\nu + \gamma^5\gamma^\lambda\frac{\not{l}}{l^2}\gamma^\nu\frac{(\not{l}+\not{p})}{(l+p)^2}\right]. \tag{1.52}$$

Now pass γ^ν through γ^5 in the second term and shift the integral over the first term according to $l\to(l+k)$:

$$iq_\mu\cdot[triangle] = e^2\int\frac{\mathrm{d}^4 l}{(2\pi)^2}\mathrm{tr}\left[\gamma^5\frac{\not{l}}{l^2}\gamma^\lambda\frac{(\not{l}+\not{k})}{(l+k)^2}\gamma^\nu - \gamma^5\frac{\not{l}}{l^2}\gamma^\nu\frac{(\not{l}+\not{p})}{(l+p)^2}\gamma^\lambda\right]. \tag{1.53}$$

This expression is manifestly antisymmetric under the interchange of (p, ν) and (k, λ), so the contribution of the second diagram in Fig. 1.3 precisely cancels.

However, because this derivation involves a shift of the integration variable, we should look closely whether this shift is allowed by the regularization. We see that the integral that must be shifted is divergent. If the diagram is regulated with a simple cutoff, or even with Pauli-Villars regularization, it turns out that the shift leaves oer a finite, nonzero term. The analysis of the axial vector current, even dimensional regularization has an extra subtlety, because γ^5 is an intrinsically four-dimensional object. In their original paper on dimensional regularization, 't-Hooft and Veltman suggested using the definition

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{1.54}$$

in d dimensions. This definition has the consequence that γ^5 anticommutes with γ^μ for $\mu = 0, 1, 2, 3$ but commutes with γ^μ for other values of μ.

In the evaluation, the external indices and the momenta p, k, q all live in the physical four dimensions, but the loop momentum l has components in all dimensions. Write

$$l = \bar{l} + \hat{l} \tag{1.55}$$

where the first term has nonzero components in dimensions 0,1,2,3 and the second term has nonzero components in the other $d-4$ dimensions. Because γ^5 commutes with the γ^μ in these extra dimensions, the identity is modified to

$$q_\mu \gamma^\mu \gamma^5 = (\not{l} + \not{k})\gamma^5 + \gamma^5(\not{l} - \not{p}) - 2\gamma^5 \hat{\not{l}}. \tag{1.56}$$

The first two terms cancel according to the argument given above; the shift is justified by the dimensional regularization. However, the third term gives an additional contribution:

$$iq_\mu \cdot [triangle] = e^2 \int \frac{\mathrm{d}^4 l}{(2\pi)^4} \mathrm{tr}\left[-2\gamma^5 \hat{\not{l}} \frac{(\not{l} - \not{k})}{(l-k)^2} \gamma^\lambda \frac{\not{l}}{l^2} \gamma^\nu \frac{(\not{l} + \not{p})}{(l+p)^2} \right]. \tag{1.57}$$

To evaluate this contribution, combine denominators in the standard way, and shift the integration variable $l \to l + P$, where $P = xk - yp$. In expanding the numerator, we must retain one factor each of $\gamma^\nu, \gamma^\lambda, \not{p}, \not{k}$ to give a nonzero trace with γ^5. This leaves over one factor of $\hat{\not{l}}$ and one factor of $\not{l}$ which must also be evaluated with components in extra dimensions in order to give a nonzero integral. The factors $\hat{\not{l}}$ anticommute with the other Dirac matrices in the problem and thus can be moved to adjacent positions. Then we must evaluate the integral

$$\int \frac{\mathrm{d}^4 l}{(2\pi)^4} \frac{\hat{\not{l}}\hat{\not{l}}}{(l^2 - \Delta)^3}, \tag{1.58}$$

where Δ is a function of k, p and the Feynman parameters. Using

$$(\hat{\not{l}})^2 = \hat{l}^2 \to \frac{d-4}{d} l^2 \tag{1.59}$$

under the symmetrical integration, we can evaluate

$$\int \frac{d^4 l}{(2\pi)^4} \frac{\hat{l}\hat{l}}{(l^2 - \Delta)^3}, = \frac{i}{(4\pi)^{d/2}} \frac{d-4}{2} \frac{\Gamma(2-\frac{d}{2})}{\Gamma(3)\Delta^{2-d/2}} \tag{1.60}$$

$$\stackrel{d\to 4}{\to} \frac{-i}{2(4\pi)^2}. \tag{1.61}$$

Notice the behaviour in which a logarithmically divergent integral contributes a factor $(d-4)$ in the denominator and allows an anomalous term, formally proportional to $(d-4)$, to give a finite contribution. The remainder of the algebra in the evaluation of the integral is straightforward. The terms involving the momentum shift P cancel, and we find

$$iq_\mu \cdot [triangle] = e^2 \left(\frac{-i}{2(4\pi)^2} \right) \mathrm{tr} \left[2\gamma^2 (-\not{k}\gamma^\lambda \not{p}\gamma^\nu \right] \tag{1.62}$$

$$= \frac{e^2}{4\pi^2} \varepsilon^{\alpha\lambda\beta\nu} k_\alpha p_\beta. \tag{1.63}$$

This term is symmetric under the interchange of (p, ν) with (k, λ), so the second diagram of Fig. 1.3 gives and equal contribution. Thus,

$$\langle p, k | \partial_\mu j^{\mu 5}(0) | 0 \rangle = -\frac{e^2}{2\pi^2} \varepsilon^{\alpha\nu\beta\lambda} (-ip_\alpha) \varepsilon^*_\nu(p) (-ik_\beta) \varepsilon^*_\lambda(k) \tag{1.64}$$

$$= -\frac{e^2}{16\pi^2} \langle p, k | \varepsilon^{\alpha\nu\beta\lambda} F_{\alpha\nu} F_{\beta\lambda}(0) | 0 \rangle, \tag{1.65}$$

as we would expect from the Adler-Bell-Jackiw anomaly equation, which is nothing else then the term $-\frac{4}{c}\vec{E} \cdot \vec{B}$ in classical electrodynamics.

Chapter 2

Chiral Symmetry in the Strong Interaction

- Peskin-Schroder, Chapter 19
- Yadurain: Theory of Quark and Gluon Interactions, Ch. 7

2.1 Chiral Symmetry of QCD

The Adler-Bell-Jackiw anomaly has a number of important implications for QCD. To describe these, we must first discuss the chiral symmetries of QCD systematically. In this discussion, we will ignore all but the lightest quarks u, d. In many analyses of the low-energy structure of the strong interactions, one also treats the s quark as light; this gives results that naturally generalize the ones we will find beloew.

The fermionic part of the QCD Lagrangian is

$$\begin{aligned} \mathcal{L} =& \bar{u} i \not{D} u + \bar{d} i \not{D} d - m_u \bar{u} u - m_d \bar{d} d && (2.1) \\ =& \bar{u}_L i \not{D} u_L + \bar{d}_L i \not{D} d_L + \bar{u}_R i \not{D} u_r + \bar{d}_R i \not{D} d_R && (2.2) \\ &+ m_u (\bar{u}_R u_L + \bar{u}_L u_R) + m_d (\bar{d}_R d_L + \bar{d}_L d_R). && (2.3) \end{aligned}$$

If the u, d quarks are very light, the last two terms are small and can be neglected. Let us study the implications of making this approximation. If we ignore the u, d masses, the Lagrangian of course has isospin symmetry, the symmetry of an $SU(2)$ unitary transformation mixing the u, d fields. However, because the classical Lagrangian for massless fermions contains no coupling between left- and right-handed quarks, this Lagrangian actually is symmetric under the seperate unitary transformations

$$\begin{pmatrix} u \\ d \end{pmatrix}_L \to U_L \begin{pmatrix} u \\ d \end{pmatrix}_L, \quad \begin{pmatrix} u \\ d \end{pmatrix}_R \to U_R \begin{pmatrix} u \\ d \end{pmatrix}_R. \tag{2.4}$$

Is is useful to seperate the $U(1)$ and $SU(2)$ parts of these transformations; then the symmetry group of the classical, massless QCD Lagrangian is $SU(2)_L \times SU(2)_R \times U(1)_L \times U(1)_R$ or, alternatively, $SU(2)_V \times SU(2)_A \times U(1)_V \times U(1)_A$. Let Q denote the quark doublet, with chiral components

$$Q_L = \frac{1-\gamma^5}{2}\begin{pmatrix} u \\ d \end{pmatrix}, \quad Q_R = \frac{1+\gamma^5}{2}\begin{pmatrix} u \\ d \end{pmatrix}. \tag{2.5}$$

Then we can write the currents associated with these symmetries as

$$j_L^\mu = \bar{Q}_L \gamma^\mu Q_L, \qquad j_R^\mu = \bar{Q}_R \gamma^\mu Q_R, \tag{2.6}$$

$$j_L^{\mu a} = \bar{Q}_L \gamma^\mu \tau^a Q_L, \qquad j_R^{\mu a} = \bar{Q}_R \gamma^\mu \tau^a Q_R, \tag{2.7}$$

where $\tau^a = \sigma^a/2$ represent the generators of $SU(2)$. The sums of left- and right-handed currents give the baryon number and isospin currents

$$j^\mu = \bar{Q}\gamma^\mu Q, \quad j^{\mu a} = \bar{Q}\gamma^\mu \tau^a Q. \tag{2.8}$$

The corresponding symmetries are the transformations given above with $U_L = U_R$. The difference of the currents give the corresponding axial vector currents $j^{\mu 5}, j^{\mu 5a}$:

$$j^{\mu 5} = \bar{Q}\gamma^\mu \gamma^5 Q, \quad j^{\mu 5a} = \bar{Q}\gamma^\mu \gamma^5 \tau^a Q. \tag{2.9}$$

In the discussion to follow, we will derive conclusions about the strong interactions by assuming that the classical conservation laws for these currents are not spoiled by anomalies. We will show below that this assumption is correct for the isotriplet currents $j^{\mu 5a}$ but not for $j^{\mu 5}$.

The vector $SU(2) \times U(1)$ transformations are manifest symmetries of the strong interactions, and the associated currents lead to familiar conservation laws. What about the orthogonal, axial vector, transformations? These do not correspond to any obvious symmetry of the strong interactions. In 1960, Nambu and Jona-Lasinio hypothesized that these are accurate symmetries of the strong interactions that are spontaneously broken. This idea has led to a correct and surprisingly detailed description of the properties of the strong interactions at low energy.

Spontaneous Breaking of Chiral Symmetry Before we describe the consequence of spontaneoulsy broken chiral symmetry, let us ask why we might expect the chiral symmetries to be spontaneously broken in the first place. In the theory of superconductivity, a small electron-electron attraction leads to the appearance of a condensate of electron pairs in the ground state of a metal. In QCD, quarks and antiquarks have strong attractive interactions, and, if these quarks are massless, the energy cost of creating an extra quark-antiquark pair is small. Thus we expect that the vacuum of QCD will contain a condensate of quark-antiquark pairs. These fermion pairs must have zero total momentum and angular momentum. Thus, as Fig XX shows, they must contain net chiral charge, pairing left-handed quarks with antiparticles of right-handed quarks. The vacuum state with a quark pari condensate is characterized by a nonzero vacuum expectation value for the scalar operator

$$\langle 0|\bar{Q}Q|0\rangle = \langle 0|\bar{Q}_L Q_R + \bar{Q}_R Q_L|0\rangle \neq 0, \tag{2.10}$$

which transforms under U_L, U_R with $U_L \neq U_R$. The expectation value signals that the vacuum mixes the two quark helicities. This allows the u, d quarks to acquire effective masses as they move through the vacuum. Inside quark-antiquark bound states, the u, d quarks would appear to move as if they had a sizeable effective mass, even if they had zero mass in the original QCD Lagrangian.

The vacuum expectation value $\langle 0|\bar{Q}Q|0\rangle$ signals the spontaneous breaking of the full symmetry group down to subgroup of vector symmetries with $U_L = U_R$. Thus there are four spontaneously broken continuous symmetries, associated with the four axial vector currents. The Goldstone theorem states that every spontaneously broken continuous symmetry of a quantum field theory leads to a massless particle with the quantum numbers of the global symmetry rotation. This means that, in QCD with massless u, d quarks, we should find four spin-zero particles with the correct quantum numbers to be created by the four axial vector currents.

The real strong interactions do not contain any massless particles, but they do contain an isospin triplet of relatively light mesons, the pions. These particles are known to have odd parity (as we expect if they are quark-antiquark bound states). Thus, they can be created by the axial isospin currents. We can parametrize the matrix element of $j^{\mu 5a}$ between the vacuum and an on-shell pion by writing

$$\langle 0|j^{\mu 5a}(x)|\pi^b(p)\rangle = -ip^\mu f_\pi \delta^{ab} e^{-ip\cdot x}, \tag{2.11}$$

where a, b are isospin indices and f_π is a constant with dimensions of $(\text{mass})^1$. We show in an exercise that the value of f_π can be determined from the rate of π^+ decay through the weak interaction; one finds $f_\pi = 98$ MeV. For this reason, f_π is often called the pion decay constant. If we contract it with p_μ and use the conservation of the axial currents, we find that an on-shell pion must satisfy $p^2 = 0$, that is, it must be massless, as required by Goldstone's theorem.

If we now restore the quark mass terms, the axial currents are no longer exactly conserved. The equation of motion of the quark field is now

$$i\not{D}Q = \mathbf{m}Q, \quad -iD_\mu \bar{Q}\gamma^\mu = \bar{Q}\mathbf{m}, \tag{2.12}$$

where

$$\mathbf{m} = \begin{pmatrix} m_u & 0 \\ 0 & m_d \end{pmatrix} \tag{2.13}$$

is the quark mass matrix. Then one can readily compute

$$\partial_\mu j^{\mu 5a} = i\bar{Q}\{\mathbf{m}, \tau^a\}Q. \tag{2.14}$$

Using this equation with the on-shell parametrization, we find

$$\langle 0|\partial_\mu j^{\mu 5a}(0)|\pi^b(p)\rangle = -p^2 f_\pi \delta^{ab} = \langle 0|i\bar{Q}\{\mathbf{m}, \tau^a\}\gamma^5 Q|\pi^b(p)\rangle. \tag{2.15}$$

The last expression is an invariant quantity times

$$\mathrm{tr}[\{\mathbf{m}, \tau^a\}\tau^b] = \frac{1}{2}\delta^{ab}(m_u + m_d). \tag{2.16}$$

Thus, the quark mass terms give the pions masses of the form

$$m_\pi^2 = (m_u + m_d)\frac{M^2}{f_\pi} = -\frac{m_u + m_d}{f_\pi^2}\langle 0|\bar{Q}Q|0\rangle. \tag{2.17}$$

The mass parameter M has been estimated to be of order 400 MeV. Thus, to give the observed pion mass of 140 MeV, one needs only $(m_u + m_d) \approx 10$ MeV, which then yields $\langle 0|\bar{Q}Q|0\rangle = -(260 \text{ MeV})^3$. This is a small perturbation on the strong interactions.

This argument has an interesting implication for the nature of the isospin symmetry of the strong interactions. In the limit in which the u, d quarks have zero mass in the Lagrangian, these quarks acquire large, equal effective masses from the vacuum with spontaneously broken chiral symmetry. As long as the masses m_u, m_d in the Lagrangian are small compared to the effective mass, the u, d quarks will behave inside the hadrons as though they are approximately degenerate. Thus the isospin symmetry of the strong interactions need have nothing to do with a fundamental symmetry linking u, d; it follows for any arbitray relation between m_u, m_d, provided that both of these parameters are much less than 300 MeV. Similarly, the approixmate $SU(3)$ symmetry of the strong interactions follows if the fundamental mass of the s quark is also small compared to the strong interaction scale. The best current estimates of the mass ratios $m_u : m_d : m_s$ are in fact $1 : 2 : 40$, so that the fundamental Lagrangian of the strong interactions shows no sign of flavour symmetry among the quark masses.

The Leptonic decay of charged pions is described by axial vector currents,

$$\sum_a j_5^{\mu a} = \sum_a \bar{Q}\gamma^\mu\gamma_5\tau^a Q = \bar{u}\gamma^\mu\gamma_5 d + \bar{d}\gamma^\mu\gamma_5 u + \bar{u}\gamma^\mu\gamma_5 u - \bar{d}\gamma^\mu\gamma_5 d, \tag{2.18}$$

is probed by weak decays (exercise)

$$\mathcal{L}_F = \frac{G_F}{\sqrt{2}}\bar{\mu}\gamma_\lambda(1-\gamma_5)\nu_\mu \bar{u}\gamma^\lambda(1-\gamma_5)d + \text{h.c.} + (\mu \leftrightarrow e). \tag{2.19}$$

The identification of the pion as the Goldstone boson of spontaneously broken chiral symmetry leads to numerous other predictions for current matrix elements and pion scattering amplitudes. In particular, the leading term of the pion-pion and pion-nucleon scattering amplitude at low energy can be computed directly in terms of f_π by arguments similar to one just given.

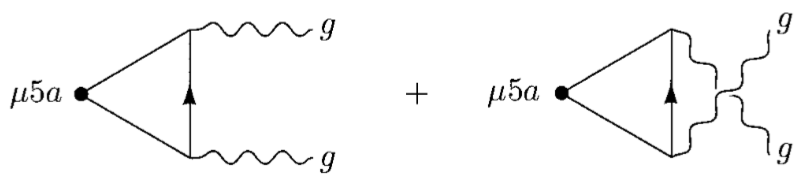

Figure 2.1: Diagrams that lead to an axial vector anomaly for a chiral current in QCD.

Anomalies of Chiral Currents Up to this point, we have discussed the chiral symmetries of QCD according to the classical current conservation equations. We must now ask whether these equations are affected by the Adler-Bell-Jackiw anomaly, and what the consequences of that modifications are.

To begin, we study the modification of the chiral conservation laws due to the coupling of the quark currents to the gluon fields of QCD. The arguments given in the previous section go through equally well in the case of massless fermions coupling to non-Abelian gauge field, so we expect that an axial vector current will receive an anomalous contribution from the diagrams shown in Fig. 2.1. The anomaly equation should be the Abelian result, supplemented by an appropriate group theory factor. In addition, since the axial current is gauge invariant, the anomaly must also be gauge invariant. That is, it must contain the full non-Abelian field strength, including its nonlinear terms. These terms are actually included in the functional derivative of the anomaly.

For the axial currents of QCD, we can read the group theory factors for the Adler-Bell-Jackiw anomaly from the diagrams in Fig. 2.1. for the axial isospin isotriplett currents,

$$\partial_\mu j^{\mu 5a} = -\frac{g^2}{16\pi^2}\varepsilon^{\alpha\beta\mu\nu}F^c_{\alpha\beta}F^d_{\mu\nu}\cdot \mathrm{tr}[\tau^a t^c t^d], \tag{2.20}$$

where $F^c_{\mu\nu}$ is a gluon field strength, τ^a is an isospin matrix, t^c is a color matrix, and the trace is taken over colors and flavours. In this case, we find

$$\mathrm{tr}[\tau^a t^c t^d] = \mathrm{tr}[\tau^a]\mathrm{tr}[t^c t^d] = 0, \tag{2.21}$$

since the trace of a single τ^a vanishes. Thus the conservation of the axial isospin currents is unaffected by the Adler-Bell-Jackiw anomaly of QCD. However, in the case of the isospin singlet axial current, the matrix τ^a is replaced by the matrix 1 on flavours, and we find

$$\partial_\mu j^{\mu 5} = -\frac{g^2 n_f}{32\pi^2}\varepsilon^{\alpha\beta\mu\nu}F^c_{\alpha\beta}F^c_{\mu\nu} \neq 0, \tag{2.22}$$

where n_f is the number of flavours; $n_f = 2$ in our current model.

Thus, the isospin singlet axial current is not in fact conserved in QCD. The divergence of this current is equal to a gluon operator with nontrivial matrix elements between hadron states. Some subtle questions remain concerning the effects of this operator. In particular, it can be shown, as we saw for the two-dimensional axial anomaly, that the right-hand side is a total divergence. Nevertheless, again in accord with our experience in two dimensions, there are physically reasonable field configurations in which the four-dimensional integral of this term takes a nonzero value. This topic is discussed further in chapter 3. In any event, the last equation indeed implies that QCD has no isosinglet axial symmetry and no associated Goldstone boson. This equation explains why the strong interaction contain no light isosinglet pseudoscalar meson with mass comparable to that of the pions, $m_{\eta'} = 960$ MeV $\geq m_\pi$.

Though the axial isospin currents have no axial anomaly from QCD interactions, they do have an anomaly associated with the coupling of quarks to electromagnetism. Again, referring to the diagrams of Fig. 2.1, we see that the electromagnetic anomaly of the axial isospin currents is given by

$$\partial_\mu j^{\mu 5a} = -\frac{e^2}{16\pi^2}\varepsilon^{\alpha\beta\mu\nu}F_{\alpha\beta}F_{\mu\nu}\cdot \mathrm{tr}[\tau_a Q^2], \tag{2.23}$$

where $F_{\mu\nu}$ is the electromagnetic field strength, Q is the matrix of quark electric charges,

$$Q = \begin{pmatrix} \frac{2}{3} & 0 \\ 0 & -\frac{1}{3} \end{pmatrix}, \tag{2.24}$$

and the trace again runs over flavours and colors. Since the matrices in the trace do not depend on color, the color sum simply gives a factor of 3. The flavor trace is nonzero only for $a = 3$,

$$\mathrm{tr}[\tau^a Q^2] = \frac{1}{2}\delta^{a3}\mathrm{tr}\begin{pmatrix} \frac{4}{9} & 0 \\ 0 & -\frac{1}{9} \end{pmatrix} = \frac{1}{6}\delta^{a3}, \quad , \tau^3 = \frac{1}{2}\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}; \tag{2.25}$$

in that case, the electromagnetic anomaly is

$$\partial_\mu j^{\mu 53} = -\frac{e^2}{32\pi^2}\varepsilon^{\alpha\beta\mu\nu}F_{\alpha\beta}F_{\mu\nu}. \tag{2.26}$$

Because the current $j^{\mu 53}$ annihilates a π^0 meson, the last equation indicates that the axial vector anomaly contributes to the matrix element for the decay $\pi^0 \to 2\gamma$. We will show that, in fact, it gives the leading contribution to this amplitude. Again, we work in the limit of massless u, d quarks, so that the chiral symmetries are exact up to the effects of the anomaly.

Consider the matrix element of the axial current between the vacuum and a two-photon state:

$$\langle p, k | j^{\mu 53} | 0 \rangle = \varepsilon^*_\nu \varepsilon^*_\lambda \mathcal{M}^{\mu\nu\lambda}(p, k). \tag{2.27}$$

This is the same matrix element that we studied in QED perturbation theory. Now, however, we will study the general properties of this matrix element by expanding it in

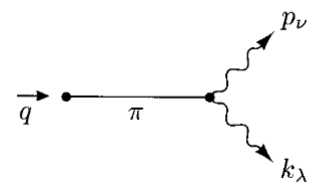

Figure 2.2: Contribution that leads to a pole in the axial vector current form factor $\mathcal{M}_1$.

form factors. In general, the amplitude can be decomposed by writing all possible tensor structures and applying the restrictions that follow from symmetry under the interchange of (p, ν) and (k, λ) and the QED Ward identities. This leaves three possible structures:

$$\mathcal{M}^{\mu\nu\lambda} = q^\mu \varepsilon^{\nu\lambda\alpha\beta} p_\alpha k_\beta \mathcal{M}_1 + (\varepsilon^{\mu\nu\alpha\beta} k^\lambda - \varepsilon^{\mu\lambda\alpha\beta} p^\nu) k_\alpha p_\beta \mathcal{M}_2 \tag{2.28}$$

$$+ [(\varepsilon^{\mu\nu\alpha\beta} p^\lambda - \varepsilon^{\mu\lambda\alpha\beta} k^\nu) k_\alpha p_\beta - \varepsilon^{\mu\nu\lambda\sigma}(p-k)_\sigma p \cdot k] \mathcal{M}_3. \tag{2.29}$$

The second term satisfies $p_\nu \mathcal{M}^{\mu\nu\lambda} = k_\lambda \mathcal{M}^{\mu\nu\lambda} = 0$ by virtue of the on-shell conditions $p^2 = k^2 = 0$.

Now contract with (iq_μ) to take the divergence of the axial vector current. We find

$$iq_\mu \mathcal{M}^{\mu\nu\lambda} = iq^2 \varepsilon^{\nu\lambda\alpha\beta} p_\alpha k_\beta \mathcal{M}_1 - i\varepsilon^{\mu\nu\lambda\sigma} q_\mu (p-k)_\sigma p \cdot k \mathcal{M}_3; \tag{2.30}$$

the other terms automatically give zero. Using $q = p + k$, $q^2 = 2p \cdot k$, we can simplify this to

$$iq_\mu \mathcal{M}^{\mu\nu\lambda} = iq^2 \varepsilon^{\nu\lambda\alpha\beta} p_\alpha k_\beta (\mathcal{M}_1 + \mathcal{M}_2). \tag{2.31}$$

The whole quantity is proportional to q^2 and apparently vanishes in the limit $q^2 \to 0$. This contrasts with the prediction of the axial vector anomaly. Taking the matrix element of the right-hand side of $\partial_\mu j^{\mu 53}$, find

$$iq_\mu \mathcal{M}^{\mu\nu\lambda} = -\frac{e^2}{4\pi^2} \varepsilon^{\nu\lambda\alpha\beta} p_\alpha k_\beta. \tag{2.32}$$

The conflicts can be resolved if one of the form factors contains a pole in q^2. Such a pole can arise through the process shown in Fig. 2.2, in which the current creates a π^0 meson with subsequently decays to two photons. The amplitude for the current to create the meson is given by $\langle 0 | j^{\mu 5a}(x) | \pi^b(x) \rangle$. Let us parametrize the pion decay amplitude as

$$i\mathcal{M}(\pi^0 \to 2\gamma) = iA \varepsilon^*_\nu \varepsilon^*_\lambda \varepsilon^{\nu\lambda\alpha\beta} p_\alpha k_\beta, \tag{2.33}$$

where A is a constant to be determined. Then the contribution of this process for Fig. 2.2 to the amplitude $\mathcal{M}^{\mu\nu\lambda}$ defined by $\langle p, k | | j^{\mu 53}(q) | 0 \rangle$ is

$$(iq^\mu f_\pi) \frac{i}{q^2} (iA \varepsilon^{\mu\lambda\alpha\beta} p_\alpha k_\beta). \tag{2.34}$$

This is a contribution to the form factor $\mathcal{M}_1$,

$$\mathcal{M}_1 = \frac{-i}{q^2} f_\pi \cdot A, \tag{2.35}$$

plus terms regular at $q^2 = 0$. Now, we can determine A in terms of the coeffecient of the anomaly:

$$A = \frac{e^2}{4\pi^2} \frac{1}{f_\pi}. \tag{2.36}$$

From the decay matrix element $i\mathcal{M}(\pi^0 \to 2\gamma)$, it is straightforward to work out the decay rate of π^0. Note that, though we have worked out the decay matrix element in the limit of a massless π^0, we must supply the physically correct kinematics which depends on the π^0 mass. Including a factor 1/2 for the phase space of identical particles, we find

$$\Gamma(\pi^0 \to 2\gamma) = \frac{1}{2m_\pi} \frac{1}{8\pi} \frac{1}{2} \sum_{\text{pols.}} |\mathcal{M}(\pi^0 \to 2\gamma)|^2 \tag{2.37}$$

$$= \frac{1}{32\pi m_\pi} \cdot A^2 \cdot 2(p \cdot k)^2 \tag{2.38}$$

$$= A^2 \cdot \frac{m_\pi^3}{64\pi}. \tag{2.39}$$

Thus, finally,

$$\Gamma(\pi^0 \to 2\gamma) = \frac{\alpha^2}{64\pi^3} \frac{m_\pi^3}{f_\pi^2}. \tag{2.40}$$

This relation, which provies a direct measurem of the coefficient of the Adler-Bell-Jackiw anomaly, is satisfied experimentally to an accuracy of a few percent.

2.2 σ-Model

It is instructive to see the abstract quantities of chiral symmetry emerging naturally in a specific example. The simplest such example is the linear σ-model. It is a seeming counterexample to effective theories because it is as renormalizable quantum field theory describing the spontaneous breaking of chiral symmetry. Rewriting it in the form of a non-decoupling effective field theory will bring the ingredients of spontaneously broken chiral symmetry to the surface. Although it has the right symmetries by construction, the linear σ-model is not general enough to describe the real world. It serves the purpose of a toy model, but it should not be mistaken for the effective field theory of QCD at low energies.

We rewrite the σ-model Lagrangian for the pion-nucleon system,

$$\mathcal{L}_\sigma = \frac{1}{2}(\partial_\mu \sigma \partial^\mu \sigma + \partial_\mu \vec{\pi} \partial^\mu \vec{\pi}) - \frac{\lambda}{4}(\sigma^2 + \vec{\pi}^2 - v^2)^2 + \bar{\psi} i \partial\!\!\!/ \psi - g\bar{\psi}(\sigma + i\vec{\tau}\vec{\pi}\gamma_5)\psi \tag{2.41}$$

with the nucleon consisting of

$$\psi = \begin{pmatrix} p \\ n \end{pmatrix} \tag{2.42}$$

in the form

$$\mathcal{L}_\sigma = \frac{1}{4}\mathrm{tr}(\partial_\mu \Sigma \partial^\mu \Sigma) - \frac{\lambda}{16}\left(\mathrm{tr}(\Sigma^\dagger \Sigma) - 2v^2\right)^2 + \bar{\psi}_L i \partial\!\!\!/ \psi_L + \bar{\psi}_R i \partial\!\!\!/ \psi_R - g\bar{\psi}_R \Sigma \psi_L - g\bar{\psi}_L \Sigma^\dagger \psi_R \tag{2.43}$$

using

$$\Sigma = \sigma \mathbb{1} - i\vec{\tau}\vec{\pi}, \quad (\sigma^2 + \vec{\pi}^2) = \frac{1}{2}\mathrm{tr}(\Sigma^\dagger \Sigma) \tag{2.44}$$

to exhibit the chiral symmetry $G = SU(2)_L \times SU(2)_R$:

$$\psi_A \xrightarrow{G} g_A \psi_A, \quad g_A \in SU(2)_A \ (A = L, R), \quad \Sigma \xrightarrow{G} g_R \Sigma g_L^{-1}.$$

For $v^2 > 0$, the chiral symmetry is spontaneously broken and the "physical" fields are the massive field $\hat{\sigma} = \sigma - v$ and the Goldstone fields $\vec{\pi}$. The Lagrangian with its non-derivative couplings for the fields $\vec{\pi}$ seems to be at variance with the Goldstone theorem predicting a vanishing amplitude whenever the momentum of a Goldstone boson goes to zero.

As an example, we will now state the linear and non-linear representation of $\mathcal{L}_\sigma$:

1. Linear representation:

   Given a linear reparametrization for our σ field, $\sigma = v + \tilde{\sigma}$, and with Goldstone bosons π, we arrive at the spontaneously broken Lagrangian given by

$$\mathcal{L}_\sigma = \frac{1}{2}((\partial_\mu \tilde{\sigma})(\partial^\mu \tilde{\sigma}) - 2\mu^2 \tilde{\sigma}^2) + \frac{1}{2}(\partial_\mu \vec{\pi})(\partial^\mu \vec{\pi}) - \lambda v \tilde{\sigma}(\tilde{\sigma}^2 + \vec{\pi}^2) \tag{2.45}$$

$$- \frac{\lambda}{4}(\tilde{\sigma}^2 + \vec{\pi}^2)^2 + \bar{\psi}(i\partial\!\!\!/ - gv)\psi - g\bar{\psi}(\tilde{\sigma} + i\vec{\tau}\vec{\pi}\gamma_5)\psi. \tag{2.46}$$

   We see the effect of spontaneous symmetry breaking, the nucleon ψ and the $\tilde{\sigma}$ fields have gained mass,

$$m_\psi = gv, \quad m_{\tilde{\sigma}} = 2\mu^2 = 2\lambda v, \tag{2.47}$$

   while the Goldstone bosons $\vec{\pi}$ have remained massless which we interpret as $SU(2)$ gauge bosons. Please note that λ and g need to be small parameters.

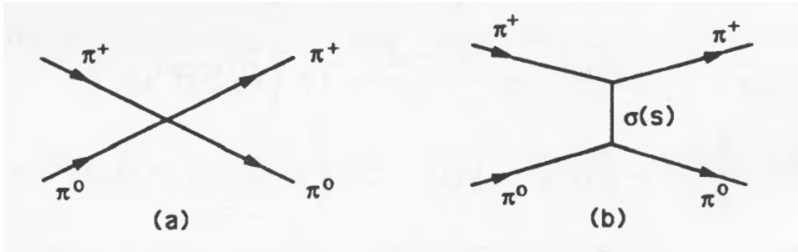

Figure 2.3: Contributions to $\pi^+\pi^0$ elastic scattering.

2. Non-linear representation:

   In order to make the Goldstone theorem manifest in the Lagrangian using a non-linear representation, we perform a field transformation from the original fields $\psi, \sigma, \vec{\pi}$ to a new set $\Psi, S, \vec{\phi}$ through a polar decomposition of the matrix field Σ where we have a small deviation of $\vec{\phi} = \vec{\pi} + \dots$:

$$U(\phi) = \exp i\frac{\vec{\tau}\vec{\phi}}{v}, \quad \Sigma = (v+S)U(\phi), \tag{2.48}$$

$$S^\dagger = S,\ U^\dagger = U^{-1},\ \det U = 1,\ u^2 = U,\ \Psi_L = u\psi_L,\ \Psi_R = u^\dagger\psi_R,\ U \to g_R U g_L^{-1} \tag{2.49}$$

   In the new fields, the σ model takes the form:

$$\mathcal{L}_\sigma = \frac{1}{2}[(\partial_\mu S)^2 - 2\mu^2 S^2] + \frac{(v+S)^2}{4}\mathrm{tr}(\partial_\mu U \partial^\mu U^\dagger) \tag{2.50}$$

$$- \lambda v S^3 - \frac{\lambda}{4}S^4 + \bar{\Psi} i\partial\!\!\!/ \Psi - g(v+s)((\bar{\Psi}_L U \Psi_R) + (\bar{\Psi}_R U^\dagger \Psi_L)). \tag{2.51}$$

   As expected, we only have derivative couplings of Goldstone bosons (π).

Representation independence We have introduced two sets of interactions with very different appearances. They are all nonlinearly related. In each of these forms the free particle sector, found by looking at terms bilinear in the field variables, has the same masses and normalizations. To compare their dynamical content, let us calculate the scattering of the Goldstone bosons of the theory, specifically $\pi^+\pi^0 \to \pi^+\pi^0$. The diagrams that enter at tree level are displayed in Fig. 2.3. The relevant terms in the Lagrangians and their tree-level amplitudes are as follows.

1. Linear representation: Given an interaction by

$$\mathcal{L}_{int} = -\frac{\lambda}{4}(\vec{\pi}^2)^2 - \lambda v \tilde{\sigma} \vec{\pi}^2, \tag{2.52}$$

we calculate an amplitude given by

$$\mathcal{M} = -2i\lambda + (-2i\lambda v)^2 \frac{i}{q^2 - m_{\tilde{\sigma}}^2} = -2i\lambda \left[1 + \frac{2\lambda v^2}{q^2 - 2\lambda v^2}\right] = \frac{iq^2}{v^2} + \mathcal{O}\left(\frac{q^4}{v^4}\right), \quad (2.53)$$

where $q = p'_+ - p_+ = p_0 - p'_0$ and the relation $m_{\tilde{\sigma}}^2 = 2\lambda v^2 = 2\mu^2$ has been used. The contributions of Fig. 2.3. are seen to cancel at $q^2 = 0$. Thus, to leading order, the amplitude is momentum-dependent even though the interaction contains no derivatives. The vanishing of the amplitudes at zero momentum is universal in the limit of exact chiral symmetry.

2. Exponential representation: Given an interaction by

$$\mathcal{L}_{int} = \frac{(v+S)^2}{4} \mathrm{tr}(\partial_\mu U \partial^\mu U^\dagger) = \frac{1}{6v^2}\left[(\vec{\phi}\partial_\mu\vec{\phi})^2 - \vec{\phi}^2(\partial_\mu\vec{\phi}\partial^\mu\vec{\phi})\right] + g(S), \quad (2.54)$$

where, again, Fig. 2.3 (b) has a higher order $\mathcal{O}(p^4)$ contribution, leaving only Fig. 2.3 (a),

$$\mathcal{M} = \frac{i(p'_+ - p_+)^2}{v^2} + \ldots. \quad (2.55)$$

The lesson to be learned is that both representations give the same answer despite very different forms and even different Feynman diagrams. A similar conclusion would follow for any other observable that one might wish to calculate.

The above analysis demonstrates a powerful field theoretic theorem, proved first by R. Haag, on representation independence. It states that if two fields are related nonlinearly, e.g. $\phi = \chi F(\chi)$ with $F(0) = 1$, then the same experimental observables result if one calculates with the field ϕ using $\mathcal{L}(\phi)$ or instead with χ using $\mathcal{L}(\chi F(\chi))$. The proof consists basically of demonstrating that two S-matrices are equivalent if they have the same single particle singularities, and since $F(0) = 1$, ϕ and χ have the same free field behaviour and single particle singularities. This result can be made plausible if we think of the scattering in non-mathematical terms. If the free particles are isolated they have the same mass and charge, and experiment cannot tell the ϕ particle from the χ particle. At this level they are in fact the same particles, due to $F(0) = 1$. The scattering experiment is then performed by colliding the particles. The results cannot depend on whether a theorist has chosen to calculate the amplitude using the ϕ or the χ names. That is, the physics cannot depend on a labeling convention.

This result is quite useful as it lets us employ nonlinear representations in situations where they can simplify the calculation. The linear sigma model is a good example. We have seen that the amplitude of this theory are momentum-dependent. Such behaviour is obtained naturally when one uses the nonlinear representations, whereas for the linear representation more complicated calculations involving assorted cancelations of constant terms are required to produce the correct momentum dependence. In addition, the non-linear representations allow one to display the low energy results of the theory without explicitly including the massive $\tilde{\sigma}$ (or S) and ψ fields.

2.3 Chiral Perturbation Theory

Now we want to couple our purely strongly interacting theory with the electroweak interactions. We know that pions, protons and neutrons interact both strongly and electroweakly.

The QCD Lagrangian with $N_f (N_f = 2$ or $3)$ massless quarks $q = (u, d, \dots)$

$$\mathcal{L}^0_{QCD} = \bar{q} i \gamma^\mu \left(\partial_\mu + i g_s \frac{\lambda_\alpha}{2} G^\alpha_\mu \right) q - \frac{1}{4} G^\alpha_{\mu\nu} G^{\alpha\mu\nu} \tag{2.56}$$

$$= \bar{q}_L i \not{D} q_L + \bar{q}_R i \not{D} q_R - \frac{1}{4} G^\alpha_{\mu\nu} G^{\alpha\mu\nu} \tag{2.57}$$

$$q_{R,L} = \frac{1}{2}(1 \pm \gamma_5) q \tag{2.58}$$

has a global symmetry

$$\underbrace{SU(N_f)_L \times SU(N_f)_R}_{\text{chiral group } G} \times U(1)_V \times U(1)_A.$$

At the effective hadronic level, the quark number symmetry $U(1)_V$ is realized as baryon number. The axial $U(1)_A$ is not a symmetry at the quantum level due to the Abelian anomaly. The Noether currents of the chiral group G are

$$J^{a\mu}_A = \bar{q}_A \gamma^\mu \frac{\lambda_a}{2} q_A \quad (A = L, R;\ \alpha = 1, \dots, N_f^2 - 1). \tag{2.59}$$

A classical symmetry can be realized in quantum field theory in two different ways depending on how the vacuum responds to a symmetry transformation. All theoretical and phenomenological evidence suggests that the chiral group G is spontaneously broken to the vectorial subgroup $SU(N_f)_V$. The axial generators of G are non-linearly realized and there are $(N_f^2 - 1)$ massles pseudoscalar Goldstone bosons. There is a well-known procedure how to realize a spontaneously broken symmetry on quantum fields. In the special case of chiral symmetry with its parity transformation, the Goldstone fields can be collected in an unitary matrix field $U(\phi)$ transforming as

$$U(\phi) \xrightarrow{G} g_R U(\phi) g_L^{-1}, \quad (g_L, g_R) \in G \tag{2.60}$$

under chiral rotations. There are different parametrizations of $U(\phi)$ corresponding to different choices of coordinates for the chiral coset space $SU(N_f)_L \times SU(N_f)_R / SU(N_f)_V$.

In the case of non-linear exponential parametrization, we can write

$$U(\phi) = \exp\left(i \frac{\vec{t} \cdot \vec{\phi}}{v} \right) \tag{2.61}$$

where $\vec{\phi}$ is a vector with the 3 components of the pion field and where the product $\vec{t} \cdot \vec{\phi}$ describes a given transformation.

For $N_f = 2$ with $\vec{t} = \vec{\tau}$ we have

$$\frac{1}{\sqrt{2}}\vec{\tau} \cdot \vec{\phi} = \begin{pmatrix} \frac{\pi^0}{\sqrt{2}} & \pi^+ \\ \pi^- & -\frac{\pi^0}{\sqrt{2}} \end{pmatrix}. \tag{2.62}$$

For $N_f = 3$ with $\vec{t} = \vec{\lambda}$ we have

$$\frac{1}{\sqrt{2}}\vec{\lambda} \cdot \vec{\phi} = \begin{pmatrix} \frac{\pi^0}{\sqrt{2}} + \frac{\eta^8}{\sqrt{6}} & \pi^+ & K^+ \\ \pi^- & -\frac{\pi^0}{\sqrt{2}} + \frac{\eta^8}{\sqrt{6}} & K^0 \\ K^- & \bar{K}^0 & -\frac{2\eta^8}{\sqrt{6}} \end{pmatrix}. \tag{2.63}$$

The Lagrangian of the Standard Model is not chiral invariant. The chiral symmetry of the strong interactions is broken by the electroweak interactions generating in particular non-zero quark masses. The basic assumptions of χPT is that the chiral limit constitutes a realistic starting point for a systematic expansion in chiral symmetry breaking interactions.

How do we include the other forces? Since all these fields are given by vector fields in the background we can parametrize by introducing external fields. We have separate left- and right-handed couling because the weak coupling is chiral. To incorporate this chiral structure, we write an effective field theory by absorbing the projector of a specific chirality state in the field definition. In this way we can fill our space with external fields for the electromagnetic and weak forces as vector and axial vector fields.

We extend the chiral invariant QCD Lagrangian by coupling to the external hermitian matrix fields v_μ, a_μ, s, p (vector, axial vector, scalar and pseudo-scalar):

$$\mathcal{L} = \mathcal{L}^0_{QCD} + \bar{q}\gamma^\mu(v_\mu + a_\mu\gamma_5)q - \bar{q}(s - ip\gamma_5)q. \tag{2.64}$$

The external field method has two major advantages:

1. External photons and W boson fields are among the gauge fields v_μ, a_μ ($N_f = 3$):

$$r_\mu = v_\mu + a_\mu = -eQA^{ext}_\mu \tag{2.65}$$

$$l_\mu = v_\mu - a_\mu = -eQA^{ext}_\mu - \frac{e}{\sqrt{2}\sin\theta_W}\left(W^{ext,+}_\mu T_+ + h.c.\right) \tag{2.66}$$

$$Q = \frac{1}{3}\mathrm{diag}(2,-1,-1), \;\; T_+ = \begin{pmatrix} 0 & V_{ud} & V_{us} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{2.67}$$

Q is the quark charge matrix, the V_{ij} are the Kobayashi-Maskawa mixing matrix elements, the only allowed charge modifying couplings are $d \to u, s \to u$ and T_+ is the raising operator in isospin state. Green functions for electromagnetic and semileptonic weak currents can be obtained as functional derivatives of a generating functional $Z[v, a, s, p]$ with respect to external photon and W boson fields. This procedure is valid for all fields π, K, just K^0 needs the Z boson field as well. Z boson fields would fit in the same structure, although it would look a bit more complicated.

2. The scalar and pseudoscalar fields s, p give rise to Green functions of (pseudo)scalar quark currents, but they also provide a very convenient way of incorporating explicit chiral symmetry breaking through the quark masses. The physically interesting Green functions are functional derivatives of the generating functional $Z[v, a, s, p]$ at

$$v_\mu = a_\mu = p = 0$$

and

$$s = \mathcal{M}_q = \mathrm{diag}(m_u, m_d, \ldots). \tag{2.68}$$

The practical advantage is that $Z[v, a, s, p]$ can be calculated in a manifestly chiral invariant way. The actual Green functions with broken chiral symmetry are then obtained by taking appropriate functional derivatives. Loosely speaking, we can use s, p to reintroduce masses as external fields which are constant over all space.

Inclusion of external fields promotes the global chiral symmetry G to a local one:

$$q \stackrel{G}{\to} g_R \frac{1}{2}(1+\gamma_5)q + g_L \frac{1}{2}(1-\gamma_5)q \tag{2.69}$$

$$r_\mu \stackrel{G}{\to} g_R r_\mu g_R^{-1} + i g_R \partial_\mu g_R^{-1} \tag{2.70}$$

$$l_\mu \stackrel{G}{\to} g_L l_\mu g_L^{-1} + i g_L \partial_\mu g_L^{-1} \tag{2.71}$$

$$s + ip \stackrel{G}{\to} g_R (s + ip) g_L^{-1}. \tag{2.72}$$

The local nature of G requires the introduction of a covariant derivative

$$D_\mu U = \partial_\mu U - i r_\mu U + i U l_\mu, \quad D_\mu \stackrel{G}{\to} g_R D_\mu U g_L^{-1}, \tag{2.73}$$

and of associated non-Abelian field strength tensors

$$F_L^{\mu\nu} = \partial^\mu l^\nu - \partial^\nu l^\mu - i[l^\mu, l^\nu] \tag{2.74}$$

$$F_R^{\mu\nu} = \partial^\mu l r \nu - \partial^\nu r^\mu - i[r^\mu, r^\nu]. \tag{2.75}$$

External fields do not have kinetic parts. Consequently, the external fields are not affected by the spontaneous breakdown of G. On a fundamental level, the electroweak gauge symmetry is broken in two steps, at the Fermi scale and at the chiral symmetry breaking scale. Thus, there is in principle a small mixing between $\bar{q}q$ and whichever fields are responsible for the electroweak breaking at the Fermi scale (Higgs, technicolour, ...). via the Higgs-Kibble mechanism, three of those states become the longitudinal components of W and Z bosons. Here, we are only interested in the light orthogonal states, the pseudoscalar pseudo-Goldstone bosons.

We consider interaction processes in which the structure of mesons and baryons is not resolved by the strong interaction.

χPT is the low-energy effective field theory of the Standard Model. The chiral Lagrangians are organized in a derivative expansion based on the chiral counting rules

$$U \qquad \mathcal{O}(p^0) \tag{2.76}$$
$$D_\mu U, v_\mu, a_\mu \qquad \mathcal{O}(p^1) \tag{2.77}$$
$$F^{\mu\nu}_{L,R} \qquad \mathcal{O}(p^2). \tag{2.78}$$

χPT is also an expansion in quark masses around the chiral limit. In principle, one can formulate χPT as an independent expansion in both derivatives and quark masses. It is convenient, however, to combine these two expansions in a single one by making use of the relations between meson and quark masses. Standard χPT is defined by the simplest choice corresponding to the counting rule

$$s, p \quad \mathcal{O}(p^2) \tag{2.79}$$

for the scalar and pseudoscalar external fields.

The locally chiral invariant Lagrangian of lowest order describing strong, electromagnetic and semileptonic weak interactions of mesons is given by

$$\mathcal{L}_2 = \frac{v^2}{4}\mathrm{tr}\left(D_\mu U D^\mu U^\dagger + \chi U^\dagger + \chi^\dagger U\right), \quad \chi = 2B(s+ip). \tag{2.80}$$

The two low energy constants (LECs) of $\mathcal{O}(p^2)$ are related to the pion decay constant and to the quark condensate in the chiral limit:

$$v = f_\pi(1+\mathcal{O}(m_q)) = 92.4 \text{ MeV}\langle 0|\bar{u}u|0\rangle = -v^2 B. \tag{2.81}$$

This theory is nonrenormalizable, meaning that at higher orders divergences show up which cannot be absorbed into parameter redefinitions. One can, however, consider the extension to $\mathcal{O}(p^4)$, then the divergences of the one-loop diagrams of the theory given by $\mathcal{L}_2$ can be absorbed into the parameters of $\mathcal{L}_2 + \mathcal{L}_4$, resulting in a consistent theory with considerably more parameters.

Baryons can be described in χPT with

$$u = U^{1/2}\exp\left(\frac{i}{2v}\vec{\tau}\cdot\vec{\pi}\right) \quad \text{for } SU(2) \tag{2.82}$$

and $\vec{\lambda} \cdot \vec{\phi}$ for $SU(3)$ correspondingly. With this we can define a vector field including the external gauge fields

$$u_\mu = i\left[u^\dagger(\partial_\mu - ir_\mu)u - u(\partial_\mu - il_\mu)u^\dagger\right] \tag{2.83}$$

which allows us to construct a chiral meson baryon Lagrangian of order $\mathcal{O}(p)$.

The chiral Lagrangians at $\mathcal{O}(p)$ for the Fermi theory of πN and for MB are given by

$$\mathcal{L}^{(1)}_{\pi N} = \bar{N}(i\not{\partial} - M + \frac{g_A}{2}\not{u}\gamma_5)N \qquad (N_f = 2) \tag{2.84}$$

$$\mathcal{L}^{(1)}_{MB} = \mathrm{tr}\left[\bar{B}(i\not{\partial} - M)B + \frac{d}{2}\bar{B}\gamma^\mu\{u_\mu, B\} + \frac{f}{2}\bar{B}\gamma^\mu\gamma_5[u_\mu, B]\right] \qquad (N_f = 3) \tag{2.85}$$

The Lagrangian $\mathcal{L}^{(1)}_{\pi N}$ is of the form expected from the discussion of the linear σ-model. B is the octet given by

$$B = \begin{pmatrix} \frac{\Sigma^0}{\sqrt{2}} + \frac{\Lambda}{\sqrt{6}} & \Sigma^+ & p \\ \Sigma^- & -\frac{\Sigma^0}{\sqrt{2}} + \frac{\Lambda}{\sqrt{6}} & n \\ \Xi^- & \Xi^0 & -\frac{2\Lambda}{\sqrt{6}} \end{pmatrix}. \tag{2.86}$$

At $\mathcal{O}(p)$, there are two (three) LECs for $N_f = 2(3)$: M is the nucleon (baryon) mass and g_A is the nucleon axial-vector coupling constant in the chiral limit (experimentally determined, $g_A = 1.25$, by pure weak processes, i.e. neutron decay). The axial $SU(3)$ coupling constants f, d are related to g_A via

$$g_A = f + d. \tag{2.87}$$

For $n_f = 3$ the baryons are in an octet representation of $SU(3)$ whereas for $n_f = 2$, the baryones (p, n) transform as a doublet under $SU(2)$.

Note that we have created mass spontaneously by chiral symmetry breaking, thus the mass-matrix is computable in terms of σ-models extended to $SU(3)$. Also recall that we are doing perturbation theory expanding quark masses in (small) momentum, and not in the coupling constant as in the usual way.

2.4 The Θ-Vacuum

Typically we can represent a finite gauge transformation by the exponential of an infinitesimal gauge transformations (see Lie Algebras and Lie Groups), which means that the gauge transformation is related to the identity continuously.

The action of such a finite gauge transformation is, generated by parameter α,

$$\psi \to U(\alpha)\psi \tag{2.88}$$

$$D_\mu\psi = (\partial_\mu + igA_\mu)\psi \tag{2.89}$$

with $A_\mu \to UA_\mu U^{-1} + \frac{i}{g}(\partial_\mu U)U^{-1}$.

For continuous transformations, we have $U = \exp\left[-i\alpha^a(x)T^a\right]$ with T^a generators.

However, for non-Abelian groups there exists a different class of gauge transformations, finite discrete transformations, i.e. parity transformation. It cannot be continuously related to identity in odd dimensions, hence discrete:

$$P = \text{diag}(-1,-1,-1),\ \ P \in O(3), P \notin SO(3). \tag{2.90}$$

The θ vacuum One is used to consider the effect on gluon fields of 'small' gauge transformations, i.e. those which are connected to the identity operator in a continuous manner. There also exists 'large' gauge transformations which change the color gauge fields in a more drastic fashion. For example the gauge transformation generated by

$$\Lambda_1(\mathbf{x}) = \frac{\mathbf{x}^2 - d^2}{\mathbf{x}^2 + d^2} + \frac{2id\tau\cdot\mathbf{x}}{\mathbf{x}^2 + d^2}, \tag{2.91}$$

where d is an arbitrary parameter and τ is an $SU(2)$ Pauli matrix in any $SU(2)$ subgroup of $SU(3)$, e.g.

$$\tau_1 = \lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tau_2 = \lambda_2, \tau_3 = \lambda_3,$$

(other embeddings possible, but for simplicity we choose this, other's are just a linear cominbation of other λ's), transforms the null potential $\mathbf{A}(\mathbf{x}) = 0$ into

$$A_j^{(1)}(\mathbf{x}) = -\frac{i}{g}(\nabla_j\Lambda_1(\mathbf{x}))\Lambda_1^{-1}(\mathbf{x}) \tag{2.92}$$

$$= -\frac{2d}{g(\mathbf{x}^2+d^2)^2}\left[\tau_j(d^2-\mathbf{x}^2) + 2x_j(\tau\cdot\mathbf{x}) - 2d(\mathbf{x}\times\tau)_j\right], A_0^{(1)}(\mathbf{x}) \quad = 0. \tag{2.93}$$

Note that for $d \to 0$ this gauge transformation does not vanish, in fact, it becomes singular at the origin. It cannot be deformed into unity continuously. Here, we are using the matrix notation

$$\mathbf{A}_\mu = A_\mu^a\frac{\lambda^a}{2}. \tag{2.94}$$

This potential lies in an $SU(2)$ subgroup of the full color $SU(3)$ group, and is 'large' in the sense that it cannot be deformed continuously into the identity. The $\tau\cdot\mathbf{x}$ factor couples the internal color indices to the spatial position such that a path in coordinate space implies a corresponding path in the $SU(2)$ color subspace. We can associate with a gauge potential A a topological charge called winding number, invariant under small gauge transformations,

$$n = \frac{ig_3^3}{24\pi^2} \int \mathrm{d}^3x \mathrm{tr}\left[A_i(\mathbf{x})A_j(\mathbf{x})A_k(\mathbf{x})\right] \varepsilon^{ijk}. \tag{2.95}$$

As can be demonstrated by direct substitution, the gauge field of $A_j^{(1)}$ corresponds to the value $n = 1$. Fields with integer vlaue of the winding number n can be obtained by repeated application of $\Lambda_1(\mathbf{x})$,

$$\Lambda_n(\mathbf{x}) = [\Lambda_1(\mathbf{x})]^n. \tag{2.96}$$

All gauge potentials can be classified into disjoint sectors labeled by their winding number.

The existence of these distinct classes has interesting consequences. For example, consider a configuration of the gluon field that starts off at $t = -\infty$ as the zero potential $\mathbf{A}(\mathbf{x}) = 0$, has some interpolating $\mathbf{A}(\mathbf{x}, t)$ for intermediate times, and ends up at $t = +\infty$ lying in the gauge equivalent configuration $\mathbf{A}(\mathbf{x}) = A^{(1)}(x)$. In other words, we discusss an adiabatic change. Then the following integral can be shown to be nonvanishing:

$$\frac{g_3^3}{32\pi^2} \int \mathrm{d}^4x \; F_{\mu\nu}^a \tilde{F}^{a\mu\nu} \quad (\tilde{F}^{a\mu\nu} = \frac{1}{2}\varepsilon^{\mu\nu\alpha\beta} F_{\alpha\beta}^a). \tag{2.97}$$

This is surprising because the integrand is a total divergence. As noted in electromagnetism, $F\tilde{F}$ can be rewritten as

$$F_{\mu\nu}^a \tilde{F}^{a\mu\nu} = \partial_\mu K^\mu, \quad K^\mu = 2\varepsilon^{\mu\nu\lambda\sigma}(A_\nu^a F_{\lambda\sigma}^a + \frac{1}{3} g f_{abc} A_\nu^a A_\lambda^b A_\sigma^c), \tag{2.98}$$

and thus the integral can be written as a surface integral at $t = \pm\infty$. For the field configuration under consideration, this reduces to the winding number integral

$$\begin{aligned}
\frac{g^2}{32\pi^2} \int \mathrm{d}^4x \; F_{\mu\nu}^a \tilde{F}^{a\mu\nu} &= \frac{g^2}{32\pi^2} \int \mathrm{d}^4x \; \partial_\mu K^\mu & (2.99)\\
&= \frac{g^2}{32\pi^2} \int \mathrm{d}^3x \; K_0|_{t=-\infty}^{t=\infty} & (2.100)\\
&= \frac{g^2}{24\pi^2} i \int \mathrm{d}^3x \; \varepsilon^{ijk} \mathrm{tr}\left[\mathbf{A}_i^{(1)}(x)\mathbf{A}_j^{(1)}(x)\mathbf{A}_k^{(1)}(x)\right] & (2.101)\\
&= 1. & (2.102)
\end{aligned}$$

More generally, the integral of $F\tilde{F}$ gives the change in the winding number

$$\frac{g^2}{32\pi^2} \int \mathrm{d}^4x \; F_{\mu\nu}^a \tilde{F}^{a\mu\nu} = \frac{g^2}{32\pi^2} \int \mathrm{d}^3x \; K_0|_{t=-\infty}^{t=\infty} = n_+ - n_- \tag{2.103}$$

between asymptotic gauge field configurations.

Thus, the vacuum state vector will be characterized by configurations of gluon fields which fall into classes labeled by the winding number. Moreover, there will exist a correspondence between the gauge transformations $\{\Lambda_n\}$ and unitary operators $\{U_n\}$ which transform the state vectors. For example, a vacuum state dominated by a field configuration in the zero winding class ('near' to $A_\mu = 0$) would be transformed by U_1 into configurations with a dominance of $n = 1$ configurations, or more generally,

$$U_1|n\rangle = |n+1\rangle. \tag{2.104}$$

This implies that a gauge-invariant vacuum state requires contributions from all classes, such as the coherent superposition

$$|\theta\rangle = \sum_n e^{-in\theta}|n\rangle, \tag{2.105}$$

where θ is an arbitrary parameter. It follows from $U_1|n\rangle = |n+1\rangle$ that this θ-vacuum is gauge-invariant up to an overall phase

$$U_1|\theta\rangle = e^{i\theta}|\theta\rangle. \tag{2.106}$$

The QCD vacuum must contain contributions from all topological classes.

The θ-term Given this nontrivial vacuum structure, one requires three ingredients to completely specify QCD: the QCD Lagrangian, the coupling constant (i.e. Λ_{QCD}), and the vacuum label θ. How can we account for the different vacua corresponding to choices of θ? In a path integral representation, the $\theta = 0$ vacuum would imply generic transition elements of the form

$$_{out}\langle\theta=0|X|\theta=0\rangle_{in} = \int [\mathrm{d}A_\mu][\mathrm{d}\psi][\mathrm{d}\bar{\psi}]Xe^{iS_{QCD}} = \sum_{n,m} {}_{out}\langle m|X|n\rangle_{in}. \tag{2.107}$$

The presence of a nonzero θ leads to an extra phase,

$$_{out}\langle\theta|x|\theta\rangle_{in} = \sum_{n,m} e^{i(m-n)\theta} {}_{out}\langle m|x|n\rangle_{in}. \tag{2.108}$$

However, this phase can be accounted for in the path integral by the addition of a new term to S_{QCD}. In particular we have, through the use of the winding number,

$$_{out}\langle\theta|X|\theta\rangle_{in} = \int [\mathrm{d}A_\mu][\mathrm{d}\psi][\mathrm{d}\bar{\psi}]Xe^{iS_{QCD}+i\frac{g_3^2}{64\pi^2}\theta\int \mathrm{d}^4x\ F^a_{\mu\nu}\tilde{F}^{a\mu\nu}} \tag{2.109}$$

$$= \sum_{n,m} e^{i(m-n)} {}_{out}\langle m|X|n\rangle_{out}, \tag{2.110}$$

where X is some operator. We see that the quantity $(m-n)$ given by the winding number difference of the fields contributing to the path integral is equivalent to a new exponential factor containing $F^a_{\mu\nu}\tilde{F}^{a\mu\nu}$. Thus a correct procedure for doing calculations involving the θ-vacua is to follow the ordinary path integral methods but with the QCD Lagrangian containing the new term,

$$\mathcal{L}_{QCD} = \mathcal{L}^{(\theta=0)}_{QCD} + \theta\frac{g_3^2}{64\pi^2}F^a_{\mu\nu}\tilde{F}^{a\mu\nu}. \tag{2.111}$$

The parameter θ is to be considered a coupling constant. Since the operator $F\tilde{F}$ is P-odd and T-odd, a nonzero θ can induce measurable T violation. Later, we shall show how to connect θ to physical observables. There is an important distinction between the various θ vacua of QCD and the many possible vacuum states of a spontaneously broken symmetry such as the Higgs sector of the electroweak theory. In the latter case, the various possible vacuum expectation values of the Higgs field label different states within the same theory. In contrast, each value of θ corresponds to a different theory, just as each value of Λ_{QCD} would label a different theory. Specifying θ and Λ_{QCD} then specifies the content of the version of QCD used by nature. One of the experimental consequences is that one can have electrical dipole moments of elementary particles, but the strongest constraint comes from the neutron ($d_n^{exp} < 8\cdot 10^{-26}$ e $\cdot\, cm$), in exercises we conclude that $|\theta| \lessapprox 10^{-9}$.

Connection with the chiral rotations There is a connection between the axial anomaly and the presence of a θ-vacuum. It involves the matrix element of $F\tilde{F}$ as follows. Consider the limit of N_f massless quarks. The $U(1)$ axial current

$$j_{5\mu} = \sum_{j=1}^{N_f}\bar{\psi}_j\gamma_\mu\gamma_5\psi_j \tag{2.112}$$

is not conserved due to the anomaly,

$$\partial^\mu j_{5\mu} = \frac{N_f\alpha_s}{8\pi}F^a_{\mu\nu}\tilde{F}^{a\mu\nu}. \tag{2.113}$$

However, because of the fact that $F\tilde{F}$ is a total divergence, one can define a new conserved current

$$\tilde{j}_{5\mu} = j_{5\mu} - \frac{N_f\alpha_s}{8\pi}K_\mu. \tag{2.114}$$

While $\tilde{j}_{5\mu}$ does form a conserved charge,

$$\tilde{Q}_5 = \int \mathrm{d}^3x\ \tilde{j}_{5\mu}(x), \tag{2.115}$$

neither $\tilde{Q}_5$ nor $\tilde{j}_{5\mu}$ is gauge-invariant. In fact, under the gauge transformation of Λ_1 it follows that that the operator $\tilde{Q}_5$ changes by a c-number integer,

$$U_1\tilde{Q}_5U_1^{-1} = \tilde{Q}_5 - 2N_f. \tag{2.116}$$

This tells us that in the world of massless quarks, the different θ-vacua are related by a chiral $U(1)$ transformation,

$$U_1 e^{i\alpha\tilde{Q}_5}|\theta\rangle = U_1 \exp^{i\alpha\tilde{Q}_5} U_1^{-1} U_1|\theta\rangle = e^{i(\theta-2N_f\alpha)} e^{i\alpha\tilde{Q}_5}|\theta\rangle, \tag{2.117}$$

or,

$$e^{i\alpha\tilde{Q}_5}|\theta\rangle = |\theta - 2N_f\alpha\rangle, \tag{2.118}$$

where α is a constant. Therefore, in the limit of massles quarks, when $\tilde{Q}_5$ is a conserved quantity, all of the θ-vacua are equivalent and one can transform away the θ dependence by a chiral $U(1)$ transformation. The same is not true if quarks have mass, as the mass terms in $\mathcal{L}_{QCD}$ are not invariant under a chiral transformation.

To summarize, one finds that the existence of topologically nontrivial gauge transformations, and of field configurations which make transitions between the different topological sectors of the theory, leads to the existence of nonvanishing effects from a new term in the QCD action. Chiral rotations can change the value of θ, allowing it to be rotated away if any of the quarks are massless. However for massive quarks, the net effect is a measurable CP violating term in the QCD Lagrangian.

2.5 The Axion

The strong CP problem: If instanton effects necessarily contribute an extra parameter to QCD, then why is θ so small?

The simplest suggested solution to why θ is so small is to invoke yet another $U(1)$ symmetry, the Peccei-Quinn symmetry. The presence of this additional $U(1)$ symmetry would be sufficient to keep $\theta = 0$.

To see how this axion hypothesis works, consider the possibilitiy of CP violation in QCD caused by introducing a complex, nondiagonal mass matrix M for the quarks transforming non-trivially under $U(1)_A$,

$$M = S_L^\dagger M' S_R, \quad \psi_L = S_L^\dagger \psi_L', \quad \psi_R = S_R^\dagger \psi_R', \tag{2.119}$$

with $S_{L,R} \in U(N_f)$. We can now factor out the $U(1)$ transformation by using

$$S_{L,R} = e^{i\phi_{L,R}} \bar{S}_{L,R} \tag{2.120}$$

with $\bar{S}_{L,R} \in SU(N_f)$, which then yields a $U(1)_A$ transformation

$$S_A = \exp i(\phi_R - \phi_L) \tag{2.121}$$

such that θ transforms under chiral rotation as

$$\theta \to \theta - 2N_f(\phi_R - \phi_L). \tag{2.122}$$

Discussing the physical phase difference $\phi_L - \phi_R$ we require to obtain a real diagonal M, which is related to the CKM phases, we find:

$$\arg\ \det\ M = \arg(\det S_L^\dagger \det M' \det S_R) \tag{2.123}$$
$$= \arg\det S_L^\dagger + \arg\det S_R + \arg\det M' \tag{2.124}$$
$$= 2N_f(\phi_R - \phi_L) + \arg\det M', \tag{2.125}$$

where the $\det M'$ is the determinant of the masses, not of the CKM matrix.

Now we can state that the effective and measured CP-violating angle $\bar{\theta}$ in strong interactions is given by

$$\bar{\theta} = \theta + \arg\det M'. \tag{2.126}$$

However, from experiments we get constraints which then imply that $\bar{\theta} < 10^{-9}$, which then raises the question: why is $\bar{\theta}$ so small? If, however, we had massless quarks, we would get $\det M' = 0$. Thus, a big question being investigated for more than two decades is whether the up-quark is in fact massless.

A dynamical explanation to eliminate this effective $\bar{\theta}$ term is given by adding a new field σ to the QCD action given by:

$$\mathcal{L}_{axion} = \bar{\psi}(Me^{-i\frac{\sigma}{v}})\psi + \frac{1}{2}\partial_\mu\sigma\partial^\mu\sigma \tag{2.127}$$

where σ is the axion field, which couples to the quark mass term via the phase factor. The axion arises as a Nambu-Goldstone boson of the new broken $U(1)$ symmetry of the quark and the Higgs sector.

Now perform another axial $U(1)$ transformation on the quark fields that eliminates the $F\tilde{F}$ term entirely and puts all CP violating terms in the mass matrix. We then find that the mass term in QCD is multiplied by

$$\exp i\left(\theta + \arg\det\left(M' - \frac{\sigma}{v}\right)\right) \tag{2.128}$$

which can be absorbed by shifting σ to

$$\sigma \to \sigma + v(\theta + \arg\det M'). \tag{2.129}$$

Since the axion is massless, the kinetic term is invariant under this shift, so the shift is sufficient to absorb all CP violating terms that appear exclusively in the mass matrix.

In this way, the introduction of a massless axion field, to lowest order, can absorb all strong CP violating effects by a shift. The motivation is again the minimization of the vacuum energy (which corresponds to the minimum of vacuum energy) because any term of the form

$$\frac{g^2}{32\pi^2}F^a_{\mu\nu}\tilde{F}^{a\mu\nu}$$

only increase the vacuum energy.

To recapitulate, we have taken the θ into the Lagrangian and we have made it dynamical. By introducing the axion field, we have added two new parameters, v and M.

Although the axion gives us a way in which the strong CP problem might be solved, experimentally the situation is still unclear. Experimental searches for the axion have been unsuccessful. In fact, the naive axion theory that we have presented can actually be experimentally ruled out. However, it is still possible to revive the axion theory if we assume that it is very light and weakly coupled. Experimentally, this "invisible axion", if it exists, should have a mass between 10^{-6} and 10^{-3} eV. The invisible axion would then be within the bounds of experiments.

We now quickly present an overview of common features of different realizations of axion models:

1. Axion is the Goldstone boson of a new, broken chiral symmetry (Peccei-Quinn Symmetry)

2. Because the chiral symmetry is not exact, the axion has mass (typically of order $m_\pi \frac{f_\pi}{v}$)

3. The effective low energy Lagrangian with $N_f = 2$ up to $\frac{1}{v}$ is given by

$$\mathcal{L}_\sigma = \frac{1}{2}\partial_\mu\sigma\partial^\mu\sigma + \frac{g}{32\pi^2}\left(\theta - \frac{\sigma}{v}\right) F^a_{\mu\nu}\tilde{F}^{a\mu\nu} - \frac{if_u}{v}\partial_\mu\sigma\bar{u}\gamma_5\gamma^\mu u - \frac{if_d}{v}\partial_\mu\sigma\bar{d}\gamma_5\gamma^\mu d. \tag{2.130}$$

Again, we can then eliminate the $F^a_{\mu\nu}\tilde{F}^{a\mu\nu}$ term by chiral rotation of the quarks given by

$$u \to e^{i\gamma_5\alpha_u}u, \quad d \to e^{i\gamma_5\alpha_d}d. \tag{2.131}$$

By choosing

$$\alpha_f = -\left(\theta - \frac{\sigma}{v}\right)\frac{c_f}{2}, \quad c_u + c_d = 1 \tag{2.132}$$

we get

$$\theta \to \theta + 2\alpha_u + 2\alpha_d. \tag{2.133}$$

Again, we have successfully eliminated the $F^\alpha_{\mu\nu}\tilde{F}^{\alpha\mu\nu}$ term. We have, however, explicit phase factors appearing in the fermionic terms,

$$\mathcal{L}_m = -m_u\bar{u}e^{-ic_u\left(\theta-\frac{\sigma}{v}\right)\gamma_5}u - m_d\bar{d}e^{-ic_d\left(\theta-\frac{\sigma}{v}\right)\gamma_5}d. \tag{2.134}$$

Since the axion field is not constant, $\sigma = \sigma(x)$, we obtain a derivative coupling as well from the quark kinetic term,

$$\mathcal{L}_{\sigma qq} = i\frac{c_u}{2v}\left(\bar{u}\gamma^\mu\gamma^5 u\right)\partial_\mu\sigma + i\frac{c_d}{2v}\left(\bar{d}\gamma^\mu\gamma^5 d\right)\partial_\mu\sigma. \tag{2.135}$$

It is also of advantage to redefine our constants,

$$f'_{u,d} = f_{u,d} - \frac{c_{u,d}}{2}. \tag{2.136}$$

4. From axion pion mixing, through the neutral π^0, in the chiral QCD Lagrangian we have the relevent parts, using $\sigma^1 = \sigma - \langle\sigma\rangle = \sigma - v\theta,\ c = -\langle 0|q\bar{q}|0\rangle$,

$$\mathcal{L} = \frac{1}{2}\partial_\mu\pi^0\partial^\mu\pi^0 + \frac{1}{2}\partial_\mu\sigma^1\partial^\mu\sigma^1 - \frac{1}{2}\begin{pmatrix}\pi^0 & \sigma^1\end{pmatrix} M_0^2 \begin{pmatrix}\pi^0 \\ \sigma^1\end{pmatrix} \tag{2.137}$$

with the non-diagonal mass matrix M (similar to the pion mass matrix)

$$M_0^2 = \begin{pmatrix} \frac{m_u+m_d}{f_\pi^2}c & \frac{-m_u c_u + m_d c_d}{f_\pi v}c \\ \frac{-m_u c_u + m_d c_d}{f_\pi v}c & \frac{m_u c_u^2 + m_d c_d^2}{v^2}c \end{pmatrix}. \tag{2.138}$$

Assuming $v \gg f_\pi$, we obtain the Eigenvalues

$$m_\pi^2 = \frac{m_d + m_u}{f_\pi^2}c \tag{2.139}$$

$$m_\sigma^2 = \frac{c}{v^2}\frac{m_u m_d}{m_u + m_d} = \frac{f_\pi^2}{v^2}\frac{m_d m_u}{(m_d + m_u)^2}m_{\pi^0}^2. \tag{2.140}$$

We thus see that the axion is indeed a very light particle,

$$m_\sigma \approx \frac{13\text{ MeV}}{v}, \tag{2.141}$$

where we remark that v is indeed not very small.

We can also use this formalism to say something about the interactions of axions with hadrons. The eigenvector of M_0^2 with eigenvalue m_σ^2 has a component along the original π^0 direction equal to $(m_u c_u - m_d c_d) f_\pi/(m_u + m_d)v$. As mentioned earlier, because of the one-pion pole this is the dominant axion-hadron coupling. We see that the ratio of the axion and pion production interaction amplitudes will typically be of order f_π/v. The fact that axions are not observed in such collisions indicates that $v > 3$ TeV, in contradiction with the original expectation that the anomalous $U(1)$ symmetry is spontaneously broken by the same scalar vacuum expectation values of order 0.3 TeV that break the electroweak $SU(2) \times U(1)$ symmetry. It is possible to explain why axions are not found in reactor or accelerator experiments by taking v as independent parameter, much larger than the electroweak breaking scale, but there are still astrophysical limitations. Limits on the rate of cooling of red giant stars give $v > 10^7$ GeV, while observations of the supernova SN1987A indicate that $v > 10^{10}$GeV. For $v > 10^7$ GeV the axion mass would be less than about 1 eV, so that stars are hot enough to produce axions. The ratio of the axion and π^0 decay rates into two photons is expected to be of order $(f_\pi/v)^2$ times a phase space ratio of order $(m_\sigma/m_\pi)^3$, or

$$\frac{\Gamma(\sigma \to 2\gamma)}{\Gamma(\pi^0 \to 2\gamma)} = \left(\frac{f_\pi}{v}\right)^2 \left(\frac{m_\sigma}{m_\pi}\right)^3 \approx \left(\frac{f_\pi}{v}\right)^5. \tag{2.142}$$

Hence for $v > 10^7$ GeV the axion lifetime is expected to be longer is expected to be longer than about 10^{24} s, which is ample time for the axion to travel even cosmological distances before decaying. Cosmological arguments suggest an upper bound $v < 10^{12}$ GeV, leaving an open but narrow window of allowed axion parameters.

2.6 Instantons

In ordinary quantum mechanics, the WKB approximation is obtained by expanding in powers of Planck's constant, $\hbar$. To zero order we have the classical trajectory; higher order yield the quantum fluctuations around this trajectory. The path integral formulation lends itself particularly well to the extension of the method to the field-theoretic case. The usefulness of the method lies in the fact that, to each order, the functional integral is of Gaussian type and can therefore be evaluated.

It is known that there are quantum mechanical situations for which no classical trajectory exists. This occurs when there is tunnelling through a potential barrier. However, one can still adapt the WKB method to cope with this situation. We will exemplify this with the typical case of a particle in one dimensions, subject to a potential $V(x)$. To leading order in the WKB approximation, the wave function is

$$\psi(x) = Ce^{i\mathcal{A}_{cl}},$$

where $\mathcal{A}_{cl}$ is now the action calculated along the classical trajectory

$$\frac{1}{2}m\ddot{x} + V(x) = E.$$

Take a potential with two minima, both corresponding to $V = 0$, and located at $x = x_0, x = x_1$, as in in figure. If $E > \max\ V$, the motion from x_0 to x_1 is possible, and $\psi(x)$ yields the "transition" or "diffusion" amplitude.

However, if $E < \max\ V$, the correct WKB analysis gives a result in which the transition amplitude

$$\langle x_1|x_0\rangle = Ce^{i\mathcal{A}_{cl}(x_1,x_0)},$$

is to be replaced by the tunnelling amplitude,

$$\langle x_1|x_0\rangle = Ce^{-\underline{A}(x_1,x_0)},$$

where the Euclidian action $\underline{\mathcal{A}}$ is not calculated along the solution of the previous equation of motion, but for

$$-\frac{1}{2}m\ddot{x} + V(x) = E.$$

We see that to obtain a tunnelling amplitude we can use the same formula as that for a transition, making only the formal replacement of t by it, both in the expression for the action,

$$\mathcal{A} = \int_{t(\xi_0)}^{x(\xi_1)} \mathrm{d}tL \rightarrow i\underline{\mathcal{A}},$$

with ξ_i the turning points, and in the equations of motions.

The tunnelling amplitude and the wave function do not give the normalization, which may, however, be disposed of by dividing by $\langle x_0|x_0\rangle$. We thus infer that, in quantum field theory, the leading tunnelling amplitude will be

$$\langle\Psi_1, t = +\infty|\Psi_0, t = -\infty\rangle \approx C\exp\left(-\int \mathrm{d}^4\underline{x}\underline{\mathcal{L}}(\underline{\phi}_{cl})\right),$$

where $\underline{\phi}_{cl}$ is the classical solution to the Euclidian equations of motions, i.e., with x^0 replaced by $\pm ix_4$, x_4 real. (The sign $\pm$ depends on the boundary conditions; the reason for the name Euclidian is that, under the transformation $x^0 \rightarrow ix_4$,the Minkowski metric becomes Euclidian, up to a global sign.)

According to the discussion at the beginning of this section, we may consider this to be the leading order of the exact expression,

$$\langle\Psi_1, t = +\infty|\Psi_0, t = -\infty\rangle = N\exp\int \mathcal{D}\underline{\phi}\left(-\int \mathrm{d}^4\underline{x}\underline{\mathcal{L}}(\underline{\phi})\right),$$

when expanding the field $\underline{\phi}$ in powers of $\hbar$ around $\underline{\phi}_{cl}$.

An important property of the states of a system in a situation when tunnelling is possible is that the stationary states (in particular, the ground state, to be identified with the vacuum in quantum field theory) are not those in which the system is localized in one minimum of the potential, but is shared by all minima. The situation is familiar in solid state theory, where the potentials are periodic.

Search: Solution of YM equation in 4d Euclidian space. Consider the energy-momentum tensor of the pure Yang-Mills QCD, leaving quarks aside, as they are irrelevant for the considerations of this. We can write it as

$$\Theta^{\mu\nu} = -\frac{1}{2}g_{\alpha\beta}\sum_a F_a^{\mu\alpha}F_a^{\nu\beta} - \frac{1}{2}g_{\alpha\beta}\sum_a \tilde{F}_a^{\mu\alpha}\tilde{F}_a^{\nu\beta}, \quad \tilde{F}^{\mu\nu} = \frac{1}{2}\varepsilon^{\mu\nu\alpha\beta}F_{\alpha,\beta}.$$

It follows that Θ^{00} is positive for real gluon fields:

$$\Theta^{00} = \frac{1}{2}\sum_{k,a}\{(F_a^{0k})^2 + (\tilde{F}_a^{0k})^2\}.$$

Therefore $\Theta^{\mu\nu} = 0$ requires $F \equiv 0$, and thus only the zero-field configurations may be identified with the vacuum. However, Θ^{00} no longer has a definite sign if we allow for complex $F^{\mu\nu}$. Particularly important is the case where a complex Minkowskian $F^{\mu\nu}$ corresponds to a real $\underline{F}^{\mu\nu}$ in Euclidian space; for this will indicate a tunnelling situation. This is the rationale for seeking solutions to the QCD equations in Euclidian space.

Another point is that in Minkowski space,

$$\tilde{\tilde{F}}_a^{\mu\nu} = -F_a^{\mu\nu},$$

so only the trivial $F = 0$ may be dual,

$$\tilde{F} = \pm F.$$

(If the sign is $(+)$ we say F is self-dual, if $(-)$ anti-dual.) However, in Euclidian space,

$$\tilde{\tilde{\underline{F}}}_a^{\mu\nu} = \underline{F}_a^{\mu\nu},$$

so nontrivial dual values of F may, and indeed do, exist. In addition, Euclidian dual F automatically satisfy the equation of motion.

This last property comes about as follows: the equations of motion for F read

$$D_\mu F_a^{\mu\nu} \equiv \partial_\mu F_a^{\mu\nu} + g \sum f_{abc} B_{b\mu} F_c^{\mu\nu} = 0;$$

the condition

$$D_\mu \tilde{F}_a^{\mu\nu} = 0$$

is the Bianchi identitiy, identically satisfied by any $F = D \times B$ whether or not B solves the equations of motion. However, if $\underline{F}$ is dual, then the Bianchi identity implies $D_\mu F_a^{\mu\nu} = 0$.

The connection with the problem of the vacuum occurs because, in the Euclidian case, the energy-momentum tensor $\Theta^{\mu\nu}$ is replaced by

$$\underline{\Theta}_{\mu\nu} = -\frac{1}{2} \sum_\lambda \{\underline{F}_{\mu\lambda}\underline{F}_{\nu\lambda} - \tilde{\underline{F}}_{\mu\lambda}\tilde{\underline{F}}_{\nu\lambda}\},$$

so for dual fields $\underline{\Theta}_{\mu\nu} = 0$: dual $\underline{F}$ may represent nontrivial vacuum states. Another property of dual fields has to do with a condition of minimum of the Euclidian action $\underline{\mathcal{A}}$. We can write

$$\begin{aligned} \underline{\mathcal{A}} &= \frac{1}{4} \int \mathrm{d}^4\underline{x} \sum \underline{F}_{\mu\nu}\underline{F}_{\mu\nu} \\ &= \frac{1}{4} \sum \int \mathrm{d}^4\underline{x} \{\frac{1}{2}(\underline{F}_{\mu\nu} \pm \tilde{\underline{F}}_{\mu\nu})^2 \mp \underline{F}_{\mu\nu}\tilde{\underline{F}}_{\mu\nu}\} \geq \frac{1}{4} | \int \mathrm{d}^4\underline{x} \sum \underline{F\tilde{F}}|. \end{aligned}$$

Thus the action is positive-definite and reaches its minimum for dual fields, where one has the equality

$$\underline{\mathcal{A}} = \frac{1}{4} | \int \mathrm{d}^4\underline{x} \sum \underline{F\tilde{F}}| = \frac{1}{4} \int \mathrm{d}^4\underline{x} \sum_{a,\mu\nu} (\underline{F}_a^{\mu\nu})^2.$$

Now, and at least in situations where the semi-classical approximation WKB holds, we know that the tunnelling amplitude is given by $\exp(-\underline{\mathcal{A}})$, so the leading tunnelling effect, if it exists, will be provided by dual configurations.

We have been talking about "nontrivial vacuum states". It is not difficult to see that nonzero values of B exist for which $G = 0$. In fact, the general form of such B is what is called a pure gauge, and my be obtained from $B = 0$ by a gauge transformation. Nontrivial solutions of the equations will be such that $G \neq 0$.

Chapter 3

Topological Aspects of Field Theory

- Ryder, Quantum Field Theory
- Srednicki, Quantum Field Theory

3.1 Topological Objects in Field Theory

It turns out that non-linear classical field theories possess extended solutions, commonly known as solitons, which represent stable configurations with a well-defined energy which is nowhere singular. May this be of relevance to particle physics? Since non-Abelian gauge theories are non-linear, it may well be, and the last ten years have seen the discovery of vortices, magnetic monopoles and 'instantons', which are soliton solutions to the gauge-field equations in two space dimensions (i.e. a 'string' in 3-dimensional space), three space dimensions (localised in space but not in time) and 4-dimensional space-time (localised in space and time). If gauge theories are taken seriously then so must these solutions be. It will be seen that they do give rise to new physics and there is even the hope that they may solve the problem of quark confinement.

Not the least interesting feature of this subject is the branch of mathematics which it involves; for the stability of these solitons arises from the fact that the boundary conditions fall into distinct classes, of which the vacuum belongs only to one. These boundary conditions are characterised by a particular correspondence (mapping) between the group space and co-ordinate space, and because these mappings are not continuously deformable into one another they are topologically distinct. The relevant notions in topology will be developed as we go along. We begin our survey with the 'sine-Gordon' equation which has no relevance to particle physics but whose soliton solutions are quite well understood, and therefore form a good introduction to the subject.

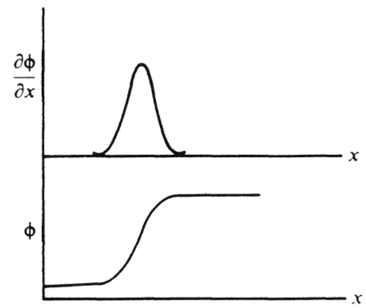

Figure 3.1: A solitary wave (soliton

1. The sine-Gordon kink The sine-Gordon equation

$$\frac{\partial^2\phi}{\partial t^2} - \frac{\partial^2\phi}{\partial x^2} + \frac{1}{b^2}\sin(b\phi) = 0 \tag{3.1}$$

describes a scalar field in one space and one time dimension. It possesses moving, as well as stationary, solutions. To find moving solutions, we want a field of the form

$$\phi(x,t) = f(x - vt) = f(\xi). \tag{3.2}$$

It is easy to check that

$$f(\xi) = \frac{4}{b}\text{atan}\exp\left[\pm\frac{\gamma}{\sqrt{b}}\xi\right] \tag{3.3}$$

is a solution, where $\gamma = (1 - v^2)^{-1/2}$. The appearance of this wave is shown in Fig. 3.1 It is a solitary wave, which moves without changing shape or size, and therefore without dissipation, in strong contrast to the waves set up when, for instance, a stone is thrown into a pond. These waves spread out and the energy is dissipated. Solitary waves (solitons) have been observed, for example, moving along canals. In this case they are solutions of the Korteweg de Vries equation.

Since solitons are solutions of non-linear wave equations the superposition principle is not obeyed. This means that when two solitons meet the resultant wave form is a complicated one, but the surprising thing is that, asymptotically, the solitons separate out again - they 'pass through' one another. This property is, of course, of interest to particle physics, though we shall not develop it any further here. Another consequence of the fact that the superposition principle does not hold is that the quantisation of solitons becomes non-trivial. We shall not follow this matter any further either. Instead, we turn to the stationary solutions of the sine-Gordon equation, which possess an interest of a different type.

It is clear that the sine-Gordon equation possesses an infinite number of constant solutions (which, as we shall see in a moment, have zero energy):

$$\phi = \frac{2\pi n}{b},\ n \in \mathbb{Z}; \tag{3.4}$$

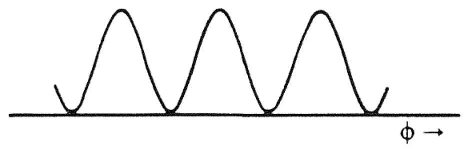

Figure 3.2: The sine-Gordon potential $V(\phi)$.

that is, the sine-Gordon equation possesses a degenerate vacuum ('Vacuum' here does not, of course mean the state in Hilbert space, but simply a classical field configuration of zero energy). The Lagrangian for the sine-Gordon equation is

$$\mathcal{L} = \frac{1}{2}\left(\frac{\partial\phi}{\partial t}\right)^2 - \frac{1}{2}\left(\frac{\partial\phi}{\partial x}\right)^2 - V(\phi) \tag{3.5}$$

with

$$V(\phi) = \frac{1}{b^2}[1 - \cos(b\phi)], \tag{3.6}$$

where the constant has been chosen so that the infinite constant solutions have $V = 0$. They therefore have zero energy since the energy density of the field configuration is

$$\mathcal{H} = \frac{1}{2}\left(\frac{\partial\phi}{\partial t}\right)^2 + \frac{1}{2}\left(\frac{\partial\phi}{\partial x}\right)^2 + V(\phi). \tag{3.7}$$

Note that we may write

$$V(\phi) = \frac{1}{2}\phi^2 - \frac{b^2}{4!}\phi^4 + \ldots, \tag{3.8}$$

or, with $\lambda = b^2$ and unit mass m

$$V(\phi) = \frac{m^2}{2}\phi^2 - \frac{\lambda}{4!}\phi^4 + \ldots, \tag{3.9}$$

and m stands for the 'particle' mass and λ for the self-interaction coupling.

The potential V is shown in Fig. 3.2 with the (zero energy) ground state given by $\phi = \frac{2\pi n}{b}$. Now construct the following configuration. Let ϕ approach one of the zeros of V (say $n = 0$) as $x \to -\infty$, but a different zero (say $n = 1$) as $x \to \infty$. Between these two there is clearly a region where

$$\phi \neq \frac{2\pi n}{b}, \quad \frac{\partial\phi}{\partial x} \neq 0, \tag{3.10}$$

and therefore, from the Hamiltonian $\mathcal{H}$, where there is a positive energy density. We assume the configuration is static, so $\frac{\partial\phi}{\partial t} = 0$. Because of the boundary conditions on ϕ, we expect the total energy to be finite. Let us find what it is. For a stationary solution to the sine-Gordon equation we have

$$\frac{\partial^2 \phi}{\partial x^2} = \frac{\partial V}{\partial \phi} \tag{3.11}$$

which gives on integration

$$\frac{1}{2}\left(\frac{\partial \phi}{\partial x}\right)^2 = V(\phi), \tag{3.12}$$

the integration constant being zero. Then the energy of the stationary soliton is

$$\begin{aligned} E &= \int \mathcal{H} \mathrm{d}x \\ &= \int \left[\frac{1}{2}\left(\frac{\partial \phi}{\partial x}\right)^2 + V(\phi)\right] \mathrm{d}x \\ &= \int 2V(\phi)\mathrm{d}x \\ &= \int_0^{2\pi/b} \sqrt{2V(\phi)}\mathrm{d}\phi \end{aligned} \tag{3.13}$$

where we have put in the integration limits given by $\phi = \frac{2\pi n}{b}$ between $n = 0$ and $n = 1$. This integral is now easily performed. We have

$$\begin{aligned} E &= \frac{\sqrt{2}}{b} \int_0^{2\pi/b} \sqrt{1 - \cos(b\phi)}\mathrm{d}\phi \\ &= \frac{\sqrt{2}}{b^2} \int_0^{2\pi} \sqrt{1 - \cos\alpha}\mathrm{d}\alpha \\ &= \frac{8}{b^2} \\ &= \frac{8m^3}{\lambda} \end{aligned} \tag{3.14}$$

where in the last step we have used the substitution given by the potential. So this soliton has a finite energy, with the interesting property that the energy is inversely proportional to the coupling constant. This may indeed be a useful property for particle physics.

There is a simple model which makes this soliton easy to visualize. consider an infinite horizontal string with pegs attached to it at equally spaced intervals, and connect each peg to its neighbour with a small spring (the 'coupling'). Each peg is also acted on by gravity. The ground state corresponds to every peg hanging vertically. The soliton we have found, with $n = 0 \rightarrow 1$, corresponds to the situation in Fig. 3.3. This soliton - and others of tihs type (see below) - is called a kink. It should be clear from the peg model that the

Figure 3.3: Pegs on a line representing the kink (soliton) solution to the sine-Gordon equation.

kink is stable and cannot decay into the ground state with $E = 0$. This would involve a (semi-)infinite number of pegs turning over, which would need a (semi-) infinite amount of energy. But what is the mathematical reason for the stability of the kink? It is to be found in the boundary conditions. 'Space' in this model is an infinite line, whose boundary is two points (the end-points). At these two points the 1-kink solution has $n = 0$ and $n = 1$, and this is not continuously deformable into $n = 0$ and $n = 0$ (the ground state). The kink, then, is a 'topological' object. Its existence depends on the topological properties of the space (in particular, its boundary, which in this case is a discrete set). This conclusion is a general one; that is to say, the stability of soliton solutions in non-linear field theories is a consequence of topology.

Finally, the stability of this soliton (kink) obviously signals a conservation law; there must be a conserved charge Q, equal to an integer N (the difference between the two integers in $\phi = \frac{2\pi n}{b}$), and a corresponding divergenceless current j^μ ($\mu = 0, 1$). They are easy to construct. With

$$j^\mu = \frac{b}{2\pi}\varepsilon^{\mu\nu}\partial_\nu\phi, \tag{3.15}$$

with the antisymmetric tensor $\varepsilon^{\mu\nu}$ normalized with $\varepsilon^{01} = 1$, we have the identity

$$\partial_\mu j^\mu = 0, \tag{3.16}$$

and the charge is

$$\begin{aligned} Q &= \int_{-\infty}^{\infty} j^0 \mathrm{d}x \\ &= \frac{b}{2\pi}\int_{-\infty}^{\infty}\frac{\partial\phi}{\partial x}\mathrm{d}x \\ &= \frac{b}{2\pi}[\phi(\infty) - \phi(-\infty)] = N. \end{aligned} \tag{3.17}$$

The interesting thing is that the current j^μ does not follow the invariance of $\mathcal{L}$ under any symmetry transformation. It is therefore not a Noether current. Its divergencelessness follows independently of the equations of motion.

We consider, in the following sections, examples of solitons in gauge theories, beginning with one in two space dimensions - the vortex.

2. The Vortex lines Now consider a scalar field in 2-dimensional space. The 'boundary' of this space is the circle at infinity, denoted S^1. We construct a field whose value on the boundary is

$$\phi = ae^{in\theta} \quad (r \to \infty) \tag{3.18}$$

where r and θ are polar coordinates in the plane, a is a constant, and, to make ϕ single-valued, n is an integer. We propose this form, rather than simply $\phi = a$, because it is a generalisation to two dimensions of the solution of the sine-Gordon equation. Taking the gradient, we have

$$\nabla\phi = \frac{1}{r}\left(inae^{in\theta}\right)\hat{\theta}. \tag{3.19}$$

The Lagrangian and Hamiltonian functions are

$$\mathcal{L} = \frac{1}{2}\left(\frac{\partial\phi}{\partial t}\right)^2 - \frac{1}{2}\left|\nabla\phi\right|^2 - V(\phi), \mathcal{H} \qquad = \frac{1}{2}\left(\frac{\partial\phi}{\partial t}\right)^2 + \frac{1}{2}\left|\nabla\phi\right|^2 + V(\phi). \tag{3.20}$$

Now let us consider a static configuration with, for example,

$$V(\phi) = \left[a^2 - \phi^*\phi\right]^2 \tag{3.21}$$

so that $V = 0$ on the boundary. Then as $r \to \infty$

$$\mathcal{H} = \frac{1}{2}\left|\nabla\phi\right|^2 = \frac{n^2a^2}{2r^2} \tag{3.22}$$

and the energy (mass) of the static configuration is

$$E = \int^{\infty} \mathcal{H} r \mathrm{d}r \mathrm{d}\theta = \pi n^2 a^2 \int^{\infty} \frac{1}{r}\mathrm{d}r. \tag{3.23}$$

This is logarithmically divergent; the kink, as it stands, cannot be generalised to two dimensions - nor to more than two, for it turns out that in all these cases the energy is divergent.

Our next goal is to change the theory in order to have such a configuration by introducing gauge bosons to cancel the bad behaviour of the scalars at infinity.

3.2 The $U(1)$ Higgs Model in $2+1$ Dimensions

To be a little more systematic, let us start from the Higgs Lagrangian:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + |(\partial_\mu + ieA_\mu)\phi|^2 - m^2\||\phi|^2 - \lambda|\phi|^4. \tag{3.24}$$

Spontaneous symmetry breaking is signalled by $m^2 < 0$, and the vacuum is then given by

$$|\phi|_{vac} = \sqrt{\frac{-m^2}{2\lambda}}. \tag{3.25}$$

The equations of motion are

$$D^\mu(D_\mu\phi) = -m^2\phi - 2\lambda\phi|\phi|^2 \tag{3.26}$$
$$\partial^\nu F_{\mu\nu} = ie(\phi\partial_\mu\phi^* - \phi^*\partial_\mu\phi) + 2e^2 A_\mu|\phi|^2 \tag{3.27}$$

Once again we have degenerate vacuum configurations with

$$|\phi| = v \;\Rightarrow \phi = ve^{i\theta}. \tag{3.28}$$

There is an extended field configuration with a finite energy which is not trivial by choosing A_μ of the form

$$\phi(r) \to ve^{i\theta} \tag{3.29}$$
$$A_\mu(r) = \frac{1}{e}\nabla\theta \quad (r\to\infty) \tag{3.30}$$

with the components for $r \to \infty$

$$A_r \to 0 \tag{3.31}$$
$$A_\theta \to -\frac{1}{er}. \tag{3.32}$$

For a static configuration $\mathcal{H} = -\mathcal{L}$ we have $\mathcal{H} \to 0$ as $r \to \infty$, making possible a field configuration of finite energy. We shall now see that the effect of adding the gauge field is to give the soliton magnetic flux. Consider the integral $\oint \vec{A}\cdot \mathrm{d}\vec{l}$ round the circle S^1 at infinity. By Stoke's theorem, this is $\int \vec{B}\cdot\mathrm{d}\vec{S} = \Phi$, the flux enclosed, hence

$$\Phi = \oint \vec{A}\cdot\mathrm{d}\vec{l} = \oint A_\theta r\mathrm{d}\theta = -\frac{2\pi}{e}, \tag{3.33}$$

and the flux is quantised. So we have, after all, constructed a 2-dimensional field configuration, consisting of a charged scalar field and a gauge field (the electromagnetic field!). It carries magnetic flux, and since $D_\mu\phi \to 0$ and $F_{\mu\nu} \to 0$ on the boundary at infinity, it appears to have finite energy. It is clear that by adding a third dimension (the z axis) on which the fields have no dependence, this configuration becomes a vortex line. Apart from the presence of the scalar field, it is the same as the solenoid discussed under the Bohm-Aharonov effect; and just as that effect was attributable to the topology of the gauge group $U(1)$, so here also we shall se that this is this same topology which ensures stability of the vortex.

Why are these solutions stable? As with the kink, the reason is topological. The Lagrangian is invariant under a symmetry group - in this case $U(1)$, the electromagnetic gauge group. The field ϕ (with boundary value given by parametrization $\phi = re^{i\theta}$ with $r \to \infty$) is a representation of $U(1)$. The group space of $U(1)$ is a circle S^1, since an element of $U(1)$ may be written $\exp(i\theta) = \exp[i(\theta + 2\pi)]$, so the space of all values of θ is a line with $\theta = 0$ identified with $\theta = 2\pi$, and the line becomes a circle S^1. The field ϕ is a representation of basis of $U(1)$, but it is the boundary value of the field in a 2-dimensional space. This boundary is clearly a circle S^1 (the circle $r \to \infty, \theta = (0 \to 2\pi)$). Hence *phi* defines a mapping of the boundary S^1 in physical space onto the group space S^1:

$$\phi : S^1 \to S^1, \tag{3.34}$$

the mapping being specified by the integer n. Now a solution characterised by one value of n is stable since it cannot be continuously deformed into a solution with a different value of n (a rubber band which fits twice round a circle cannot be continuously deformed into one which goes once round the circle). This is to say that the first homotopy group of S^1, the group space of $U(1)$, is not trivial:

$$\pi_1(S^1) = \mathbb{Z}. \tag{3.35}$$

$\mathbb{Z}$ is the additive group of integers.

The status of a topological argument like this is that it provides a very general condition which must be fulfilled in order that solitons exist in a particular model. If, as in the model above, the topological argument indicates that soliton solutions are possible in principle then one goes to the equations of motion to find them. Topology therefore provides existence arguments.

3.3 The Dirac Monopole

Consider a magnetic monopole of strength g at the origin. The magnetic field is radial and is given by a Coulomb-type law

$$\mathbf{B} = \frac{g}{r^3}\mathbf{r} = -g\nabla(\frac{1}{r}) \tag{3.36}$$

(we are using Gaussian units). Since $\nabla^2(\frac{1}{r}) = -4\pi\delta^{(3)}(r)$, we have

$$\nabla \cdot \mathbf{B} = 4\pi g\delta^{(3)}(r) \tag{3.37}$$

corresponding to a point magnetic charge, as desired. Since $\mathbf{B}$ is radial, the total flux through a sphere surrounding the origin is

$$\Phi = 4\pi r^2 B = 4\pi g. \tag{3.38}$$

Consider a particle with electric charge e in the field of this monopole. The wave function for a free particle is

$$\psi = |\psi| \exp\left[\frac{i}{\hbar}(\mathbf{p} \cdot \mathbf{r} - Et)\right]. \tag{3.39}$$

In the presence of an electromagnetic field, $\mathbf{p} \rightarrow \mathbf{p} - \frac{e}{c}\mathbf{A}$, so

$$\psi \rightarrow \psi \exp\left(-\frac{ie}{\hbar c}\mathbf{A}\cdot\mathbf{r}\right); \tag{3.40}$$

or the phase α changes by

$$\alpha \rightarrow \alpha - \frac{e}{\hbar c}\mathbf{A}\cdot\mathbf{r}. \tag{3.41}$$

Consider a closed path at fixed r, θ, with ϕ ranging from 0 to 2π. The total change in phase is

$$\begin{aligned}
\Delta\alpha &= \frac{e}{\hbar c}\oint \mathbf{A}\cdot \mathrm{d}\mathbf{L} &(3.42)\\
&= \frac{e}{\hbar c}\int \mathrm{curl}\mathbf{A}\cdot \mathrm{d}\mathbf{S} &(3.43)\\
&= \frac{e}{\hbar c}\int \mathbf{B}\cdot \mathrm{d}\mathbf{S} &(3.44)\\
&= \frac{e}{\hbar c}\Phi(r,\theta); &(3.45)
\end{aligned}$$

$\Phi(r, \theta)$ is the flux through the cap defined by a particular r and θ, as shown by the shaded area in the figure. As θ is varied the flux through the cap varies. As $\theta \rightarrow 0$ the loop shrinks to a point and the flux passing through the cap approaches zero:

$$\Phi(r, 0) = 0.$$

As the loop is lowered over the sphere the cap encloses more and more flux until, eventually, at $\theta \rightarrow \pi$ we should have

$$\Phi(r, \pi) = 4\pi g. \tag{3.46}$$

However, as $\theta \rightarrow \pi$ the loop as again shrunk to a point so the requirement that $\Phi(r, \pi)$ is finite entails that A is singular at $\theta = \pi$. This argument holds for all spheres of all possible radii, so it follows that $\mathbf{A}$ is singular along the entire negative z axis. This is known as the Dirac string. It is clear that by a suitable choice of coordinates the string may be chosen to be along any direction, and, in fact, need not be straight, but must be continuous.

The singularity in $\mathbf{A}$ gives rise to the so-called Dirac veto - that the wave function vanish along hte negative z axis. Its phase is therefore indeterminate there and there is no necessity that $\theta \rightarrow \pi, \Delta\alpha \rightarrow 0$. We must have $\Delta\alpha = 2\pi n$, however, in order for ψ to be single-valued. We then have

$$\begin{aligned}
2\pi n &= \frac{e}{\hbar c}4\pi g, &(3.47)\\
eg &= \frac{1}{2}n\hbar c. &(3.48)
\end{aligned}$$

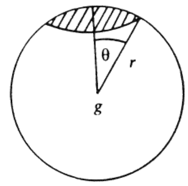

Figure 3.4:

This is the Dirac quantisation condition. It implies that the product of any electric with any magnetic charge is given by the above. Then, in principle, if there exists a magnetic charge anywhere in the universe all electric charges will be quantised:

$$e = n\frac{\hbar c}{2g}.$$

This is a possible explanation for the observed quantisation of electric charge, though nowadays this is more commonly ascribed to the existence of quarks and non-Abelian symmetry groups. Note, however, that the quantisation condition has an explicit dependence on Planck's constant, and therefore on the quantum theory. In natural units, $\hbar = c = 1$, it becomes

$$eg = \frac{1}{2}n.$$

Let us now derive an expression for the vector potential A_μ. As seen above, it is singular. This much is clear, for if $\mathbf{B} = \text{curl}\mathbf{A}$ and $\mathbf{A}$ is regular $\text{div}\mathbf{B} = 0$, and no magnetic charges may exist. From the argument above, A is constructed by considering the pole as the end-point of a string of magnetic dipoles whose other end is at infinity. This gives

$$A_x = g\frac{-y}{r(r+z)},\ \ A_y = g\frac{x}{r(r+z)},\ \ A_z = 0 \tag{3.49}$$

or

$$A_r = A_\theta = 0,\ \ A_\phi = -\frac{g}{r}\frac{1+\cos\theta}{\sin\theta}. \tag{3.50}$$

$\mathbf{A}$ is clearly singular along $r = -z$. If, on the other hand, the Dirac string were chosen to be along $r = z$, we should have

$$A_r = A_\theta = 0,\ \ A_\phi = -\frac{g}{r}\frac{1+\cos\theta}{\sin\theta}. \tag{3.51}$$

The rationale for writing the alternative expression is that the Dirac string singularity is clearly unphysical, and in these expressions it is in different places. The only physical singularity in $\mathbf{A}$ is at the origin, where $\text{div}\mathbf{B} = \text{divcurl}\mathbf{A}$ is singular. Since it is obviously desirable to get rid of unphysical singularities, this suggests the following construction.

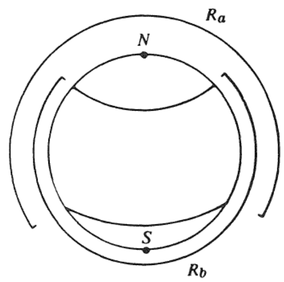

Figure 3.5: R_a and R_b are overlapping domains on the sphere. R_a excludes the S pole, R_b the N pole.

Divide the space surrounding the monopole, - the sphere, essentially - into two overlapping regions R_a and R_b as shown in the figure. R_a excludes the negative z axis (S pole) and R_b excluded the positive z axis (N pole). In each region $\mathbf{A}$ is defined differently:

$$A^a_r = A^a_\theta = 0,\ A^a_\phi = \frac{g}{r}\frac{1-\cos\theta}{\sin\theta}, \tag{3.52}$$

$$A^b_r = A^b_\theta = 0,\ A^b_\phi = -\frac{g}{r}\frac{1+\cos\theta}{\sin\theta}. \tag{3.53}$$

It is clear that $\mathbf{A}^a$ and $\mathbf{A}^b$ are both finite in their own domain. In the region of overlap, however, they are not the same, but are related by a gauge transformation ($\hbar = c = 1$):

$$A^b_\phi = A^a_\phi - \frac{2g}{r\sin\theta} = A^a_\phi - \frac{i}{e}S\nabla_\phi S^{-1} \tag{3.54}$$

with

$$S = \exp(2ige\phi). \tag{3.55}$$

The covariant form is

$$A^b_\mu = A^a_\mu - \frac{i}{e}S\partial_\mu S^{-1}. \tag{3.56}$$

The requirement that the gauge transform function S be single-valued as $\phi \to \phi + 2\pi$ is clearly the Dirac quantisation condition. To check that they really do represent a monopole, we calculate the total magnetic flux through a sphere surrounding the origin.

$$\begin{aligned}\Phi &= \int F_{\mu\nu}\mathrm{d}x^{\mu\nu} \\ &= \oint \mathrm{curl}\mathbf{A}\cdot\mathrm{d}\mathbf{S} \\ &= \int_{R_a}\mathrm{curl}\mathbf{A}\cdot\mathrm{d}\mathbf{S} + \int_{R_b}\mathrm{curl}\mathbf{A}\cdot\mathrm{d}\mathbf{S}.\end{aligned}$$

Here we take R_a and R_b as not actually overlapping, but having a common boundary, which for convenience is taken to be the equator $\theta = \pi/2$. Since R_a, R_b have boundaries Stokes' theorem is applicable, and since the equator bounds R_a in a positive orientation and R_b in a negative one we have

$$\Phi = \oint_{\theta=\pi/2} \mathbf{A}^a \cdot \mathrm{d}\mathbf{l}^a - \oint_{\theta=\pi.2} \mathbf{A}^b \cdot \mathrm{d}\mathbf{l}^b \tag{3.57}$$

$$= \frac{i}{e} \oint \frac{\mathrm{d}}{\mathrm{d}\phi} (\log S^{-1}) \mathrm{d}\phi \tag{3.58}$$

$$= 4\pi g. \tag{3.59}$$

This construction is due to Wu and Yang, and is, in essence, a fibre bundle formulation of the magnetic monopole. The base space (3-dimensional space $\mathbb{R}^3$ minus the origin $\sim S^2 \times \mathbb{R}$) is parameterised in two independent ways, corresponding to two overlapping but not identical regions. In each region the vector potential is given by a different expression. Readers familiar with the Moebius strip will recognize a similarity here. there is no unique parameterisation of the Moebius strip; locally it is the direct product of an interval $(0, 1)$ and a circle, but globally the circle has to be divided into two distinct overlapping regions, with a different parameterisationof the strip in each region.

There is thus a fibre-bundle formulation of the Dirac monopole. The base space is essentially S^2 (the sphere surrounding the monopole) and the group space is S^1 (since the gauge group is $U(1)$). The fibre bundle is not $S^2 \times S^1$ but S^3, whih is locally the same as $S^2 \times S^1$ but is globally distinct.

3.4 The 't-Hooft-Polyakov Monopole

The previous discussion of magnetic monopoles, although interesting, was not compelling, because ordinary electrodynamics does not require that monopoles exist. Electrodynamics withouth monopoles is perfectly consistent. However, in certain gauge theories, we will find that spontaneous symmetry breaking is intimately connected with the existence of monopole solutions. Hence, monopoles must exist for these theories as a consequence of broken gauge symmetry.

It can be shown that pure gauge theory does not, by itself, possess any static non-singular monopole configurations. However, a more general case, such as a gauge theory coupled to scalar fields, does possess monopole solutions.

We now want to look for a finite-energy solution of the classical field equations with nonzero winding number, but we already know that these will not exist unless we introduce gauge fields. We therefore take ϕ^a to be in the adjoint representation of an $SU(2)$ gauge group (we can also say that it is in the fundamental representation of $O(3)$). The

Lagrangian is given by

$$\mathcal{L} = \frac{1}{2}(D_\mu\phi)^a(D^\mu\phi)^a - \frac{1}{4}F^a_{\mu\nu}F^\alpha_{\mu\nu} - V(\phi),\ a = 1,2,3, \tag{3.60}$$

$$(D_\mu\phi)^a = \partial_\mu\phi^\alpha + e\varepsilon^{abc}A^b_\mu\phi^c, \tag{3.61}$$

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + e\varepsilon^{abc}A^b_\mu A^c_\nu, \tag{3.62}$$

with potential $V(\phi)$

$$V(\phi = \frac{1}{8}\lambda(\phi^a\phi^a - v^2)^2. \tag{3.63}$$

The potential has its minimum at $|\phi| = v$, thus the gauge symmetry is spontaneously broken to $U(1)$. The system will choose on of the possible minima we denote by for $\langle\phi^a\rangle = {}^t(0,0,v)$

$$\langle\phi^a\rangle = v\delta^{a3}. \tag{3.64}$$

We can now expand around the vacuum,

$$\phi^a = \begin{pmatrix} x^1 \\ x^2 \\ x^3 + v \end{pmatrix} = \langle\phi^a\rangle + x^a. \tag{3.65}$$

and find that the A^3_μ field remains massless (interpreted as electromagnetic field). Taking a linear combination of the other vactor fields, we can define the complex vector fields

$$W^\pm_\mu = \frac{A^a_\mu \pm iA^2_\mu}{\sqrt{2}} \tag{3.66}$$

with mass $m_W = ev$ and electrical charge $\pm e$. This theory, known as Georgi-Glashow model, was once considered as an alternative to the Standard Model of electroweak interactions, but it has been ruled out due to a missing Z^0 boson.

We discuss this model because the mass is proportional to the charge as well as to the vacuum expectation value, $m_W = ev$. In other words, the mass depends on the coupling constant. If one had a way to measure the vacuum expectation value, the mass would be completely determined by the coupling. In this model we would have to measure the Higgs boson x to measure the vacuum expectation value. In the Standard Model of electro-weak interaction (Glashow-Salam-Weinberg model), however, we can calculate the vacuum expectation value without knowing the Higgs mass.

Returning to topological phenomena, we try to use the massless property of the A^3_μ field to construct a field strength tensor. We can write down a gauge-invariant expression that reduces to the electromagnetic field strength tensor $F_{\mu\nu} = \partial_\mu A^3_\nu - \partial_\nu A^3_\mu$ when setting $\phi^a = v\delta^{a3}$:

$$F_{\mu\nu} = \hat{\phi}F^a_{\mu\nu} - \frac{1}{e}\varepsilon^{abc}(D_\mu\hat{\phi})^b(D_\nu\hat{\phi})^c, \tag{3.67}$$

with

$$\hat{\phi}^a = \frac{\phi^a}{|\phi|} \tag{3.68}$$

. In this procedure we have only used the first term and ignored the second term. It is properly derived in Voloshin (1982). We can now use this as the definition of the electromagnetic field strength at any spacetime point where $|\phi| \neq 0$. (If $|\phi| = 0$, the $SU(2)$ symmetry is unbroken, and there is no gauge-invariant way to pick out a particular component of the non-Abelian field strength $F^a_{\mu\nu}$.) Using repeatedly $\hat{\phi}^a\hat{\phi}^a = 1$, we can rewrite our electromagnetic field strength tensor as

$$F_{\mu\nu} = \partial_\mu(\hat{\phi}^a A^a_\nu) - \partial_\nu(\hat{\phi}^a A^a_\mu) - \frac{1}{e}\varepsilon^{abc}\hat{\phi}^a\partial_\mu\hat{\phi}^b\partial_\nu\hat{\phi}^c. \tag{3.69}$$

This allows us to write the magnetic field components easily as

$$B^i = \frac{1}{2}\varepsilon^{ijk}F_{jk} \tag{3.70}$$

$$= \varepsilon^{ijk}\partial_j(\hat{\phi}^a A^a_k) - \frac{1}{2e}\varepsilon^{ijk}\varepsilon^{abc}\hat{\phi}^a\partial_j\hat{\phi}^b\partial_k\hat{\phi}^c. \tag{3.71}$$

Let us consider the magnetic flux through a sphere at spatial infinity; this is given by

$$\Phi = \int \mathbf{B}\cdot \mathrm{d}\mathbf{S} \tag{3.72}$$

$$= \int \mathrm{d}S_i\varepsilon^{ijk}\partial_j(\hat{\phi}^a A^a_k) - \frac{1}{2e}\varepsilon^{ijk}\varepsilon^{abc}\hat{\phi}^a\partial_j\hat{\phi}^b\partial_k\hat{\phi}^c \tag{3.73}$$

$$= \int \mathrm{d}S_i\frac{1}{2e}\varepsilon^{ijk}\varepsilon^{abc}\hat{\phi}^a\partial_j\hat{\phi}^b\partial_k\hat{\phi}^c, \tag{3.74}$$

where the first term is a curl and has thus zero divergence. The second term, however, yields with the winding number n

$$\Phi = -\frac{4\pi n}{e}. \tag{3.75}$$

This flux implies that any soliton with nonzero winding number is a magnetic monopole with magnetic charge $Q_M = \Phi$.

If we add a field in the fundamental representation of $SU(2)$, then the component fields have electric charges $\pm\frac{1}{2}e$. This is the smallest electric charge we can get, and all possible electric charges are integer multiples of it. All magnetic charges are integer multiples of $4\pi/e$. Thus the possible electric and magnetic charges obey the Dirac charge quantization condition, which is

$$Q_E Q_M = 2\pi k, \tag{3.76}$$

where k is an integer. This condition can be derived from general considerations of the quantum properties of monopoles.

Now let us turn to the explicit construction of a soliton solution. This simplest case to consider is provided by the identity map (which has winding number $n = 1$); the soliton we will find is the 't-Hooft-Polyakov monopole. We will follow the procedure from our earlier discussions on vortex solutions, where we had a scalar field getting its vacuum expectation value at infinity and a magnetic field falling on the boundary. In other words, we need a gauge field A which is a total derivative.

The boundary condition on the scalar field is

$$\lim_{r\to\infty} \phi^a(\mathbf{x}) = v\frac{x^a}{r}. \tag{3.77}$$

We can find the appropriate boundary condition on the gauge field by requiring $(D_\mu\phi)^a = 0$ for large r which yields

$$\partial_i \left(\frac{x^a}{r}\right) + e\varepsilon^{abc} A_i^b \frac{x^c}{r} = 0. \tag{3.78}$$

Solving this equation by writing out the first derivative explicitly and multiplying by $rx_j\varepsilon^{jda}$, we will find

$$A_i^d = \varepsilon^{dij}\frac{x_j}{er^2}. \tag{3.79}$$

We finally find the static solution given by

$$\phi^a(\mathbf{x}) = vf(r)\frac{x^a}{r}, \tag{3.80}$$

$$\mathcal{A}_i^a(\mathbf{x}) = \frac{1}{e}a(r)\varepsilon^{aiy}\frac{x_j}{r^2}, \tag{3.81}$$

with $f(\infty) = a(\infty) = 1$ (so that A_i^a and ϕ^a have the desired asymptotic limits) and $f(0) = a(0) = 0$ (so that A_i^a and ϕ^a are well defined at $r = 0$).

The total energy of the soliton (which we call M because it is the mass of the monopole) is given by

$$M = \int \mathrm{d}^3\mathbf{x}\ \mathcal{H} \tag{3.82}$$

$$= \int \mathrm{d}^3\mathbf{x}\ \left[\frac{1}{2}B_i^a B_i^a + \frac{1}{2}(D_i\phi)^a(D_i\phi)^a + V(\phi)\right] \tag{3.83}$$

$$= \int \mathrm{d}^3\mathbf{x}\ \left[\frac{1}{2}\left(B_i^a + (D_i\phi)^2\right)^2 - B_i^a(D_i\phi)^a - V(\phi)\right] \tag{3.84}$$

where we completed the square in the last line. Using the distribution rule for covariant derivatives,

$$B_i^a(D_i\phi)^a = \partial_i(B_i^a\phi^a) - (D_iB_i)^2\phi^2 \tag{3.85}$$

and the implication of the Bianchi identity of

$$(D_i B_i)^a = 0, \tag{3.86}$$

we can write

$$M = \int \mathrm{d}^3\mathbf{x} \left[\frac{1}{2}\left(B_i^a + (D_i\phi)^a\right)^2 + V(\phi)\right] - \int \mathrm{d}^3\mathbf{x}\partial_i(B_i^a\phi^a) \tag{3.87}$$

and observe that the squared bracket, $\left[\frac{1}{2}\left(B_i^a + (D_i\phi)^a\right)^2 + V(\phi)\right]$, is positive.

This means that the mass is constrained by

$$M \geq -\int \mathrm{d}^3\mathbf{x}\partial_i(B_i^a\phi^a) = -\int_{S^1} \mathrm{d}S_i\ B_i^a\phi^a \tag{3.88}$$

where S^1 denotes the sphere at infinity and we used Stoke's theorem. At space infinity, however, we have

$$\phi^a = v\hat{x}^a \tag{3.89}$$

and

$$B_i^a\phi^a = vB_i, \tag{3.90}$$

where vB_i is the generalized magnetic field form $F_{\mu\nu}$ in non-Abelian gauge theory. More loosely writing,

$$M \geq -v \cdot \text{flux} = -v\int_{S^1} \mathrm{d}\mathbf{S}\cdot\mathbf{B}, \tag{3.91}$$

and see that the lowest value is related to the flux times the vacuum expactation value. With fixed **B** from the boundary conditions on the sphere, we find find the flux

$$\int \mathrm{d}\mathbf{S}\cdot\mathbf{B} = -\frac{4\pi}{e} \tag{3.92}$$

and the limit for the magnetic monopole,

$$M \geq \frac{4\pi v}{e} = \frac{4\pi(ev)}{e^2} = \frac{m_W}{\alpha}, \tag{3.93}$$

using $m_W = ev, \alpha = \frac{e^2}{4\pi}$. Since $\alpha \ll 1$, the monopole is much heavier than the W boson.

Alas, the Georgi-Glashow model, which has monopole solutions, is not in accord with nature, while the Standard Model, which is in accord with nature, does not have monopole solutions. This is because the Standard Model, electric charge is a linear combination of an $SU(2)$ generator and the $U(1)$ hypercharge generator. Nothing prevents us from introducing an $SU(2)$ singlet field with an arbitrarily small hypercharge. Such a field would have an arbitrary small electric charge (in units of e), and then the Dirac charge quantization condition would preclude the existence of magnetic monopoles.

This disappointing situation is remedied in unified theories, where the gauge group has a single non-Abelian factor like $SU(5)$. In unified theories, the monopole mass is of order M_x/α, where m_X is the mass of a superheavy vector boson; typically $m_X \sim 10^{15}$ GeV.

In other words, the physics here is that we have finite static energy solutions which correspond to massive objects. We have a radial magnetic field, a magnetic monopole. The mass of the magnetic monopole is related to the mass of the gauge boson that came out of the SSB and the coupling constant. We saw that with current accelerators it is impossible to have enough energy to produce it. However, in the early universe enough energy was present, so magnetic monopoles would possibly exist given a correct gauge group. In a phase transition where spontaneous symmetry breaking occurs magnetic monopoles can be created with the correct boundary conditions.

In a cooling universe, it is possible to create an extended field configuration with a different vacuum expectation value for ϕ for different space regions. The second ingredient is having a correct gauge group with possibilities for degenerate vacua. Here we discussed $O(3) \cong SU(2) \to U(1)$ breaking, another possible scenario would be $SU(5) \to SU(3) \times SU(2) \times U(1)$.

Thus, if you embed the Standard Model in a greater theory unifying all forces, it is possible to have magnetic monopoles.

3.5 Phase Transitions in the Early Universe

Elementary particle theory possesses gauge symmetries that are spontaneously broken by scalar fields belonging to non-trivial representations of the gauge group when these fields develop non-zero expectation values at the minimum of the effective potential. In particular, the $SU(2)_L \times U(1)_Y$ gauge group of electroweak theory is spontaneously broken to the $U(1)_{em}$ of electromagnetism by the electroweak Higgs scalar expectation value. If grand unification to a group larger than the $SU(3)_C \times SU(2)_L \times U(1)_Y$ of the Standard Model, e.g. to $SU(5)$, occurs at some energy scale, then the grand unified gauge group breaks spontaneously to the Standard Model gauge group before this gauge group in turn breaks to the $U(1)_{em}$ gauge group. Things may be more complicated than this, with a sequence of spontaneous symmetry breakings to subgroups of the original grand unified group.

As we shall see later, finite temperature effects may result in some other minimum of the effective potential being deeoper than the absolute minimum of the zero-temperature theory. Then, as the universe cools, it may undergo a series of first- and second-order phase transitions between different minima of the effective potential. Such transitions will occur at temperatures corresponding to the scales of energy associated with the various spontaneous symmetry breakings. In the case of electroweak phase transition to the phase with only $U(1)_{em}$ unbroken, the scale of temperature for the phase transition will be or order $10^2 - 10^3$ GeV and, in the case of a grand unified phase transition to a phase with $SU(3)_C \times SU(2)_L \times U(1)_Y$ in stages through a sequence of phase transitions, the additional phase transitions will occur at intermediate scales.

It should be noted that the phase transitions can have profound effect on the history of the universe through a number of different processesses. For example, topologically stable objects such as domain walls, cosmic strings and magnetic monopoles can be formed when the 'alignment' of the spontaneous symmetry breaking expectation value is different in adjacent causal domains. These can make substational contributions to the energy density of the universe. Moreoever, if supercooling occurs before the phase transition is completed, the reheating that takes place when the phase transition occurs can greatly modify pre-existing particle densities. In addition, if the universe spends some time with positive vacuum energy (cosmological constant) before relaxing to a minimum with zero vacuum energy, then rapid expandion can occur. Such an 'inflationary' stage in the history of the universe may explain the extreme isotropy, homogeneity and flatness of the present day observed universe. For all of these reasons it is important to understand any phase transitions that may have occured as the universe cooled.

After a short review of the partition function of quantum statistical mechanics and the effective potential, we calculate the effective potential with nonzero temperature and discuss different phase transition phenomena.

Summary of Thermal Field Theory In quantum statistical mechanics, the partition function is the trace,

$$Z = \mathrm{tr} e^{-\beta H} = \sum_n \langle n | e^{-\beta \hat{H}} | n \rangle \tag{3.94}$$

over a complete set of states $|n\rangle$. It is basically the trace multiplied by a probability factor given by the exponential. Because we want to have an equilibrium we have to start from $|n\rangle$ and return back to it, $\langle n|$. After you make a small transition in temperature, you see that it corresponds just to time. So, if you really want to see how it evolves in time, where we do not have $\exp -it$ as in quantum mechanics, but $\beta \equiv \frac{L}{k_B T} = \frac{1}{T}$, we see the probability of the state return back to it.

$$Z = \int \mathrm{d}^3\mathbf{x} \mathrm{d}^3 y \ \sum_n \langle n|x\rangle \langle x| e^{-\beta \hat{H}} |y\rangle \langle y|n\rangle, \tag{3.95}$$

where we have inserted twice the unity, $1 = |y\rangle\langle y|$, to get with the delta function $\langle x|y\rangle = \delta(x-y)$

$$= \int \mathrm{d}^3 x \mathrm{d}^3 y \ \langle x|y\rangle \langle x| e^{-\beta \hat{H}} |y\rangle \tag{3.96}$$

$$\rightsquigarrow Z = \int \mathrm{d}^3x \langle x | e^{-\beta \hat{H}} | x \rangle. \tag{3.97}$$

We know something nice about those transitions: we can formulate the pathintegral. Let's introduce time:

$$Z = \int^3 \mathrm{d}^3\mathbf{x} \langle x, t = -i\beta | x, t = 0 \rangle \tag{3.98}$$

So, the partition function, is nothing else than the integral over all transition amplitudes of starting from a point back to the same point, after an imaginary time $-i\beta$ has elapsed (Wick-rotated).

Only two things are different from path integral: first, we have evolution to imaginary time, the second is that it returns to the same point.

In the path integral formulation we have the same partition function written as

$$Z = \int_{\text{periodic boundary conditions}} \mathcal{D}q(z)\ e^{-\int_0^\beta \mathrm{d}\tau L(q)}, \tag{3.99}$$

with $\tau = -i\beta$. If you are surprised by this, remember that we have to Wick-rotate in field theory as well. In fact, if you go to infinity, you go exactly back to field theory.

However, here we have periodic boundary conditions, because we start from x and want to return to the exact same point x, $q(0) = q(\beta)$.

In statistical quantum field theory we have the partition function

$$Z = \mathrm{tr} e^{-\beta \hat{H}} = \int_{PBC} \mathcal{D}\phi e^{-\int_0^\beta \mathrm{d}\tau \int \mathrm{d}^3\mathbf{x} \mathcal{L}[\phi]} \tag{3.100}$$

where we have to take the trace over all states and Hamiltonian (which is dependent on the field). Otherwise, nothing else has changed.

The periodic boundary conditions (PBC) are given by $\phi(\mathbf{x}, \tau = 0) = \phi(\mathbf{x}, \tau = \beta)$. So, statistical field theory is nothing more than restricting your time interval.

In the zero T limit, $\beta \to \infty$, we recover the standard Wick-rotated field theory. With the partition function we can compute many things: temperature, entropy, etc.

Here, however, we want to compute something else, the vacuum expectation value. How do you compute the vacuum expectation value in field theory with given Z. We take the functional derivative in respect to sources, in other words, from the effective potential. So nothing changes if we want to compute the vacuum expectation value.

Effective Potential At Zero Temperature Given a scalar field theory defined by

$$Z = e^{iW[J]} = \int \mathcal{D}\phi e^{i[S(\phi) + J\phi]}, \tag{3.101}$$

with $J\phi = \int \mathrm{d}^4x J(x)\phi(x)$. Then the vacuum expectation value is given by

$$\langle 0\rangle_J = \frac{\int \mathcal{D}\phi\ \phi e^{i[S+J\phi]}}{\int \mathcal{D}\phi\ e^{i[S+J\phi]}}. \tag{3.102}$$

For a physical theory we would later put $J = 0$. In the presence of sources we have

$$\langle\phi(x)\rangle_J = \langle 0|\phi|0\rangle_J = \frac{\delta W}{\delta J(x)}. \tag{3.103}$$

This means that we have reduced the task of computing the vacuum expectation value to computing the path integral, which, unfortunately, is rather hard.

Given a functional W of J we can perform a Legendre transform to obtain a functional Γ of $\langle\phi\rangle$. Legendre transform is just the fancy term for the simple relation

$$\Gamma[\langle\phi\rangle] = W[J] - \int \mathrm{d}^4x J(x)\langle\phi(x)\rangle \tag{3.104}$$

which implies the simple equation

$$\frac{\delta\Gamma[\langle\phi\rangle]}{\delta\langle\phi\rangle(x)} = -J(x) \tag{3.105}$$

and for $J = 0$,

$$\frac{\delta\Gamma[\langle\phi\rangle]}{\delta\langle\phi\rangle(x)} = 0. \tag{3.106}$$

Expanding the functional $\Gamma[\langle\phi\rangle]$ gives

$$\Gamma[\langle\phi\rangle] = \int \mathrm{d}^4x[-V_{eff}(\langle\phi\rangle) + Z(\langle\phi\rangle)(\partial\langle\phi\rangle)^2 + \ldots] \tag{3.107}$$

which yields

$$V'_{eff}(\langle\phi\rangle) = J \tag{3.108}$$

and without external sources

$$V'_{eff}(\langle\phi\rangle) = 0. \tag{3.109}$$

In other words, the vacuum expectation value of $\langle\phi\rangle$ in the absence of an external source is determined by minimizing $V_{eff}(\langle\phi\rangle)$.

All of these formal manipulations are not worth much if we cannot evaluate $W[J]$. In fact, in most cases we can only evaluate $\exp iW[J] = \int \mathcal{D}\phi \exp i[S[\langle\phi\rangle] + J\langle\phi\rangle]$ in the steepest descent approximation, namely the solution of

$$\frac{\delta[S(\phi) + \int \mathrm{d}^4y J(y)\phi(y)]}{\delta\phi(x)}|_{\phi_s} = 0 \tag{3.110}$$

or more explicitly,

$$\partial^2\phi_s(x) + V'[\phi_x(x)] = J(x). \tag{3.111}$$

We write the dummy integration variable as $\phi = \phi_s + \tilde{\phi}$ and expand to quadratic order in $\tilde{\phi}$ to obtain

$$Z = \exp\frac{i}{\hbar}W[J] = \int \mathcal{D}\phi \exp\frac{i}{\hbar}[S[\phi] + J\phi] \tag{3.112}$$

$$\approx e^{\frac{i}{\hbar}[S[[phi_s]+J\phi_s]]} \int \mathcal{D}\tilde{\phi} e^{\frac{i}{\hbar}\int \mathrm{d}^4x \frac{1}{2}[(\partial\tilde{\phi})^2 - V''(\phi_s)\tilde{\phi}^2]} \tag{3.113}$$

$$= \exp\left(\frac{i}{\hbar}[S[\phi_s] + J\phi_s] - \frac{1}{2}\mathrm{tr}\ \log[\partial^2 + V''(\phi_s)]\right). \tag{3.114}$$

Now we have determined

$$W[J] = [S[\phi_s] + J\phi_s] + \frac{i\hbar}{2}\mathrm{tr}\ \log[\partial^2 + V''(\phi_s)] + \mathcal{O}(\hbar^2), \tag{3.115}$$

and need to calculate:

$$\langle\phi\rangle = \frac{\delta W}{\delta J} = \frac{\delta[S[\phi_s] + J\phi_s]}{\delta\phi_s}\frac{\delta\phi_s}{\delta J} + \phi_s + \mathcal{O}(\hbar) = \phi_s + \mathcal{O}(\hbar). \tag{3.116}$$

To leading order in $\hbar$, $\langle\phi\rangle$ is equal to ϕ_s. Thus we obtain

$$\Gamma[\phi] = S[\phi] + \frac{i\hbar}{2}\mathrm{tr}\ \log[\partial^2 + V''(\phi)] + \mathcal{O}(\hbar^2). \tag{3.117}$$

Nice though this formula looks, in practice it is impossible to evaluate the trace for arbitrary $\phi(x)$: We have t ofind all the eigenvalues of the operator $\partial^2 + V''(\phi)$, take their log, and sum. Our task simplifies drastically if we are content with studying $\Gamma[\phi]$ for ϕ independent of x, in which case $V''(\phi)$ is a constant and the operator $\partial^2 + V''(\phi)$ is translation invariant and easily treated in momentum space,

$$\mathrm{tr}\ \log[\partial^2 + V''(\phi)] = \int \mathrm{d}^4x \langle x|\log[\partial^2 + V''(\phi)]|x\rangle \tag{3.118}$$

$$= \int \mathrm{d}^4x \int \frac{\mathrm{d}^4k}{(2\pi)^4}\langle x|k\rangle\langle k|\log[\partial^2 + V''(\phi)]|k\rangle\langle k|x\rangle \tag{3.119}$$

$$= \int \mathrm{d}^4x \int \frac{\mathrm{d}^4l}{(2\pi)^4}\log[-k^2 + V''(\phi)], \tag{3.120}$$

where we took the Fourier transformation and used the plane wave $\langle k|x\rangle = e^{ikx}$. The result is not divergent, but still infinite. We can do renormalization and regulate the integral, impose counter-terms to cancel the divergences, but a finite piece will be remaining. Since we are talking about the effective potential we can always add constants (global shifts) to the effective potential,

$$V_{eff}(\phi) = V(\phi) - \frac{i\hbar}{2}\int \frac{\mathrm{d}^4k}{(2\pi)^4}\log\left[\frac{k^2 - V''(\phi)}{k^2}\right] + \mathcal{O}(\hbar^2), \tag{3.121}$$

and get the result known as the Coleman-Weinberg effective potential.

The effective potential at nonzero temperature In quantum field theory at zero temperature, the expectation value ϕ_c of a scalar field ϕ (also referred to as the classical field) is determined by minimizing the effective potential $V(\phi_c)$. The effective potential contains a tree-level potential term, which can be read off from the Hamiltonian, and quantum corrections form various loop orders. The one-loop quantum correction is calculated from various loop orders. The one-loop quantum correction is calculated by shifting the fields ϕ by their expectation values ϕ_c and isolating the terms $\mathcal{L}_{quad}(\phi_c, \tilde{\phi})$ in the Lagrangian which are quadratic in the shifted fields $\tilde{\phi}$. If we write

$$V(\phi_c) = V_0(\phi_c) + V_1(\phi_c) \tag{3.122}$$

where V_0 is the tree-level contribution and V_1 is the one-loop quantum correction then, for a single scalar field,

$$\exp\left(-i\int \mathrm{d}^4x V_1(\phi_c)\right) = \int \mathcal{D}\tilde{\phi}\exp\left(i\int \mathrm{d}^4x\mathcal{L}_{quad}(\phi_c,\tilde{\phi})\right) \tag{3.123}$$

where $\int \mathcal{D}\tilde{\phi}$ denotes a path integral.

At finite temperature, scalar fields $\phi(t, \mathbf{x})$ are replaced by fields $\phi(\tau, \mathbf{x})$ periodic in τ with period β, where β is given by T^{-1}. We now write the finite-temperature effective potential $\tilde{V}(\phi_c)$ as

$$\tilde{V}(\phi_c) = \tilde{V}_0(\phi_c) + \tilde{V}_1(\phi_c) \tag{3.124}$$

where $\tilde{V}_0$ and $\tilde{V}_1$ are the tree-level and one-loop terms and the expectation value ϕ_c is now a thermal average. Then we get

$$\exp\left(-\int_0^\beta \mathrm{d}\tau\int \mathrm{d}^3x\tilde{V}_1(\phi_c)\right) = \int_{\text{periodic}} \mathcal{D}\tilde{\phi}\exp\left(\int_0^\beta \mathrm{d}\tau\int \mathrm{d}^3x\mathcal{L}_{quad}(\phi_c,\tilde{\phi})\right). \tag{3.125}$$

If gauge fields and fermion fields are included (but with only scalar fields being given expectation values to avoid breaking Lorentz invariance), then path integrals over gauge fields and their associated Fadeev-Popov ghosts and over antiperiodic fermion fields are included as well.

The finite-temperature Lagrangian of the $U(1)$ Higgs model is given by

$$\mathcal{L} = (D_\mu\phi)(D^\mu\phi)^* - m^2\phi^*\phi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2f}(\partial_\mu A^\mu)^2 + \partial_\mu\eta^*\partial^\mu\eta + V(\phi), \tag{3.126}$$

and the potential of the usual Higgs model,

$$V(\phi) = \frac{\lambda}{16}(|\phi|^2 - v^2). \tag{3.127}$$

We then get for the effective potential to one-loop order,

$$V_{eff} = -\frac{4\pi^2T^4}{90} + \frac{1}{2}m^2(T)\phi_c^2 - \frac{CT}{3}\phi_c^3 + \frac{\lambda}{16}\phi^4 \tag{3.128}$$

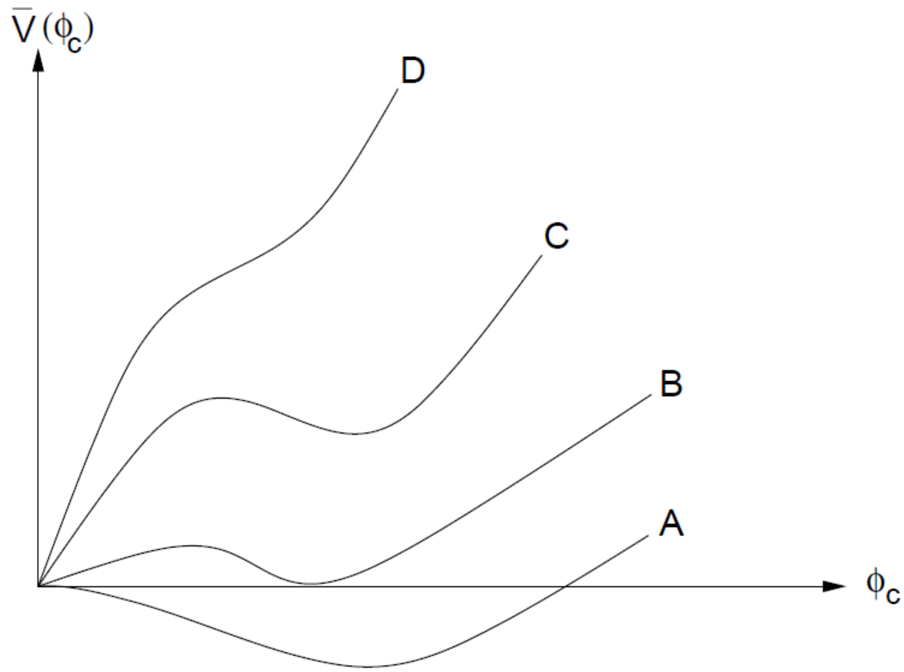

Figure 3.6: $T_d > T_c > T_b > T_a$

with $e^4 \gg \lambda$ and

$$m^2(T) = m^2 + \frac{\lambda + 3e^2}{12} T^2 \tag{3.129}$$

and

$$4\pi C \approx \left(\frac{3\lambda}{4}\right)^{\frac{3}{2}} + \left(\frac{\lambda}{4}\right)^{\frac{3}{2}} + 3e^3. \tag{3.130}$$

We observe that $m^2(T)$ can be positive or negative at zero energy. However, at high energy it will be much smaller.

On the figure, we observe on Fig. 3.6 that we have phase transitions, jumping from one minimum to another, and perhaps, finally, to the global one, as the vacuum relaxes. These are transitions from symmetric system down to system with spontaneously broken symmetries.

This is a viable way of breaking down the models of the unvierse. It cools down and dynamical symmetry breaking occurs. The consequences is that it can give rise to topological objects and extended field configurations. We found earlier that they exist and have finite energy. Recall that we studied kinks (domain walls), vortices (cosmic strings) and magnetic monopoles earlier.

The kink was an interpolation between two points with $\langle\phi\rangle = v$ and $\langle\phi\rangle = -v$ (see subsection 3.1). In a three-dimensional geometry, the kink is interpreted as domain wall.

It is of interest that computing the energy over the line is equivalent to computing the energy density on the surface of the domain wall, and we remark that the energy of solitons can be computed. We have an upper constraint by the size of the universe too. Knowing the energy and the limit of the size, we can estimate the energy density due to the domain walls. However, it turns out that this rough estimate is too much for domain walls, vortices or magnetic monopoles.

One possible solution might be that those topological defects would not have formed, although topology tells us that they do, or that they might have existed at one point in the universe. During an inflation, however, all those densities were diluted, we then can compute the gravitational effects and discuss phenomenological implications. This small density also suggests that we had an inflationary period in our universe.

3.6 Baryon and Lepton Number Violation

Baryon and Lepton number are considered, in a generalised sense, to be charges, which are space integrals of densities. So we start with the definition of baryonic and leptonic currents:

$$j_B^\mu = \frac{1}{3}\sum(\bar{q}L\gamma^\mu q_L + \bar{u}_R\gamma^\mu u_R + \bar{d}_R\gamma^\mu d_R) j_L^\mu \quad = \sum(\bar{l}\gamma^\mu l + \bar{e}_R\gamma^\mu e_R + \bar{\nu}_R\gamma^\mu \nu_R) \tag{3.131}$$

with l doublett and f_r singletts, such that

$$\text{Baryon Number}: \quad B = \int \mathrm{d}^3 x j_B^0 \tag{3.132}$$

$$\text{Lepton Number}: L = \int \mathrm{d}^3 x j_L^0. \tag{3.133}$$

We know that electroweak only couples to the left-handed components, not to the right-handed ones. With this, it contains an axial vector coupling. This may lead to a chiral anomaly. In fact, both currents are non-conserved because of the electroweak ($SU(2)_L$) anomaly. We are coupling here to the left-handed field strength tensor. Analogously for leptons.

The anomalies are given by,

$$\partial_\mu j_B^\mu = \frac{N_f}{16\pi^2} g^2 F_{\mu\nu}^a \tilde{F}^{\mu\nu,a}, \tag{3.134}$$

$$\partial_\mu j_L^\mu = \frac{N_f}{16\pi^2} g^2 F_{\mu\nu}^a \tilde{F}^{\mu\nu,a}, \tag{3.135}$$

$$F_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g\varepsilon^{abc} W_\mu^b W_\nu^c \tag{3.136}$$

$$\tilde{F}_{\mu\nu}^a = \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma} F^{\rho\sigma,r} \tag{3.137}$$

We observe, however, that $B-L$ is conserved while $B+L$ is violated,

$$B-L: \ \partial_\mu(j^\mu_B - j^\mu_L) = 0 \tag{3.138}$$

$$B+L: \partial_\mu(j^\mu_B + j^\mu_L) = \frac{N_f}{8\pi^2} gF\tilde{F}. \tag{3.139}$$

This statement is that $B+L$ is violated, as seen in the context of instantons in QCD, while $B-L$ has to be conserved. Then we can construct field configurations where instantons might exist. So, there might be electroweak instantons mediating this anomaly.

However, if we look at transition rates of the $B+L$ violation with $\Delta(B+L) = 2N_f n$ with n topological charge, we get

$$\Gamma \sim \exp i\frac{16\pi^2}{g^2} \sim l^{-170} \tag{3.140}$$

which is an negligible value. This rate is far too small. However, we have oversimplified our discussion in an important point: the nature of electroweak interactions. We held the discussion analogously to the discussion of unbroken gauge theories, however electroweak theory is a spontaneously broken gauge theory with the Higgs mechanism present. The Higgs does not charge B or L, but has $SU(2)$ charge which indicates that there may exist other, not like instanton, vacuum configurations which might be baryon number violating which, we hope, will provide a stronger violation.

3.7 Sphalerons

We will now discuss vacuum configurations of an $SU(2)$ model with Higgs field,

$$\mathcal{L} = -\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} + |D_\mu\phi|^2 - V(\phi) \tag{3.141}$$

with

$$D_\mu\phi = (\partial_\mu - \frac{ig}{2}\tau^a W^a_\mu)\phi \tag{3.142}$$

$$V(\phi) = \lambda(\phi^\dagger\phi - \frac{1}{2}v^2)^2, \ \lambda > 0. \tag{3.143}$$

How do these vacuum configurations look like? This discussion is different from the usual treatment of the Higgs mechanism. We are interested in more general solution: In the presence of the Higgs potential, do new extended field configurations exist for the W field?

Proceeding in the usual manner, we use the Hamiltontian approach: Choose gauge $W^a_\mu = 0$

$$\mathcal{L} = \frac{1}{2}(\dot{W}^a_i)^2 - \frac{1}{4}F^a_{ij}F^{aij} - |D_i\phi|^2 - V(\phi) \tag{3.144}$$

and canonical variables,

$$\pi_{W_i} = \frac{\delta \mathcal{L}}{\delta W_i} = \dot{W}_i \tag{3.145}$$

$$\pi_\phi = \dot{\phi}^* \tag{3.146}$$

$$\pi_{\phi^*} = \dot{\phi} \tag{3.147}$$

which yields the Hamiltonian density (which is positive definite)

$$\mathcal{H} = \pi_{W_i}\dot{W}_i + \pi_\phi\dot{\phi} + \pi_{\phi^*}\dot{\phi}^* - \mathcal{L} \tag{3.148}$$

$$= \frac{1}{2}(\dot{W}_i)^2 + |\dot{\phi}|^2 + \frac{1}{4}F^a_{ij}F^{aij} + |D_i\phi|^2 + V(\phi) \tag{3.149}$$

and we get the vacuum solution

$$\dot{W}^a_i = 0, \quad \dot{\phi} = 0, \quad F^2_{ij} = 0, \quad D_i\phi = 0, \quad V(\phi) = 0. \tag{3.150}$$

This solution exists and is realized by a pure gauge configuration, where a hatted operator acts in $SU(2)$ space, is:

$$\hat{W}_i(x) = W^a_i\tau^a = -\frac{2i}{g}(\partial_i\hat{\Lambda}(x))\hat{\Lambda}^{-1}(x), \tag{3.151}$$

$$\phi(x) = \frac{v}{\sqrt{2}}\hat{\Lambda}(x)\begin{pmatrix}0\\1\end{pmatrix}, \tag{3.152}$$

where the gauge function $\hat{\Lambda}$ variies in space. $\hat{W}_i(x)$ would not be a solution alone. By introducing the Higgs field $\phi(x)$ we get a stable topologically non-trivial gauge field solution by the interaction of those two.

How do we classify the solutions? They are, again, classified by topology and we use the same definition of the charge as in QCD. Thus, the topology of field configurations is classified by topological charge Q_k. We have time-independent solutions of the Yang-Mills equations of $SU(2)$ changing topological charge. It shifts $B+L$ number by $2N_f$ units (the same situation as in the axial anomaly in QED).

In other words, sphalerons are static solutions of field equations connecting $B+L=0$ state at $r \to 0$ with $B+L = 2N_f$ state at $r \to \infty$. Recall that for instantons we had $t \to \infty$ instead of the space variable r.

The next question is how do we excite these field configurations. Do they have themselves a contribution to the energy? How much energy does it take to generate such a configuration? It is not clear that they can be generated out of the void. Perhaps there are barriers like in the case of instantons, in other words, an energy threshold.

In computing the energy, we start with the field equations,

$$(D_iF^a_{ij}) = \frac{i}{2}g(\phi^\dagger\tau^a(D_j\phi) - (D_j\phi^\dagger\tau^a\phi)) \tag{3.153}$$

$$D^2_i\phi = 2\lambda(\phi^\dagger\phi - \frac{v^2}{2})\phi \tag{3.154}$$

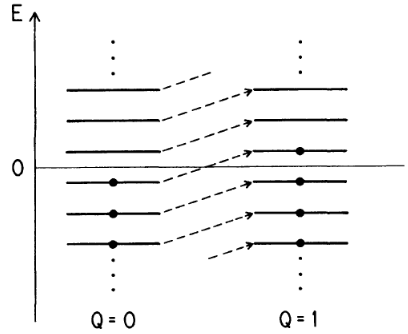

Figure 3.7: Fermion energy levels for two states with different topological numbers, showing that the state with $Q = 1$ has unit baryon number.

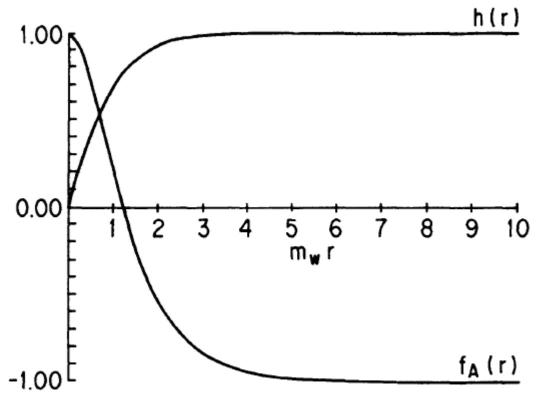

Figure 3.8: The functions $f(r)$ and $h(r)$ for the sphaleron.

and calculate the solutions,

$$W_i^a = \frac{2i}{g} f(r)(i\varepsilon_{iaj} x_j \frac{1}{r^2}) \tag{3.155}$$

$$\phi = \frac{v}{\sqrt{2}} h(r) i \frac{\vec{\tau} \cdot \vec{x}}{r} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{3.156}$$

by connecting a group index with a space index, and we need

$$h(r \to 0) = 0, f(r \to 0) = 1, h(r \to \infty) = 1, f(r \to \infty) = -1 \tag{3.157}$$

Then for a specified field configuration, we can calculate the topological charge,

$$\frac{1}{2}, \tag{3.158}$$

and the mass,

$$M_{sph} = \int \mathrm{d}^3 x \mathcal{H} = c m_W \frac{8\pi}{g^2} \geq\sim 2 \text{ TeV}^2 \tag{3.159}$$

where c is $\mathcal{O}(1)$, and $c = 1.6$ for $\lambda \to 0$ and $c = 2.7$ for $\lambda \to \infty$. It mediates transitions across a potential barrier between $B + L = 0$ and $B + L = 2N_f$ vacua with the height of this potential barrier given by the mass, $\sim M_{sph}$. The transition rate is proportional to the exponential, $\sim \exp -\frac{M_{sph}}{T}$. This means that they become significant for high temperature, $T \gg M_W$. This means that these sphyerons may be created at the LHC which are not point-like particles, but are intermediate field configurations mediating between different states. It would be possible, with sufficient energy, to produce a field configuration to mediate the violation. The signature would not be a resonance, but a multi-particle state that violates the $B + L$ conservation. One may imagine a $t\bar{b}$ state which violates for all generations. They are, however, hard to detect and they are dominant in the sense that they are equivalent to electroweak at higher energy, where one might see substantial sphyerons processes at higher energies. They are hard to detect because we have to look at a global property of the final state.

How does this translate into practice for baryon symmetry.

3.8 Baryogenesis

We have observed a baryon asymmetry,

$$\eta = \frac{n_B - n_{\bar{B}}}{n_\gamma} \sim 5 \cdot 10^{-10} \tag{3.160}$$

where n_γ are the photons in the cosmic microwave background where the number of photon and baryon-antibaryon is assumed to be equivalent in the early universe. If we start with a baryon-antibaryon symmetric universe, we would expect it to be roughly equal to the number of the photon. However, we observe, that most baryon-antibaryon have violated. We observe, however, a slight left-over.

We try to formulate necessary conditions for generation of this η (Sakharov). We find the condition for our field theory Lagrangian:

- baryon number violating interactions
- C and CP violation (otherwise B excess $= \bar{B}$ excess $= 0$)
- universe out of thermal equilibrium at least at one point in its history. Thermal equilibrium means that globally particle numbers are conserved while locally reaction rate to creation and annihilation are the same. Otherwise we would have $\Gamma(X_1 X_2 \overset{\not{B}}{\to} Y_1 Y_2) = \Gamma(Y_1 Y_2 \overset{\not{B}}{\to} X_1 X_2)$. We find in the standard model for the $SU(2)_L$
  - C violation (only e_L and $\bar{e}_R$ couple)
  - CP violation (CKM for quarks)

– B violation through sphaleron processes, non-perturbative tunneling between different vacuum states, equivalent to $V_1 \to V_2 = V_1 \pm \Pi_{i=1}^{n_f} |u_{L_i} d_{L_i} d_{L_i} \nu_i\rangle$. They change B, L but conserve all other charges. Charges $|u_L d_L d_L \nu\rangle$ given by

charge	u_L	d_L	d_L	ν	total
T_3	$\frac{1}{2}$	-0.5	$-.5$	$.5$	0
Y	1/3	1/3	1/3	-1	0
Q	2/3	$-1/3$	$-1/3$	0	0
B	1/3	1/3	1/3	0	1
L	0	0	0	1	1
$B-L$	1/3	1/3	1/3	-1	0

(3.161)

where we observe that $B-L$ conserved and $B+L$ violated. sphaleron processes in thermal equilibrium for $T \gg M_W$. If in thermal equilibrium, statistical rate of lowering and raising are the same and they are thus not able to generate a B, L violation. So we have to start with a non-vanishing $B-L$ at high energy in order to reshuffle by thermal processes. So, we ned either of:

- need mechanism to generate $B-L$ anomaly at high temperatures
- baryon asymmetry generated below $T \leq\sim M_W$ (electroweak baryogenesis)

Sphaleron processes in thermal equilibrium do not change the global number. However, if we start with a $B-L$ asymmetry at high temperature, the sphareons would redistribute that asymmetry and we could use it, e.g., to redistribute a leptonic to baryonic. Only for energies below m_W these asymmetry can be generated.

We know the physics up to the scale M_W pretty well and have constraints by C, CP violation as well. In other words, we have to look into electroweak baryogenesis, worked out in detail with standard model constraints (with Higgs) and has been ruled out because it was too low for the baryon asymmetry by the order of 10^2.

This leaves us with baryogenesis from leptogenesis. Now, if we have

$$\langle B\rangle = 0, \quad \langle L\rangle \neq 0 \Rightarrow \langle L\rangle = -\langle B-L\rangle \text{ at high } T \tag{3.162}$$

Then sphaleron processes redistribute in thermal equilibrium $\langle B-L\rangle \neq 0$ between baryons and leptons.

The constraints on baryon number and CP violaton in the baryonic sector are much stronger than those in the leptonic sector. At present, the CKM mechanism is the only we have for CP violations. We have much weaker constraints on the leptonic sector. In fact, CP violation has not been observed in the leptonic sector. One can postulate by ultra heavy Majorana-neutrinos.

$\langle L\rangle \neq 0$ from CP violation in decay of ultra-heavy Majorana-neutrinos ($N = \bar{N}$).

Then we get the total reaction rate,

$$\Gamma_{tot} = \Gamma(N_i \to H^+ l_L^-) + \Gamma(N_i \to H^- + l_R^+) \sim (hh^\dagger)_{ii} M_i, \tag{3.163}$$

where the reactions are out of equilibrium for a temperature below the mass of the Majorana-Neutrino, $T \ll M_i$. Then we have

- C violated
- CP violated if $\rangle\Updownarrow h \neq 0$ for a Higgs (not necessarily standard model Higgs)

Chapter 4

(Quantum) Fields in Curved Spacetime

Literature:

- L. Bergstrom, A. Goober: Cosmology & Particle Astrophysics

We have a number of unexplained phenomena in the standard model of cosmology observed by experiments. They include:

- Absence of GUT monopoles and topological defects:

  The production rate of GUT monopoles and its flux can be computed. However, the computed quantity is of several orders of magnitudes larger than current observation.

- Flatness of the universe:

  Measured by red-shift of distant supernovae, size of angular defects in cosmic microwave background, etc.

- Homogeneity of the universe:

  From the cosmic microwave background we have measured a homogeneity of the order $\frac{\Delta T}{T} \approx 10^{-5}$.

These phenomena suggest an epoch of exponential expansion of the universe induced by vacuum energy, also known as 'inflation'. One of the results of inflation is the diluted density of monopoles as well as the homogeneity of the universe. It is assumed that during the inflation epoch a small thermally connected region is being expanded exponentially which yields in flat universe after the expansion.

4.1 Evolution of the Universe

We discuss the evolution of the universe by applying Einstein's field equations (we have absorbed the cosmological constant in the matter density $T_{\mu\nu}$),

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu}, \tag{4.1}$$

using the Robertson-Walker metric,

$$\mathrm{d}s^2 = \mathrm{d}t^2 - a^2(t)\left(\frac{\mathrm{d}r^2}{1-kr^2} + r^2\mathrm{d}\theta^2 + r^2\sin^2\theta\mathrm{d}\phi^2\right), \tag{4.2}$$

to an homogeneous, isotropic medium (due to Friedman-Robertson-Lemaitre-Walker, FRLW). The matter density for such an homogeneous, isotropic medium is given by,

$$T_{\mu\nu} = (p+\rho)u_\mu u_\nu - p\ g_{\mu\nu}, \tag{4.3}$$

with p pressure, ρ density, $u_\mu = \frac{\mathrm{d}x_\mu}{\mathrm{d}\tau}$ flow-velocity and $u_\mu = {}^t(1,\vec{0})$ in comoving frame.

We discuss different media categorized by

- vacuum with $p_\Lambda = -\rho_\Lambda$ and

$$T^\Lambda_{\mu\nu} = \mathrm{diag}(\rho_\Lambda, -\rho_\Lambda, -\rho_\Lambda, -\rho_\Lambda), \tag{4.4}$$

- (non-relativistic) diluted matter, $p_m = 0$, given by

$$T^m_{\mu\nu} = \mathrm{diag}(\rho_m, \mathbf{0}). \tag{4.5}$$

- radiation with $p_r = \frac{\rho_r}{3}$ and

$$T^r_{\mu\nu} = \mathrm{diag}(\rho_r, \frac{\rho_r}{3}, \frac{\rho_r}{3}, \frac{\rho_r}{3}). \tag{4.6}$$

We arrive at the Friedman equations with the only non-zero components given by

$$\text{for}\quad \mu=\nu=0\quad 3\left(\frac{\dot{a}}{a}\right) + \frac{3K}{a^2} = 8\pi G\rho, \tag{4.7}$$

$$\mu=\nu=(1,2,3) 2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2} = -8\pi G p. \tag{4.8}$$

We now proceed by discussing several types of different universes:

Flat and vacuum dominant universe In a flat, $K = 0$, and vacuum dominant, $\rho = \rho_\Lambda$, universe we get a Hubble Parameter $H(t) := \frac{\dot{a}}{a}$ characterised by the differential equation

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{\Lambda}{3} = [H(t)]^2, \tag{4.9}$$

which corresponds to

$$\frac{\mathrm{d}H(t)}{\mathrm{d}t} = 0, \tag{4.10}$$

which has the solution of an exponential growth

$$a(t) = a_0 \exp Ht. \tag{4.11}$$

Flat universe with radiation or matter In a flat universe ($k = 0$) with radiation or matter we have the following differential equation expressing energy conservation in a given volume $\mathrm{d}V \approx a^3$,

$$\mathrm{d}E = -p\mathrm{d}V \tag{4.12}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(\rho a^3) = -p\frac{\mathrm{d}}{\mathrm{d}t}a^3. \tag{4.13}$$

We find for the universe with radiation,

$$\dot{\rho} + 4\rho\frac{\dot{a}}{a} = 0 \ \Rightarrow \rho = \frac{\rho_0}{a^4}, \tag{4.14}$$

and matter,

$$\dot{\rho} + 3\rho\frac{\dot{a}}{a} = 0 \ \Rightarrow \rho = \frac{\rho_0}{a^3}. \tag{4.15}$$

Applying these results in the Friedmann equation,

$$\dot{a}^2 = \frac{8\pi G\rho(t)}{3}a^3, \tag{4.16}$$

yields for the radiation dominated universe,

$$a(t) = \left(\frac{t}{t_0}\right)^{\frac{1}{2}} \tag{4.17}$$

$$\rho(t) = \frac{\rho_0 t_0^2}{t^3} \tag{4.18}$$

and for the matter dominated,

$$a(t) = \left(\frac{t}{t_0}\right)^{2/3} \tag{4.19}$$

$$\rho(t) = \frac{\rho_0 t_0^2}{t^2}. \tag{4.20}$$

Universe with curvature For a universe characterised by curvature $k = \pm 1$, the equation

$$H^2(t) = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G\rho}{3} - \frac{k}{a^2} \tag{4.21}$$

has a fix point solution. When defining the cirtical density $\rho_c = \frac{3H^2(t)}{8\pi G}$ and normalizing up to $\Omega_i = \frac{\rho_i}{\rho_c}$ we find

$$\frac{\rho - \rho_c}{\rho} = \frac{3k}{8\pi G\rho a^2}. \tag{4.22}$$

From present day observation, however, we have

$$\left|\frac{\rho - \rho_c}{\rho}\right| \approx \mathcal{O}(1). \tag{4.23}$$

When summarizing the evolutiong of the different epochs for $\rho \to \rho_c$,

$$\text{vacuum dominated:} \quad \frac{\rho - \rho_c}{\rho} \approx e^{-2Ht}, \tag{4.24}$$

$$\text{radiation dominated:} \quad \frac{\rho - \rho_c}{\rho} \approx t, \tag{4.25}$$

$$\text{matter dominated:} \quad \frac{\rho - \rho_c}{\rho} \approx t^{2/3}, \tag{4.26}$$

we note that in the radiation and matter dominated epochs any deviation from $\rho = \rho_c$ is amplifed by time.

4.2 Inflation

We will first discuss this epoch in a classical field theoretic framework. Later, we modify it to discuss it in a quantum field theory picture.

The field theoretic model for an inflationary epoch in the early universe (chaotic inflation, A. Linde).

- at $t \cong t_{pl} - 5.4 \cdot 10^{-44}$ s, have an area of size $l_{pl}^3 \cong (1.63 \cdot 10^{-35}\text{ m})^3$ filled with a homogeneous scalar field ϕ (inflaton) with potential $V(\phi)$. The action is given by

$$S = \int \sqrt{-g}\mathrm{d}^4x \left(\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial^\nu\phi - V(\phi)\right) \tag{4.27}$$

and homogeneity is expressed by $\Delta\phi \ll V(\phi)$. This means that we can apply FRLW metric and find

$$S = \int \mathrm{d}t a^3 (\frac{1}{2}\dot{\phi}^2 - V(\phi)) = \int \mathrm{d}t \mathcal{L}. \tag{4.28}$$

This Lagrangian allows us to find classical equations of motions for ϕ,

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\phi}} - \frac{\partial \mathcal{L}}{\partial \phi} = 0 \tag{4.29}$$

$$\Rightarrow \ddot{\phi} + 3\frac{\dot{a}}{a}\dot{\phi} = -\frac{\partial V}{\partial \phi} \tag{4.30}$$

is a classical motion in potential V, with a time-dependent friction term $3\frac{\dot{a}}{a} = 3H$.

We calculate the content of the Friedman equations with $k = 0$,

$$\rho = T_{00} = \frac{1}{2}\dot{\phi}^2 + V(\phi), \tag{4.31}$$

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3M_{pl}^2}(\frac{1}{2}\dot{\phi}^2 + V(\phi)), \tag{4.32}$$

where ρ behaves like a vacuum energy density, then ϕ varies only slowly with time; i.e. $\frac{\dot{a}}{a}\dot{\phi} \gg \ddot{\phi}$ and $V(\phi) \gg \dot{\phi}^2$ (slow-roll conditions).

With these slow-roll conditions our equations of motions simplify to

$$3\frac{\dot{a}}{a}\dot{\phi} = -\frac{\partial V}{\partial \phi} \tag{4.33}$$

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3M_{pl}^2}V(\phi) \tag{4.34}$$

and we have for large $V(\phi)$ a large H and thus the slow-roll condition. Looking at the dynamical evolution of the ϕ field we observe that for some time there is a violation. How large?

$$\dot{\phi}^2 \sim (\partial_\phi V)^2 \frac{1}{(\dot{a}/a)^2} \sim (\partial_\phi V)^2 \frac{M_{pl}^2}{V}. \tag{4.35}$$

We then have for the slow-roll condition

$$\partial_\phi V \ll \frac{V}{M_{pl}^2} \tag{4.36}$$

$$\partial_\phi V \sim \frac{V}{\phi} \text{ for any polynomial potential} \tag{4.37}$$

and finally

$$\phi \gg M_{pl} \tag{4.38}$$

We see that our slow-roll conditions are fulfilled when the field is much larger than the Planck-Mass.

However, it does not mean that the mass energy density beyond the Planck-density $\rho_{pl} \sim M_{pl}^2$, since $V = \lambda\phi^4$

$$V(\phi \sim M_{pl}) - \lambda M_{pl}^4 \ll M_{pl}^4 \text{ for } \lambda \ll 1. \tag{4.39}$$

In other words, the start of inflation $V(\phi_{in}) \sim M_{pl}^4$

$$\phi_{in} \sim \lambda^{-1/4} M_{pl} \gg M_{pl}$$

and our of inflation for $\phi_{out} \sim M_{pl}$ yields the end of slow-roll and we have a damped oscillations around $\phi = 0$ (reheating).

The moment where it ends it sets the initial conditions for our universe, i.e. the curvature. In late-time epoch of inflation we want annihilation of inflaton and creation of SM particles.

4.3 Quantum theory of the inflaton

The free Lagrangian of the massive scalar inflaton field in curved spacetime is given by,

$$\mathcal{L} = \frac{1}{2}\sqrt{-g}(g^{\mu\nu}(x)\partial_\mu\phi\partial_\nu\phi - m^2\phi^2) \tag{4.40}$$

where we use the FRLW metric $g_{\mu\nu} = \text{diag}(1, -a^2, -a^2, -a^2)$ such that $\sqrt{-g} = a^3$. We use the Euler-Lagrange equations,

$$\frac{\partial}{\partial x^\mu}\left[\frac{\partial\mathcal{L}}{\partial(\partial\phi/\partial x^\mu)}\right] - \frac{\partial\mathcal{L}}{\partial\phi} = 0, \tag{4.41}$$

to obtain the equation of motion,

$$[\partial_g^2 + m^2]\phi(x) = 0 \tag{4.42}$$

where

$$\partial_g^2\ \phi(x) = g^{\mu\nu}D_\mu D_\nu\phi(x) = \frac{1}{\sqrt{-g}}\partial_\mu[\sqrt{-g}g^{\mu\nu}(x)\partial_\nu\phi(x)]. \tag{4.43}$$

Rewritten in the FRLW metric,

$$\ddot{\phi} + 3\frac{\dot{a}}{a}\dot{\phi} - \frac{\Delta\phi}{a^2} + m^2\phi(x) = 0. \tag{4.44}$$

We now evaluate how the field evolves for different behaviour of the scale factor $a(t)$. For an generic inflationary period with a nearly constant Hubble parameter, $H = \dot{a}/a \sim$ const, and the exponential scale factor,

$$a(t) = a_0 \exp H(t - t_0). \tag{4.45}$$

Using conformal time,

$$\eta = \int^t \frac{\mathrm{d}t}{a(t)} = -\frac{1}{a_0 H \exp H(t - t_0)} + c = -\frac{1}{a(\eta)H}, \tag{4.46}$$

we can rewrite our metric as being proportional to the Minkowski metric,

$$\mathrm{d}s^2 = a^2(\eta)[\mathrm{d}\eta^2 - \mathrm{d}\vec{r}^2]. \tag{4.47}$$

To quantize the field, we calculate the canonical momentum,

$$\pi_\phi = \frac{\partial \mathcal{L}}{\partial(\frac{\partial \phi}{\partial \eta})} = a^2(\eta)\frac{\partial \phi}{\partial \eta}, \tag{4.48}$$

and impose the canonical quantization condition at equal conformal time,

$$[\phi, \pi]_\eta = a^2(\eta)[\phi(\eta, \vec{r}), \frac{\partial \phi}{\partial \eta}(\eta, \vec{r}')] = i\delta^{(3)}(\vec{r} - \vec{r}'). \tag{4.49}$$

We then compute the mode expansion in terms of momentum eigenfunctions,

$$\phi(\eta, \vec{r}) = \int \frac{\mathrm{d}^3 k}{(2\pi)^{3/2}}[a_{\vec{k}}\phi_{\vec{k}}(\eta)e^{i\vec{k}\cdot\vec{r}} + a^\dagger_{\vec{k}}\phi^*_{\vec{k}}(\eta)e^{-i\vec{k}\cdot\vec{r}}] \tag{4.50}$$

to find that it fulfills the quantization condition if

$$a^2(\eta)(\phi_{\vec{k}}\frac{\partial \phi^*_{\vec{k}}}{\partial \eta} - \phi_{\vec{k}}\frac{\partial \phi_{\vec{k}}}{\partial \eta}) = i. \tag{4.51}$$

Here we have the problem that η is proportional changing with time. So, the differential equations determing our eigenfunctions are much more complicated than the Klein-Gordon equations because we have an explicit dependence on the parameter η. The problem is finding momentum eigenfunctions, taking into account explicit dependence on η.

To find the momentum eigenfunctions we use the ansatz

$$\phi_{\vec{k}}(\eta) = \chi_{\vec{k}}(\eta)\frac{1}{a(\eta)} \tag{4.52}$$

in the equation of motions,

$$\ddot{\chi}_{\vec{k}} + \left(\frac{m^2}{H^2\eta^2} - \frac{\ddot{a}}{a} \right) \chi_{\vec{k}} = 0 \tag{4.53}$$

with $\frac{\ddot{a}}{a} = \frac{2}{\eta^2}$. This differential equation has a solution in terms of Hankel functions,

$$\chi_{\vec{k}} = \sqrt{-\eta} \left(c_1 H_\nu^{(1)}(-k\eta) + c_2 H_\nu^{(2)}(-k\eta) \right) \tag{4.54}$$

with $H_\nu^{(1)*} = H_\nu^{(2)}$ and $\nu^2 = \frac{9}{4} - \frac{m^2}{H^2}$.

Looking at asymptotic behavior at early times we find,

$$\chi_{\vec{k}}(k\eta \to -\infty) \to \sqrt{\frac{2}{\pi k}}(c_1 e^{-ik\eta} + c_2 e^{ik\eta}) \tag{4.55}$$

which tells us what positive and negative frequency modes at early times were. In ordinary field theory one would then define the asymptotic frequencies as in and out states for any given times. Here we cannot use this any more. However, we can use that they do not mix at asympotically early or late times and thus derive the inflaton field operator, with positive frequency modes $c_1 = \frac{\sqrt{\pi}}{2}, c_2 = 0$ and vice-versa. It then yields the inflaton field operator,

$$\phi = \frac{\sqrt{-\pi\eta}}{2a(\eta)} \int \frac{\mathrm{d}^3k}{(2\pi)^{3/2}} [a_{\vec{k}} H_\nu^{(1)}(-k\eta) e^{i\vec{k}\cdot\vec{r}} + a_{\vec{k}}^\dagger H_\nu^{(2)}(-k\eta) e^{-i\vec{k}\cdot\vec{r}}] \tag{4.56}$$

This is a new feature of fields on curved spacetime because it yields the creation of the standard model particles. However, it does not allow to determine feedback on the curved background; in other words, it is basically statically curved.

Bibliography

[1] Anastasiou, *Quantum Field Theory 2*, Lecture Notes at ETH Zurich 2008.

[2] Bailin, *Cosmology In Gauge Field Theory And String Theory*, IOP Publishing 2004.

[3] Bergstrom, *Cosmology and Particle Astrophysics*, Springer 2004.

[4] Bertlmann, *Anomalies in Quantum Field Theory*, Oxford University Press 1996.

[5] Coleman, *Aspects Of Symmetry*, Cambridge University Press 1985.

[6] Collins, *Renormalization*, Cambridge University Press 1984.

[7] Donoghue, *Dynamics Of The Standard Model*, Cambridge University Press 1992.

[8] Ecker, *Chiral Perturbation Theory*, arXiv:hep-ph/9501357

[9] Frampton, *Gauge Field Theories*, John Wiley & Sons 2000.

[10] Fukugita, *Physics of Neutrinos and Application to Astrophysics*, Springer 2003.

[11] Huang, *Quarks Leptons & Gauge Fields*, World Scientific Publishing 1982.

[12] Kaku, *Quantum Field Theory*, Oxford University Press 1993.

[13] Peskin, *Introduction To Quantum Field Theory*, Perseus Book Publishing 1995.

[14] Ryder, *Quantum Field Theory*, Cambridge University Press 1996.

[15] Srednicki, *Quantum Field Theory*, Cambridge University Press 2007.

[16] Weinberg, *The Quantum Theory Of Fields II: Modern Applications*, Cambridge University Press 1996.

[17] Yndurain, *The Theory Of Quark And Gluon Interaction*, Springer 2006.

[18] Zee, *Quantum Field Theory In A Nutshell*, Princeton University Press 2003.