Instantons

Student: Jacob Shapiro
Supervisor: Dr. Philippe de Forcrand

May 30, 2014

Abstract

This document encompasses a report corresponding to a presentation given in ETHZ in the spring semester of 2014 as part of a proseminar on topological objects in physics, directed by Dr. Philippe de Forcrand. Its purpose is to make precise that which was shown in the presentation.

After reviewing some “toy” models in quantum mechanics which allow us to exhibit the important concepts of instantons in a familiar environment–essentially a tunneling description between distinct vacua–a discussion of instantons in non-Abelian gauge field theory follows. Finally a brief chapter on the consequences of instantons in QCD concludes the report.

Note that this document exists in two versions: an unabridged version which contains all proofs with full detail and is not meant to be printed, and an abridged version which contains *no proofs*. You are now reading the unabridged version.

Contents

Part I

Instantons in Quantum Mechanics

Chapter 1

The Cumbersome Harmonic Oscillator

In this chapter we shall derive the ground state energy eigenvalue for the harmonic oscillator in a cumbersome way, following very closely the presentation of [12]. This way shall prove very useful for quantum field theory.

1.1 Path-Integral Formulation of Quantum Mechanics

A point particle of mass $m \equiv 1$ (in properly chosen units) moves in two spacetime dimensions. Its position, varying with time, is described by a function $x \in \mathbb{R}^{\mathbb{R}}$, and we further assume that the particle is under the influence of some potential $V[x](t)$. The Lagrangian for this system is given by $L[x](t) \equiv \frac{1}{2}(\dot{x}(t))^2 - V[x](t)$. Let $\{x_i, x_f, t_i, t_f\} \subset \mathbb{R}$.

1.1.0.1 Fact

According to Feynman's path-integral formulation of quantum mechanics [8], if the particle was initially at x_i at time t_i, the *transition amplitude* of the particle to be found finally at x_f at time t_f–conventionally denoted by $\langle x_f, t_f \,|\, x_i, t_i\rangle$–is given by:

$$\boxed{\langle x_f, t_f \,|\, x_i, t_i\rangle = \mathcal{N} \int_{\{x\in\mathbb{R}^{\mathbb{R}}:\, x(t_f)=x_f \wedge x(t_i)=x_i\}} \mathcal{D}x \exp\left\{ i\frac{\int_{t_i}^{t_f} dt L[x](t)}{\hbar} \right\}} \tag{1.1}$$

1.1.0.2 Remarks

1. $\mathcal{N}$ is a normalization factor which will be determined later.

2. This formulation gives a natural scale for the action, $\hbar$. Now we can answer what is a "large" action, whereas in classical mechanics that notion had no meaning. From here on we shall choose our units such that $\hbar \overset{!}{=} 1$, in order to simplify the formulas.

3. Usually in quantum mechanics the transition amplitude $\langle x_f,\, t_f \,|\, x_i,\, t_i\rangle$ is in fact denoted by $\left\langle x_f \,\middle|\, \exp\left\{i\hat{H}\left(t_f - t_i\right)\right\} \,\middle|\, x_i\right\rangle$ where $\hat{H}$ is the time-evolution operator on the Hilbert space of states of the particle (and the various $|x\rangle$'s are vectors in this space). However, the path-integral formulation was conceived exactly in order to blur the distinction between space and time (which is impossible in the Hamiltonian formulation of mechanics and only possible in the Lagrangian formulation of mechanics) and so we may operate agnostically to the existence of $\hat{H}$ and proceed dealing only with the functional L, which is now the fundamental object of the theory.

1.2 Imaginary Time

An invaluable tool in path integral computations is the Euclidean path integral, which is a formal analytic continuation of the path integral to complex-valued time $\langle x_f,\, -it_f \,|\, x_i,\, -it_i\rangle$.

1.2.1 Ground State and Energy for Large Times

1.2.1.1 Claim

$\lim_{T\to\infty} \left\langle x_f,\, -i\frac{T}{2} \,\middle|\, x_i,\, i\frac{T}{2}\right\rangle = \lim_{T\to\infty} e^{-E_0 T}\psi_0\left(x_f\right)\psi_0^*\left(x_i\right)$ where $\psi_0\left(x\right)$ is the wave function corresponding to the lowest lying energy eigenstate.

Proof

- In the usual quantum mechanical notation, $\left\langle x_f,\, -i\frac{T}{2} \,\middle|\, x_i,\, i\frac{T}{2}\right\rangle = \left\langle x_f \,\middle|\, e^{-T\hat{H}} \,\middle|\, x_i\right\rangle$.
- Let $\left\{|n\rangle\right\}_{n\in\mathbb{N}}$ be a complete orthonormal set of eigenstates of $\hat{H}$ with eigenvalues $\left\{E_n\right\}_{n\in\mathbb{N}}$. Assume that $\exists E_0 = \min\left(\left\{E_j \,:\, j\in\mathbb{N}\right\}\right)$.
- Then

$$\begin{aligned}\left\langle x_f \,\middle|\, e^{-T\hat{H}} \,\middle|\, x_i\right\rangle &= \left\langle x_f \,\middle|\, e^{-T\hat{H}}\left(\sum_{n\in\mathbb{N}}|n\rangle\langle n|\right) \,\middle|\, x_i\right\rangle \\ &= \sum_{n\in\mathbb{N}} e^{-TE_n}\left\langle x_f \,|\, n\right\rangle\left\langle n \,|\, x_i\right\rangle \\ &\overset{T\to\infty}{\longrightarrow} \lim_{T\to\infty} e^{-TE_0}\left\langle x_f \,|\, 0\right\rangle\left\langle 0 \,|\, x_i\right\rangle\end{aligned}$$

■

This is half of why it is useful to work in imaginary time.

1.2.2 Euclidean Time Path Integral

1.2.2.1 Claim

$\left\langle x_f,\, -i\frac{T}{2} \,\middle|\, x_i,\, i\frac{T}{2}\right\rangle = \mathcal{N}\int_{\left\{x\in\mathbb{R}^\mathbb{R} :\, x\left(\frac{T}{2}\right)=x_f \wedge x\left(-\frac{T}{2}\right)=x_i\right\}}\mathcal{D}x\exp\left\{-\int_{-\frac{T}{2}}^{\frac{T}{2}}dt\left[\frac{1}{2}\left(\dot{x}\left(t\right)\right)^2 + V\left[x\right]\left(t\right)\right]\right\}$ for any $T\in\mathbb{R}$.

Proof

- We start with the technical definition of the symbolic expression $\langle x_f,\, t_f\,|\, x_i,\, t_i\rangle = \mathcal{N}\int_{\{x\in\mathbb{R}^{\mathbb{R}}:\, x(t_f)=x_f\wedge x(t_i)=x_i\}}\mathcal{D}x\exp\left\{i\frac{\int_{t_i}^{t_f}dtL[x](t)}{\hbar}\right\}$:
  - We know that
$$\left\langle x_f,\,\frac{\varepsilon}{2}\,\middle|\,x_i,\,-\frac{\varepsilon}{2}\right\rangle \equiv \left\langle x_f\,\middle|\,e^{-i\hat{H}\varepsilon}\,\middle|\,x_i\right\rangle$$
$$\approx \sqrt{\frac{m}{2\pi i\varepsilon}}\exp\left\{i\varepsilon\left[\frac{m}{2}\left(\frac{x_f-x_i}{\varepsilon}\right)^2-V\left(x_f\right)\right]\right\}+\mathcal{O}\left(\varepsilon^2\right)$$
    In the derivation of this formula (expanding the exponent, inserting $\int dp\,|p\rangle\langle p|$ and solving $\int\frac{dp}{2\pi}e^{-i\frac{\varepsilon}{2m}p^2}=\sqrt{\frac{m}{2\pi i\varepsilon}}$), ε is used as a parameter whose character (real or imaginary) is not important. Thus this formula is valid (up to $\mathcal{O}\left(\varepsilon^2\right)$) for $\varepsilon\in\mathbb{C}$ as well.
  - The general limit formula for the transition amplitude, which is, $\left\langle x_f\,\middle|\,e^{-i\hat{H}T}\,\middle|\,x_i\right\rangle = \lim_{n\to\infty}\left(\frac{mn}{2\pi iT}\right)^{\frac{n}{2}}\int dx_1\ldots\int dx_{n-1}\exp\left\{i\sum_{j=1}^n\frac{T}{n}\left[\frac{m}{2}\left(\frac{x_j-x_{j-1}}{T/n}\right)^2-V\left(x_j\right)\right]\right\}$ where $x_n\equiv x_f$ and $x_0\equiv x_i$, is derived from the formula for infinitesimal times, and should thus stay valid also for imaginary values of T.
  - We identify $\lim_{n\to\infty}\sum_{j=1}^n\frac{T}{n}\left[\frac{m}{2}\left(\frac{x_j-x_{j-1}}{T/n}\right)^2-V\left(x_j\right)\right]$ as being equal to $\int_{-\frac{T}{2}}^{\frac{T}{2}}dt\left[\frac{m}{2}\dot{x}\left(t\right)^2-V\left(x\left(t\right)\right)\right]$, where $x\left(t\right)$ is some integrable function in $\mathbb{R}^{\mathbb{R}}$ which interpolates between the fixed points $\left\{x_j\right\}_{j=0}^n$, that is, $x\left(-\frac{T}{2}+j\frac{T}{n}\right)\overset{!}{=}x_j$ for all $j\in\{0,\,\ldots,\,n\}$.
  - The n integrals over x_j are interpreted as spanning the space of all paths $x\left(t\right)$ which have their end points fixed at x_i and x_f and this is written symbolically as $\lim_{n\to\infty}\left(\frac{mn}{2\pi iT}\right)^{\frac{n}{2}}\int dx_1\ldots\int dx_{n-1}=\mathcal{N}\int_{\{x\in\mathbb{R}^{\mathbb{R}}:\, x(\frac{T}{2})=x_f\wedge x(-\frac{T}{2})=x_i\}}\mathcal{D}x$.

- When going to Euclidean spacetime, we plug into the path integral formula $-iT$ instead of T and obtain:
$$\left\langle x_f,\,-i\frac{T}{2}\,\middle|\,x_i,\,i\frac{T}{2}\right\rangle = \left\langle x_f\,\middle|\,e^{-\hat{H}T}\,\middle|\,x_i\right\rangle$$
$$= \lim_{n\to\infty}\left(\frac{mn}{2\pi T}\right)^{\frac{n}{2}}\int dx_1\ldots\int dx_{n-1}e^{\left\{\sum_{j=1}^n\frac{T}{n}\left[\frac{m}{2}\left(\frac{x_j-x_{j-1}}{-iT/n}\right)^2-V(x_j)\right]\right\}}$$
$$= \lim_{n\to\infty}\left(\frac{mn}{2\pi T}\right)^{\frac{n}{2}}\int dx_1\ldots\int dx_{n-1}e^{\left\{-\sum_{j=1}^n\frac{T}{n}\left[\frac{m}{2}\left(\frac{x_j-x_{j-1}}{T/n}\right)^2+V(x_j)\right]\right\}}$$

- The identification for the expression in the exponent is still valid, in fact, the only thing that is different about it is the sign of the potential. $x\left(t\right)$ is still an integrable function in $\mathbb{R}^{\mathbb{R}}$ defined in the very same way. So we have:
$$\left\langle x_f,\,-i\frac{T}{2}\,\middle|\,x_i,\,i\frac{T}{2}\right\rangle = \mathcal{N}'\int_{\{x\in\mathbb{R}^{\mathbb{R}}:\, x(\frac{T}{2})=x_f\wedge x(-\frac{T}{2})=x_i\}}\mathcal{D}x\,\exp\left\{-\int_{-\frac{T}{2}}^{\frac{T}{2}}dt\left[\frac{m}{2}\dot{x}\left(t\right)^2+V\left(x\left(t\right)\right)\right]\right\}$$

- In conclusion when going to Euclidean spacetime, we have *three* differences:

  1. The i in the exponent becomes -1.
  2. There is an extra -1 multiplying the potential in the action integral.
  3. The normalization constant is different: $\lim_{n\to\infty}\left(\frac{mn}{2\pi iT}\right)^{\frac{n}{2}} \mapsto \lim_{n\to\infty}\left(\frac{mn}{2\pi T}\right)^{\frac{n}{2}}$.

  ■

1.2.2.2 Remarks

The effect of Euclidean time on the path integral is two fold:

1. The exponent in the path integral is now real. If we assume that $\min(\{V[x] \mid x\}) = 0$ (and we may do that without loss of generality by choosing the energy scale appropriately) , then the exponent is always negative, which means we may be more optimistic about the convergence of the path integral.

2. It turns out that to calculate path-integrals it is worthwhile to investigate the classical paths of the action first (as we shall see soon). In classical mechanics, what we have achieved by the complex-valued time is equivalently an inverted potential: $V[x] \mapsto -V[x]$. This identification will help us "read off" classical paths using preexisting intutition in classical mechanics (although now we shall employ it on inverted potentials).

1.3 Approximating the Path Integral

In a crude way, assume that for our problem at hand $\int_{-\frac{T}{2}}^{\frac{T}{2}} dt \left\{\frac{1}{2}(\dot{x}(t))^2 + V[x](t)]\right\} \gg 1$ $\forall$ paths to be integrated on. This corresponds to the semiclassical approximation where we assume $\hbar$ is very small compared to the action. Then it is clear that most of the contribution to the path integral will come from those paths which minimize the Euclidean action. This is exactly the definition of the classical paths with potential $(-V[x])$. If there is only *one* such extremum path for the action, which we denote by $x_{cl}(t)$ (that is, we assume that $\ddot{x}_{cl}(t) = V'[x]$), then we can estimate $\mathcal{N}\int_{\{x\in\mathbb{R}^{\mathbb{R}}:\, x(\frac{T}{2})=x_f \wedge x(-\frac{T}{2})=x_i\}} \mathcal{D}x \exp\{-S[x]\} \sim \exp\{-S[x_{cl}]\}$ (by using the symbol S we really mean S_E–this will be true till the end of this text).

If $\exists$ more than one such extremal path, then we would make a reasonable approximation by *summing* over the contribution from each path, assuming these extremum points are well separated in function space.

An analogy can be made to integrating over two Gaussians which are well separated (see Figure 1.1).

For our $V[x]$, the inverted harmonic oscillator, we know that there exists only one solution to the equation of motion which has finite action (infinite action doesn't interest us because its contribution would be zero anyway): $x_{cl}(t) = 0$, with boundary conditions $x_{cl}\left(\pm\frac{T}{2}\right) = 0$.

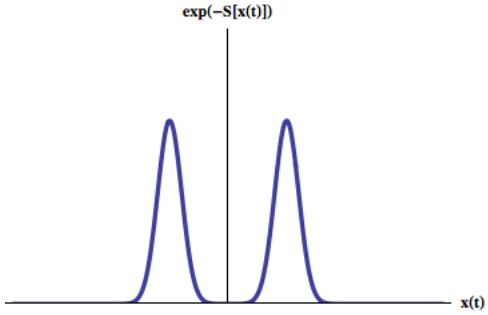

Figure 1.1: The volume under the curve can be approximated as the sum of two separate Gaussian integrals.

1.3.1 Taylor Expanding the Action

Pick some $\varepsilon \in (0,\, 1)$. Define $\eta\,(t) := \frac{1}{\varepsilon}\,[x\,(t) - x_{cl}\,(t)]$, where $x\,(t) \in \left\{x \in \mathbb{R}^{\mathbb{R}} :\, x\left(\frac{T}{2}\right) = x_f \wedge x\left(-\frac{T}{2}\right) = x_i\right\}$ is the set of paths we would be integrating over. Because $x_{cl}\,(t)$ is a solution to the classical equation of motion with the same boundary conditions, we have that $\eta\left(\pm\frac{T}{2}\right) = 0$.

1.3.1.1 Claim

The action can be approximated as $S\,[x_{cl} + \varepsilon\eta] \approx S\,[x_{cl}] + \frac{1}{2}\sum_{n\in\mathbb{N}} c_n^2 \left(\left(\frac{\pi n}{T}\right)^2 + \frac{\partial^2 V[x_{cl},\, \dot{x}_{cl}]}{\partial x^2}\right) + \mathcal{O}\left(\varepsilon^3\right)$ where c_n are expansion coefficients of $\varepsilon\eta$ in a complete set of eigenfunctions of the differential operator $-\frac{d^2}{dt^2} + \frac{\partial^2 V[x_{cl},\, \dot{x}_{cl}]}{\partial x^2}$.

Proof

- We have that $x\,(t) = x_{cl}\,(t) + \varepsilon\eta\,(t)$.
- Treat $S\,[x]$ as an analytic function of ε to make a Taylor expansion of it around $\varepsilon = 0$. Thus formally we have:
  $S\,[x] = S\,[x_{cl} + \varepsilon\eta] \approx S\,[x_{cl}] + \frac{d}{d\varepsilon}S\,[x_{cl} + \varepsilon\eta]\big|_{\varepsilon=0}\,\varepsilon + \frac{1}{2}\,\frac{d^2}{d\varepsilon^2}S\,[x_{cl} + \varepsilon\eta]\Big|_{\varepsilon=0}\,\varepsilon^2 + \mathcal{O}\left(\varepsilon^3\right)$
- Observe that $\frac{d}{d\varepsilon}S\,[x_{cl} + \varepsilon\eta]\big|_{\varepsilon=0} = 0$ by definition. Verification:

$$\begin{aligned}
\left.\frac{d}{d\varepsilon}S\left[x_{cl}+\varepsilon\eta\right]\right|_{\varepsilon=0} &= \left.\frac{d}{d\varepsilon}S\left[x_{cl}+\varepsilon\eta\right]\right|_{\varepsilon=0} \\
&= \left.\int_{-\frac{T}{2}}^{\frac{T}{2}} dt\frac{d}{d\varepsilon}L\left[x+\varepsilon\eta,\,\dot{x}+\varepsilon\dot{\eta}\right]\right|_{\varepsilon=0} = \\
&= \left.\int_{-\frac{T}{2}}^{\frac{T}{2}} dt\left(\frac{\partial L}{\partial x}\left[x+\varepsilon\eta,\,\dot{x}+\varepsilon\dot{\eta}\right]\eta+\frac{\partial L}{\partial \dot{x}}\left[x+\varepsilon\eta,\,\dot{x}+\varepsilon\dot{\eta}\right]\dot{\eta}\right)\right|_{\varepsilon=0} = \\
&= \underbrace{\left(\frac{\partial L}{\partial \dot{x}}\eta\right)_{-\frac{T}{2}}^{\frac{T}{2}}}_{=0 \text{ because } \eta\left(\pm\frac{T}{2}\right)=0} + \int_{-\frac{T}{2}}^{\frac{T}{2}} dt\left(\frac{\partial L}{\partial x}-\frac{d}{dt}\frac{\partial L}{\partial \dot{x}}\right)\eta = \\
&= \int_{-\frac{T}{2}}^{\frac{T}{2}} dt\left(\frac{\partial L}{\partial x}\left[x_{cl},\,\dot{x}_{cl}\right]-\frac{d}{dt}\frac{\partial L}{\partial \dot{x}}\left[x_{cl},\,\dot{x}_{cl}\right]\right)\eta
\end{aligned}$$

Since η was arbitrary we get the classical equation of motion, $\frac{\partial L}{\partial x}\left[x_{cl},\,\dot{x}_{cl}\right]-\frac{d}{dt}\frac{\partial L}{\partial \dot{x}}\left[x_{cl},\,\dot{x}_{cl}\right]=0$.

- $\left.\frac{d^2}{d\varepsilon^2}S\left[x_{cl}+\varepsilon\eta\right]\right|_{\varepsilon=0}$ is slightly more involved:

$$\begin{aligned}
\left.\frac{d^2}{d\varepsilon^2}S\left[x_{cl}+\varepsilon\eta\right]\right|_{\varepsilon=0} &= \left.\frac{d}{d\varepsilon}\int_{-\frac{T}{2}}^{\frac{T}{2}} dt\left(\frac{\partial L}{\partial x}\left[x_{cl}+\varepsilon\eta,\,\dot{x}_{cl}+\varepsilon\dot{\eta}\right]\eta+\frac{\partial L}{\partial \dot{x}}\left[x_{cl}+\varepsilon\eta,\,\dot{x}_{cl}+\varepsilon\dot{\eta}\right]\dot{\eta}\right)\right|_{\varepsilon=0} = \\
&= \int_{-\frac{T}{2}}^{\frac{T}{2}} dt\left(\frac{\partial^2 L}{\partial x^2}\eta^2+2\frac{\partial^2 L}{\partial \dot{x}\partial x}\eta\dot{\eta}+\frac{\partial^2 L}{\partial \dot{x}^2}\dot{\eta}^2\right)
\end{aligned}$$

  - For most Lagrangians (including the harmonic oscillator) $\frac{\partial^2 L}{\partial \dot{x}\partial x}=0$.
  - For Lagrangians of the form $L\left[x\right]=\frac{1}{2}\dot{x}^2-V\left[x\right]$ we have $\frac{\partial^2 L}{\partial \dot{x}^2}=1$ and so $\int_{-\frac{T}{2}}^{\frac{T}{2}} dt\left(\frac{\partial^2 L}{\partial \dot{x}^2}\dot{\eta}^2\right)=\int_{-\frac{T}{2}}^{\frac{T}{2}} dt\dot{\eta}\dot{\eta}=\underbrace{\dot{\eta}\eta|_{-\frac{T}{2}}^{\frac{T}{2}}}_{0}-\int_{-\frac{T}{2}}^{\frac{T}{2}} dt\ddot{\eta}\eta$.
  - Thus we find that the second derivative with respect to ε is: $\int_{-\frac{T}{2}}^{\frac{T}{2}} dt\eta\left(-\left(-\frac{\partial^2 V\left[x_{cl},\,\dot{x}_{cl}\right]}{\partial x^2}\eta\right)-\ddot{\eta}\right)=\int_{-\frac{T}{2}}^{\frac{T}{2}} dt\eta\left(-\frac{d^2}{dt^2}+\frac{\partial^2 V\left[x_{cl},\,\dot{x}_{cl}\right]}{\partial x^2}\right)\eta$

- All together we find that: $S\left[x_{cl}+\varepsilon\eta\right]\approx S\left[x_{cl}\right]+\frac{1}{2}\int_{-\frac{T}{2}}^{\frac{T}{2}} dt\varepsilon\eta\left(t\right)\left(-\frac{d^2}{dt^2}+\frac{\partial^2 V\left[x_{cl},\,\dot{x}_{cl}\right]}{\partial x^2}\right)\varepsilon\eta\left(t\right)+\mathcal{O}\left(\varepsilon^3\right)$
- Let $\left\{y_n\right\}_{n\in\mathbb{N}}$ be a complete set of orthonormal functions which all vanish at $\pm\frac{T}{2}$: $\int_{-\frac{T}{2}}^{\frac{T}{2}} dty_n\left(t\right)y_m\left(t\right)=\delta_{nm}$, which are eigenfunctions of the differential operator $-\frac{d^2}{dt^2}+\frac{\partial^2 V\left[x_{cl},\,\dot{x}_{cl}\right]}{\partial x^2}$. Using the boundary conditions we can find the eigenvalues:
  - Make an Ansatz with $y\left(t\right)=A\cos\left(\lambda t\right)+B\sin\left(\lambda t\right)$

- $\left(-\frac{d^2}{dt^2}+\frac{\partial^2 V[x_{cl},\,\dot{x}_{cl}]}{\partial x^2}\right)y\left(t\right)=\left(\lambda^2+\frac{\partial^2 V[x_{cl},\,\dot{x}_{cl}]}{\partial x^2}\right)y\left(t\right)$. Thus the eigenvalues are $\lambda^2+\frac{\partial^2 V[x_{cl},\,\dot{x}_{cl}]}{\partial x^2}$. But what is λ?
- Employ the boundary conditions: $\begin{cases} A\cos\left(\lambda\frac{T}{2}\right)+B\sin\left(\lambda\frac{T}{2}\right)=0\\ A\cos\left(\lambda\frac{T}{2}\right)-B\sin\left(\lambda\frac{T}{2}\right)=0\end{cases}$
- The only way to get a nontrivial solution is if $-\cos\left(\lambda\frac{T}{2}\right)\sin\left(\lambda\frac{T}{2}\right)-\sin\left(\lambda\frac{T}{2}\right)\cos\left(\lambda\frac{T}{2}\right)\overset{!}{=}0$, which means $\sin\left(\lambda T\right)\overset{!}{=}0$ which is true whenever $\lambda_n T=\pi n$ for any $n\in\mathbb{Z}$.
- Thus our eigenvalues are $\left(\frac{\pi n}{T}\right)^2+\frac{\partial^2 V[x_{cl},\,\dot{x}_{cl}]}{\partial x^2}$ for all $n\in\mathbb{N}$ (we don't need to take negatives values of n since that term is anyway squared).

- Because $\{y_n\}_{n\in\mathbb{N}}$ is a complete set, we may expand any given function using it. In particular, write $\varepsilon\eta\left(t\right):=\sum_{n\in\mathbb{N}}c_n y_n\left(t\right)$ where $\{c_n\}_{n\in\mathbb{N}}$ are the expansion coefficients.
- Plugging this into $S\left[x\right]$ we get:

$$\begin{aligned} S\left[x_{cl}+\varepsilon\eta\right] &\approx S\left[x_{cl}\right]+\frac{1}{2}\int_{-\frac{T}{2}}^{\frac{T}{2}}dt\sum_{n\in\mathbb{N}}c_n y_n\left(t\right)\left(-\frac{d^2}{dt^2}+\frac{\partial^2 V\left[x_{cl},\,\dot{x}_{cl}\right]}{\partial x^2}\right)\sum_{l\in\mathbb{N}}c_l y_l\left(t\right)\\ &= S\left[x_{cl}\right]+\frac{1}{2}\sum_{n\in\mathbb{N}}\sum_{l\in\mathbb{N}}\int_{-\frac{T}{2}}^{\frac{T}{2}}dt c_n y_n\left(t\right)\left(\left(\frac{\pi l}{T}\right)^2+\frac{\partial^2 V\left[x_{cl},\,\dot{x}_{cl}\right]}{\partial x^2}\right)c_l y_l\left(t\right)\\ &= S\left[x_{cl}\right]+\frac{1}{2}\sum_{n\in\mathbb{N}}c_n^2\left(\left(\frac{\pi n}{T}\right)^2+\frac{\partial^2 V\left[x_{cl},\,\dot{x}_{cl}\right]}{\partial x^2}\right)\end{aligned}$$

■

1.3.2 Path-Integral Approximation

1.3.2.1 Claim

$\left\langle 0,\,i\frac{T}{2}\,\middle|\,0,\,-i\frac{T}{2}\right\rangle\approx\sqrt{\frac{1}{2\pi}\frac{\omega}{sinh(\omega T)}}$ for the simple harmonic oscillator.

Proof We have taken $x_f=x_i=0$ in order to simplify the formulas.

- Going back to our path-integral, because we know there is only one unique solution to the harmonic oscillator potential, we need to approximate around only one classical path. Thus our path integral, given our approximation becomes:

$$\begin{aligned}\left\langle 0,\,i\frac{T}{2}\middle|0,\,i-\frac{T}{2}\right\rangle &\approx \mathcal{N}\int_{\{x\in\mathbb{R}^{\mathbb{R}}:\,x\left(\frac{T}{2}\right)=0\wedge x\left(-\frac{T}{2}\right)=0\}}\mathcal{D}x\,\times\\ &\times\exp\left\{-S\left[x_{cl}\right]-\frac{1}{2}\sum_{n\in\mathbb{N}}c_n^2\left(\left(\frac{\pi n}{T}\right)^2+\frac{\partial^2 V\left[x_{cl},\,\dot{x}_{cl}\right]}{\partial x^2}\right)\right\}\end{aligned}$$

- Now we can reap the fruits of the approximation to the action which we have made. Because $S[x_{cl}]$ does not depend on x, we can pull it out of the integral (in fact it is zero but we keep writing it for a while none the less). We may make a "change of variable" $x \mapsto \eta$ which is a mere "translation" in function space, so that $\mathcal{D}x = \mathcal{D}(\varepsilon\eta)$. But now, because $\varepsilon\eta$ is expanded in terms of expansion coefficients $\{c_n\}_{n\in\mathbb{N}}$, integrating over all possible η is really the same as integrating over all possible values of c_n, for each $n \in \mathbb{N}$. Thus we get now $|\mathbb{N}|$ *ordinary* integrals:

$$\begin{aligned}\left\langle 0,\, i\frac{T}{2}\middle|\, 0,\, -i\frac{T}{2}\right\rangle &\approx \underbrace{e^{-S[x_{cl}]}}_{1}\mathcal{N}\int_{\mathbb{R}} dc_0 \int_{\mathbb{R}} dc_1 \ldots \exp\left\{-\frac{1}{2}\sum_{n\in\mathbb{N}} c_n^2\left(\left(\frac{\pi n}{T}\right)^2 + \frac{\partial^2 V[x_{cl},\, \dot{x}_{cl}]}{\partial x^2}\right)\right\}\\ &= \mathcal{N}\int_{\mathbb{R}} dc_0 \int_{\mathbb{R}} dc_1 \ldots \prod_{n\in\mathbb{N}} \exp\left\{-\frac{1}{2}c_n^2\left(\left(\frac{\pi n}{T}\right)^2 + \frac{\partial^2 V[x_{cl},\, \dot{x}_{cl}]}{\partial x^2}\right)\right\}\\ &= \mathcal{N}\prod_{n\in\mathbb{N}}\left\{\int_{\mathbb{R}} dc_n \exp\left\{-\frac{1}{2}\left(\left(\frac{\pi n}{T}\right)^2 + \frac{\partial^2 V[x_{cl},\, \dot{x}_{cl}]}{\partial x^2}\right)c_n^2\right\}\right\}\end{aligned}$$

- Because we have separated our $|\mathbb{N}|$-dimensional integral into a product of $|\mathbb{N}|$ integrals, we may perform them one by one. They are easy, being ordinary Gaussian integrals which we can readily solve:

$$\begin{aligned} &= \mathcal{N}\prod_{n\in\mathbb{N}}\left\{\sqrt{2\pi}\left(\left(\frac{\pi n}{T}\right)^2 + \frac{\partial^2 V[x_{cl},\, \dot{x}_{cl}]}{\partial x^2}\right)^{-\frac{1}{2}}\right\} = \\ &= \mathcal{N}\prod_{n\in\mathbb{N}}\left\{\sqrt{2\pi}\left(\left(\frac{\pi n}{T}\right)^2\right)^{-\frac{1}{2}}\left(1 + \left(\left(\frac{\pi n}{T}\right)^2\right)^{-1}\frac{\partial^2 V[x_{cl},\, \dot{x}_{cl}]}{\partial x^2}\right)^{-\frac{1}{2}}\right\} = \\ &= \mathcal{N}\prod_{n\in\mathbb{N}}\left\{\sqrt{2\pi}\left(\left(\frac{\pi n}{T}\right)^2\right)^{-\frac{1}{2}}\right\}\prod_{n\in\mathbb{N}}\left\{\left(1 + \left(\left(\frac{\pi n}{T}\right)^2\right)^{-1}\frac{\partial^2 V[x_{cl},\, \dot{x}_{cl}]}{\partial x^2}\right)^{-\frac{1}{2}}\right\} = \\ &= \mathcal{N}\underbrace{\prod_{n\in\mathbb{N}}\left\{\sqrt{2\pi}\frac{T}{\pi n}\right\}}_{\text{free particle result}}\underbrace{\left\{\prod_{n\in\mathbb{N}}\left(1 + \frac{T^2}{\pi^2 n^2}\frac{\partial^2 V[x_{cl},\, \dot{x}_{cl}]}{\partial x^2}\right)\right\}^{-\frac{1}{2}}}_{\text{contribution from potential}}\end{aligned}$$

- Now we take advantage of the fact we haven't defined $\mathcal{N}$ yet. Observe that the first infinite product is the result we would have obtained for a free particle (had $V = 0$). But for this case, we know how to compute $\left\langle 0,\, i\frac{T}{2}\,\middle|\, 0,\, -i\frac{T}{2}\right\rangle$, and so we can calculate what $\mathcal{N}$ must be:

$$\begin{aligned}
\left\langle 0,\, i\frac{T}{2} \middle| 0,\, -i\frac{T}{2} \right\rangle_{free} &= \left\langle 0 \middle| \exp\left\{-\frac{\hat{p}^2}{2}T\right\} \middle| 0 \right\rangle \\
&= \left\langle 0 \middle| \exp\left\{-\frac{\hat{p}^2}{2}T\right\} \left(\int_{\mathbb{R}} dp\, |p\rangle \langle p|\right) \middle| 0 \right\rangle \\
&= \int_{\mathbb{R}} dp\, e^{-\frac{p^2}{2}T} \psi_p(0)\, \psi_p^*(0) \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} dp\, e^{-\frac{p^2}{2}T} \\
&= \sqrt{\frac{1}{2\pi T}}
\end{aligned}$$

- Thus we define $\mathcal{N} := \frac{\sqrt{\frac{1}{2\pi T}}}{\prod_{n\in\mathbb{N}}\left\{\sqrt{2\pi}\frac{T}{\pi n}\right\}}$ so that our results match in the case of $V = 0$.
- Another trick is the identity: $\prod_{n\in\mathbb{N}\backslash\{0\}} \left(1 + \left(\frac{\alpha}{\pi n}\right)^2\right) = \frac{sinh(\alpha)}{\alpha}$ which follows from Euler's infinite product formula for sin [2].
- Thus we find that the path integral is equal to: $\left\langle 0,\, i\frac{T}{2} \,\middle|\, 0,\, -i\frac{T}{2} \right\rangle \approx \sqrt{\frac{1}{2\pi} \frac{\omega}{sinh(\omega T)}}$ as desired, where $\omega^2 = \frac{\partial^2 V[x_{cl},\, \dot{x}_{cl}]}{\partial x^2}$.

■

1.3.3 Ground State Energy for the Harmonic Oscillator

To get an actual expression for the energy eigenvalues, we take the limit $T \to \infty$:

$$\begin{aligned}
\lim_{T\to\infty} \left\langle 0,\, i\frac{T}{2} \middle| 0,\, -i\frac{T}{2} \right\rangle &\approx \lim_{T\to\infty} \sqrt{\frac{1}{2\pi} \frac{\omega}{sinh(\omega T)}} = \\
&= \lim_{T\to\infty} \sqrt{\frac{\omega}{\pi} \frac{e^{-\omega T}}{1 - e^{-2\omega T}}} \\
&\approx \sqrt{\frac{\omega}{\pi}} \lim_{T\to\infty} e^{-\frac{\omega}{2}T} \left(1 + \frac{1}{2}\mathcal{O}\left(e^{-2\omega T}\right)\right)
\end{aligned}$$

- Comparing this result with the expression from usual quantum mechanics we find: $\boxed{\lim_{T\to\infty} e^{-E_0 T} |\psi_0(0)|^2 \approx \sqrt{\frac{\omega}{\pi}} \lim_{T\to\infty} e^{-\frac{\omega}{2}T}}$
  The sense in which this is an approximation is that higher order terms on the left hand side will give us contributions from higher states of the system (only $n \in 2\mathbb{N}$ though because for $n \in 2\mathbb{N}+1$, $\psi_n(0) = 0$).
- Thus we deduce that $\boxed{E_0 = \frac{\omega}{2}}$ and $\boxed{|\psi_0(0)|^2 = \sqrt{\frac{\omega}{\pi}}}$, which thankfully agrees with the usual canonical quantization computation, as in [9].

1.4 Conclusion

In conclusion, to find the ground state energy of a system and the ground state wave function evaluated at $x = 0$, we have found the formula: $\lim_{T\to\infty} e^{-E_0 T} \left|\psi_0(0)\right|^2 \approx e^{-S[x_{cl}]} \sqrt{\frac{\omega}{\pi}} \lim_{T\to\infty} e^{-\frac{\omega}{2}T}$ where $\omega \equiv \sqrt{\frac{\partial^2 V[x_{cl}, \dot{x}_{cl}]}{\partial x^2}}$.

Chapter 2

The Double-Well Potential

Our setup is the same as before, however now our potential is given by $V[x] = \frac{\omega^2}{8a^2}\left(x^2 - a^2\right)^2$ where $\{a,\, \omega\} \subset (0,\, \infty)$.

In order to simplify the expressions to come, define $\lambda := \frac{\omega^2}{8a^2}$ so that $V[x] = \lambda\left(x^2 - a^2\right)^2$.

This is the quartic double well potential, which we study as an archetype for tunneling phenomena in quantum mechanics. It is a prominent example for a system where employment of perturbation theory (for example, expansion in λ) will not reveal tunneling, and we must use another approach to see the effect.

2.1 Computation With Usual QM Methods

The Schroedinger equation reads (if we assume $m = 1$ and define $\hbar \equiv 1$):

$$\left[\frac{d^2}{dx^2} - 2\lambda\left(x^2 - a^2\right)^2 + 2E\right]\psi(x) = 0 \tag{2.1}$$

2.1.0.1 Claim

The two lowest energy eigenvalues of this differential equation are given by:

$$\begin{cases} E_0 = \frac{\omega}{2}\left[1 - \sqrt{\frac{2\omega^3}{\pi\lambda}}\exp\left(-\frac{\omega^3}{12\lambda}\right)\right] \\ E_1 = \frac{\omega}{2}\left[1 + \sqrt{\frac{2\omega^3}{\pi\lambda}}\exp\left(-\frac{\omega^3}{12\lambda}\right)\right] \end{cases}$$

Note: $E_0 - E_1 \sim \exp\left(-\frac{\omega^3}{12\lambda}\right)$, which cannot be expanded in perturbation series for small λ.

Proof

- The solution of this problem is presented in full detail in chapter 50 of [5] page 183 problem 3.

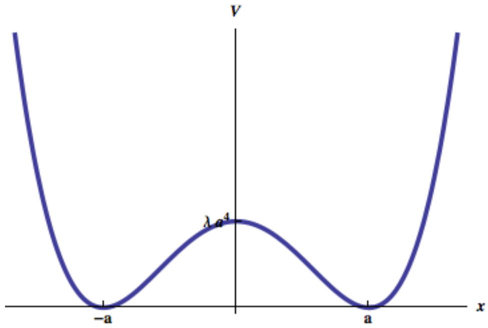

Figure 2.1: The double well.

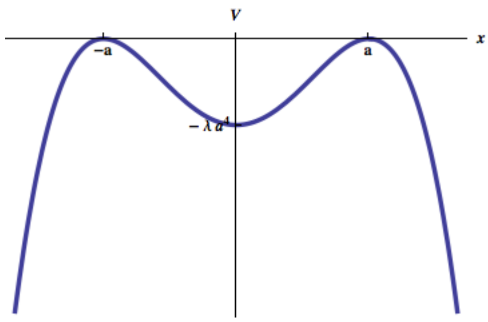

Figure 2.2: In Euclidean spacetime the potential is inverted.

2.2 Computation With Euclidean Path Integral

Our strategy for the double-well is the same as the quantum harmonic oscillator: find the classical paths (which minimize the action), and expand the action around them. Finally plug this expansion into the path-integral. This means that classical paths whose action is infinite will have no contribution to the path integral. We shall now try to compute $\left\langle -a \middle| e^{-T\hat{H}} \middle| a \right\rangle$ (from here on referred to as $BC1$) and $\left\langle a \middle| e^{-T\hat{H}} \middle| -a \right\rangle$ (from here on referred to as $BC2$), two transitions which we associate with quantum tunneling.

2.2.1 The Classical Euclidean Paths of Finite Action

Because we are working in Euclidean spacetime, our potential is inverted and we are looking for solutions for the following equation of motion:

$$\boxed{\ddot{x}_{cl}\left(t\right) = 4\lambda\left[x_{cl}\left(t\right)^2 - a^2\right]x_{cl}\left(t\right)} \tag{2.2}$$

2.2.1.1 Claim

The *only* solutions to this equation are $\boxed{a \cdot \tanh\left(\pm\frac{\omega}{2}\left(t-\mathcal{T}\right)\right)}$ for all $\mathcal{T} \in \mathbb{R}$, where the $+$ variant corresponds to $BC1$ and the $-$ variant corresponds to $BC2$.

Proof

- Multiply the equation by $2\dot{x}_{cl}(t)$ to obtain a differential equation which we integrate to get an integration constant C_1:

$$\begin{aligned} 2\dot{x}_{cl}(t)\cdot\ddot{x}_{cl}(t) &= 2\dot{x}_{cl}(t)\cdot 4\lambda\left(x_{cl}(t)^2-a^2\right)x_{cl}(t) \\ \frac{d}{dt}\left[\dot{x}_{cl}(t)\right]^2 &= 2\lambda\frac{d}{dt}\left[x_{cl}(t)^2-a^2\right]^2 \\ \left[\dot{x}_{cl}(t)\right]^2 &= 2\lambda\left[x_{cl}(t)^2-a^2\right]^2+C_1 \end{aligned}$$

- Determine C_1 using the boundary conditions (by the way, C_1 is the energy): At $\pm\frac{T}{2}$, we require that $x_{cl}=\pm a$ and that $\dot{x}_{cl}=0$ (asymptotically go to $\pm a$) and so C_1 *must* be 0.

- Write $x_{cl}(t) := a\tanh[u(t)]$ and plug it into the equation to get:

$$\begin{aligned} \left\{a\,sech\,[u(t)]^2\,\dot{u}(t)\right\}^2 &= 2\lambda\left\{a^2\tanh[u(t)]^2-a^2\right\}^2 \\ a^2\left\{1-\tanh[u(t)]^2\right\}^2\left[\dot{u}(t)\right]^2 &= 2\lambda a^4\left\{\tanh[u(t)]^2-1\right\}^2 \\ \left[\dot{u}(t)\right]^2 &= 2\lambda a^2 \\ \dot{u}(t) &= \pm a\sqrt{2\lambda}=\pm a\sqrt{2\frac{\omega^2}{8a^2}}=\pm\frac{\omega}{2} \\ u(t) &= \pm\frac{\omega}{2}t+C_2 \end{aligned}$$

  where C_2 is some integration constant.

- Define $\mathcal{T} := \frac{2}{\omega}C_2$

- Thus we find that the most general solution is: $a\tanh\left(\pm\frac{\omega}{2}t+\frac{\omega}{2}\mathcal{T}\right)$ for any $\mathcal{T}\in\mathbb{R}$.

■

2.2.1.2 Remarks

- These solutions' boundary conditions are:
  - For + version, $x_{cl}(-\infty)=-a$ and $x_{cl}(\infty)=a$. This solution is called an *instanton* at time $\mathcal{T}$ and shall be denoted from here until the end of this chapter as $I_{\mathcal{T}}(t)$.
  - For the $-$ version $x_{cl}(-\infty)=a$ and $x_{cl}(\infty)=-a$. This solution is called an *anti-instanton* at time $\mathcal{T}$ and shall be denoted $A_{\mathcal{T}}(t)$.

- Thus they are only *approximate* solutions if our boundary conditions are $x_{cl}\left(\pm\frac{T}{2}\right)=\pm a$, and become precise solutions only when $T\to\infty$.

- These solutions, which correspond to tunneling are only made possible by the fact we are working in the Euclidean spacetime framework. In Minkowski spacetime $\nexists$ such solutions, because classically $\nexists$ tunneling!

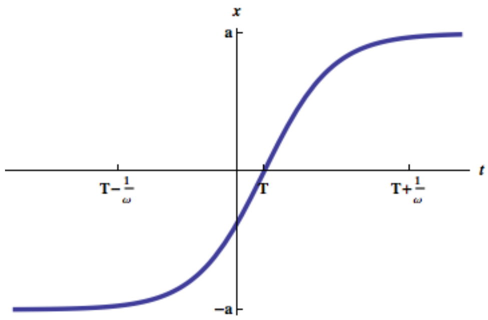

Figure 2.3: An instanton at $\mathcal{T}$.

- Observe that $\lim_{t\to\infty}\left(I_{\mathcal{T}}(t)-a\right)=\lim_{t\to\infty}-2ae^{-\omega(t-\mathcal{T})}$ and so we can imagine the "width" of an (anti) instanton in time is proportional $\propto\frac{1}{\omega}$. In other words, it happens within a time window of width about $\frac{1}{\omega}$ centered at $\mathcal{T}$, where before and after nothing happens.

2.2.1.3 Claim

Let $n\in 2\mathbb{N}+1$. Let $\mathcal{T}_1\in\left(-\frac{T}{2},\,\frac{T}{2}\right)$, $\mathcal{T}_2\in\left(\mathcal{T}_1,\,\frac{T}{2}\right)$, $\dots$, $\mathcal{T}_n\in\left(\mathcal{T}_{n-1},\,\frac{T}{2}\right)$. Then $\boxed{a\prod_{j=1}^{n}\tanh\left[\pm\frac{1}{2}\omega\left(t-\mathcal{T}_j\right)\right]}$ are *approximate* solutions to the equation of motion. The $+$ $(-)$ variant corresponding to $BC1$ $(BC2)$.

Proof

- Work on the $+$ version first.
- For brevity denote $\tau_j:=\frac{1}{2}\omega\left(t-\mathcal{T}_j\right)$.
- Denote our suggested approximate solution by $y(t)=a\prod_{j=1}^{n}\tanh\left(\tau_j\right)$.
- Then $\dot{y}(t)=a\sum_{l=1}^{n}\left[\prod_{j\neq l}\tanh\left(\tau_j\right)\frac{1}{2}\omega\left(sech\left(\tau_l\right)\right)^2\right]$
- and

$$\begin{aligned}\ddot{y}(t) &= \frac{1}{2}\omega\,a\sum_{l=1}^{n}\left[-\omega\prod_{j}\tanh\left(\tau_j\right)\left(sech\left(\tau_l\right)\right)^2+\sum_{k\neq l}\left(\prod_{j\neq l,\,j\neq k}\tanh\left(\tau_j\right)\right)\left(sech\left(\tau_k\right)\right)^2\right]\\ &= \frac{1}{2}\omega\,a\sum_{l=1}^{n}\left[-\omega\frac{1}{a}y(t)\left(sech\left(\tau_l\right)\right)^2+\sum_{k\neq l}\left(\prod_{j\neq l,\,j\neq k}\tanh\left(\tau_j\right)\right)\left(sech\left(\tau_k\right)\right)^2\right]\end{aligned}$$

- Plugging this into the equation of motion we have:

$$
\begin{aligned}
0 &\stackrel{?}{=} \ddot{y}(t) - y(t)\,4\lambda\left(y(t)^2 - a^2\right) \\
&= \frac{1}{2}\omega\, a \sum_{l=1}^{n}\left[-\omega\frac{1}{a}y(t)\,(sech(\tau_l))^2 + \sum_{k\neq l}\left(\prod_{j\neq l,\, j\neq k}\tanh(\tau_j)\right)(sech(\tau_k))^2\right] - \\
&\quad -\frac{1}{2}\omega^2 y(t)\left(\left(\prod_j \tanh(\tau_j)\right)^2 - 1\right)
\end{aligned}
$$

1. First case, $t \in (\mathcal{T}_{j_0} - \varepsilon,\, \mathcal{T}_{j_0} + \varepsilon,)$ for some $j_0 \in \{1,\, \ldots,\, n\}$ and for some $\varepsilon \ll 1$. For such values of t, $y(t)$ is almost zero. In addition, $sech(\tau_j) \approx 0 \forall j \neq j_0$ and 1 otherwise and so we have (neglecting already the terms with $y(t)$):

$$
\begin{aligned}
&\approx \frac{1}{2}\omega\, a \sum_{l=1}^{n}\left[\sum_{k\neq l}\left(\prod_{j\neq l,\, j\neq k}\tanh(\tau_j)\right)(sech(\tau_k))^2\right] \\
&\approx \frac{1}{2}\omega\, a \sum_{l=1}^{n}\left[\sum_{k\neq l}\left(\prod_{j\neq l,\, j\neq k}\tanh(\tau_j)\right)\delta_{k,\, j_0}\right] \\
&= \frac{1}{2}\omega a \sum_{l=1}^{n}\begin{cases} 0 & l = j_0 \\ \left(\prod_{j\neq l,\, j\neq j_0}\tanh(\tau_j)\right) & l \neq j_0 \end{cases} \\
&= \frac{1}{2}\omega a \sum_{l\neq j_0}\left(\prod_{j\neq l,\, j\neq j_0}\tanh(\tau_j)\right) \\
&\approx \underbrace{0}_{\text{sum of 1 and -1 even number of times}}
\end{aligned}
$$

2. Second case, when $|t - \mathcal{T}_j| \gg \frac{1}{\omega}\ \forall j \in \{1,\, \ldots,\, n\}$, that is, it is "very" far away from any instanton-event at any of the $\mathcal{T}_j$'s. Then $sech(\tau_j) \approx 0 \forall j$, $|\tanh(\tau_j)| \approx 1 \forall j$ and so our equation is fulfilled.

■

We call the solution with the plus "n instantons", denoted by $I^n_{\mathcal{T}_1,\, \ldots,\, \mathcal{T}_n}(t)$. Observe that it obeys exactly the same boundary conditions as the single instanton. The solution with the minus is called "n anti-instantons", denoted by $A^n_{\mathcal{T}_1,\, \ldots,\, \mathcal{T}_n}(t)$, and obeys the same boundary conditions as the single anti-instanton.

2.2.2 Features of Instanton Solutions

2.2.2.1 Claim

$E[I_{\mathcal{T}}] \equiv \frac{1}{2}\dot{I_{\mathcal{T}}}(t)^2 - V[I_{\mathcal{T}}] = 0$ and the same for $A_{\mathcal{T}}$. Thus we would see that the instantons and anti-instantons have zero total (Euclidean) energy.

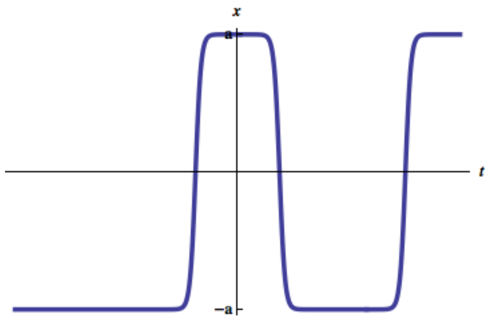

Figure 2.4: A 3-instanton.

Proof $\frac{1}{2}a^2\left(\frac{\omega}{2}\right)^2\left[sech\left(\pm\frac{\omega}{2}(t-\mathcal{T})\right)\right]^4-\frac{\omega^2}{8a^2}\left(-a^2\left[sech\left(\pm\frac{\omega}{2}(t-\mathcal{T})\right)\right]^2\right)^2=0.$
■

2.2.2.2 Claim

$S[I_{\mathcal{T}}]=S[A_{\mathcal{T}}]=\frac{\omega^3}{12\lambda}$ (in particular, the action of instantons and anti-instantons is independent of their parameter $\mathcal{T}$!)

Proof

- Using the preceding claim, we can simplify the computation of the action:

$$\begin{aligned}
S[I_{\mathcal{T}}] &= \int_{-\infty}^{\infty} dt\left(\frac{1}{2}\dot{I_{\mathcal{T}}}(t)^2+V[I_{\mathcal{T}}]\right)\\
&= \int_{-\infty}^{\infty} dt\dot{I_{\mathcal{T}}}(t)^2\\
&= =a^2\left(\frac{\omega}{2}\right)^2\underbrace{\int_{-\infty}^{\infty} dt\left[sech\left(\pm\frac{\omega}{2}(t-\mathcal{T})\right)\right]^4}_{\frac{4}{3\left(\frac{\omega}{2}\right)}}\\
&= \frac{2}{3}a^2\omega\\
&= \frac{\omega^3}{12\lambda}
\end{aligned}$$

- Because the result of the integral is independent of the plus or minus sign, this computation holds also for $A_{\mathcal{T}}$.
■

Define $\boxed{S_0:=\frac{\omega^3}{12\lambda}}$.

2.2.2.3 Claim

$S\left[I^n_{\mathcal{T}_1,\ldots,\mathcal{T}_n}(t)\right]=S\left[A^n_{\mathcal{T}_1,\ldots,\mathcal{T}_n}(t)\right]=n\,S_0$

Proof Using the fact that instantons have width $\frac{1}{\omega}$ we may separate the integral: $S\left[I^n_{\mathcal{T}_1,\ldots,\mathcal{T}_n}(t)\right] = \int_{-\infty}^{\infty} dt\left(\frac{1}{2}\left(\frac{d}{dt}I^n_{\mathcal{T}_1,\ldots,\mathcal{T}_n}(t)\right)^2 + V\left[I^n_{\mathcal{T}_1,\ldots,\mathcal{T}_n}(t)\right]\right) = \sum_{j=1}^n \int_{t_j-\varepsilon}^{t_j+\varepsilon}\left(\frac{1}{2}\dot{I_{\mathcal{T}_j}}(t)^2 + V\left[I_{\mathcal{T}_j}\right]\right) = nS_0$, where ε is chosen such that $\frac{1}{\omega} < \varepsilon < min\left(\{|\mathcal{T}_j - \mathcal{T}_k| : \{j, k\} \subset \{1, \ldots, n\}\}\right)$. Later we would see that this should always be possible for such n that we care about.
■

2.2.3 Structure of Transition Amplitude Approximation

We find that the most general classical *approximate* solution to the equation of motion with either $BC1$ or $BC2$ is indexed by some *odd* integer n, together with n consecutive numbers in the interval $\left(-\frac{T}{2}, \frac{T}{2}\right) \subset \mathbb{R}$.

As we remarked in the first chapter, if we want to make an approximation for the path integral around stationary paths, and if $\exists$ more than one stationary path, then in general we could approximate the path integral as a sum of approximations around the various stationary paths.

Thus we expect

$$\left\langle -a \left| e^{-T\hat{H}} \right| a \right\rangle \tag{2.3}$$
$$\approx \sum_{n\in 2\mathbb{N}+1} \int_{-T/2}^{T/2} d\mathcal{T}_1 \ldots \int_{\mathcal{T}_{n-1}}^{T/2} d\mathcal{T}_n \left[\text{path integral approximation around } I^n_{\mathcal{T}_1,\ldots,\mathcal{T}_n}(t)\right]$$

and the anti-instanton approximation for $\left\langle a \left| e^{-T\hat{H}} \right| -a \right\rangle$.

Furthermore, by similar procedures it is clear that an even number of instantons or anti-instantons obey the boundary conditions of $x_{cl}\left(\pm\frac{T}{2}\right) = \pm a$ respectively, and thus, they form an *approximate* solution with these boundary conditions, and it is also clear that $\pm a$ (the constant map sitting always at either a or $-a$) is an *exact* solution to the classical equations of motion with zero action. We will use these solutions to compute $\left\langle a \left| e^{-\hat{H}T} \right| a \right\rangle$ or $\left\langle -a \left| e^{-\hat{H}T} \right| -a \right\rangle$ respectively.

2.3 Single Instanton Contributions

Our next step would be to compute the contribution of a *single* instanton (of some given $\mathcal{T}$) to $\left\langle -a \left| e^{-T\hat{H}} \right| a \right\rangle$ for which we expand the action around $I_{\mathcal{T}}$. (Or vice versa of an anti-instanton to $\left\langle a \left| e^{-T\hat{H}} \right| -a \right\rangle$). Using our experience with the harmonic oscillator, we can readily write down that contribution:

$$\begin{aligned} \left\langle -a \left| e^{-T\hat{H}} \right| a \right\rangle_{n=1} &\approx e^{-S[I_{\mathcal{T}}]}\mathcal{N}\prod\left\{\sqrt{2\pi}\varepsilon_n^{-\frac{1}{2}}\right\} \\ &= e^{-S_0}\mathcal{N}'\left\{\prod \varepsilon_n\right\}^{-\frac{1}{2}} \end{aligned}$$

where ε_n are the eigenvalues of the operator $-\frac{d^2}{dt^2} + \frac{\partial^2 V\left[I_{\mathcal{T}}, \dot{I}_{\mathcal{T}}\right]}{\partial x^2}$.

Next, to make the following expressions shorter, define det (operator) := $\prod$ (eigenvalues of operator). This definition is not without sense, because as we've seen in the previous chapter, the product of the eigenvalues represents the contribution to the transition amplitude of quadratic quantum fluctuations around the classical path.

Then we can rewrite the contribution as:

$$\left\langle -a \left| e^{-T\hat{H}} \right| a \right\rangle_{n=1} = e^{-S_0}\mathcal{N}' \left\{ \det\left(-\frac{d^2}{dt^2} + \frac{\partial^2 V\left[I_{\mathcal{T}}, \dot{I}_{\mathcal{T}}\right]}{\partial x^2} \right) \right\}^{-\frac{1}{2}}$$

For convenience, we take the harmonic oscillator, $\left\{\det\left(-\frac{d^2}{dt^2}+\omega^2\right)\right\}^{-\frac{1}{2}}$, as a reference for our computations:

$$\begin{aligned} \left\langle -a \left| e^{-T\hat{H}} \right| a \right\rangle_{n=1} &= e^{-S_0}\mathcal{N}' \left\{ \det\left(-\frac{d^2}{dt^2} + \omega^2 \right) \right\}^{-\frac{1}{2}} \left\{ \frac{\det\left(-\frac{d^2}{dt^2} + \frac{\partial^2 V\left[I_{\mathcal{T}}, \dot{I}_{\mathcal{T}}\right]}{\partial x^2} \right)}{\det\left(-\frac{d^2}{dt^2}+\omega^2\right)} \right\}^{-\frac{1}{2}} \\ &= e^{-S_0}\sqrt{\frac{1}{2\pi}\frac{\omega}{sinh\left(\omega T\right)}} \left\{ \frac{\det\left(-\frac{d^2}{dt^2} + \frac{\partial^2 V\left[I_{\mathcal{T}}, \dot{I}_{\mathcal{T}}\right]}{\partial x^2} \right)}{\det\left(-\frac{d^2}{dt^2}+\omega^2\right)} \right\}^{-\frac{1}{2}} \end{aligned}$$

Next, a simple calculation shows that

$$\begin{aligned} \frac{\partial^2 V\left[I_{\mathcal{T}}, \dot{I}_{\mathcal{T}}\right]}{\partial x^2} &= 4\lambda\left(3x^2 - a^2\right)\Big|_{x=I_{\mathcal{T}}} \\ &= \omega^2 - \frac{3}{2}\frac{\omega^2}{\left\{cosh\left[\frac{1}{2}\omega\left(t-\mathcal{T}\right)\right]\right\}^2} \end{aligned}$$

so we obtain finally:

$$\left\langle -a \left| e^{-T\hat{H}} \right| a \right\rangle_{n=1} = e^{-S_0}\sqrt{\frac{1}{2\pi}\frac{\omega}{sinh\left(\omega T\right)}} \left\{ \frac{\det\left(-\frac{d^2}{dt^2} + \omega^2 - \frac{3}{2}\frac{\omega^2}{\left\{cosh\left[\frac{1}{2}\omega\left(t-\mathcal{T}\right)\right]\right\}^2} \right)}{\det\left(-\frac{d^2}{dt^2}+\omega^2\right)} \right\}^{-\frac{1}{2}}$$

where we have used our computation from the first chapter for the harmonic oscillator operator determinant, and "all" that is left for us is to compute $\frac{\det\left(-\frac{d^2}{dt^2} + \omega^2 - \frac{3}{2}\frac{\omega^2}{\left\{cosh\left[\frac{1}{2}\omega\left(t-\mathcal{T}\right)\right]\right\}^2} \right)}{\det\left(-\frac{d^2}{dt^2}+\omega^2\right)}$.

Thus the eigenvalue equation reads: $\boxed{\left(-\frac{d^2}{dt^2} + \omega^2 - \frac{3}{2}\frac{\omega^2}{\left\{cosh\left[\frac{1}{2}\omega\left(t-\mathcal{T}\right)\right]\right\}^2} \right) y_n\left(t\right) = \varepsilon_n y_n\left(t\right)}$.

It is clear that as $\omega^2 - \varepsilon_n > 0$ we will have a discrete set of eigenvalues and when $\omega^2 - \varepsilon_n < 0$ there will be a continuous spectrum. However, with our boundary conditions $y_n\left(\pm\frac{T}{2}\right) = 0$ the eigenvalues which are bigger than ω^2 will *also* be discrete, and only when taking the limit $T \to \infty$ they will "become" continuous.

2.3.0.1 Claim

$\left\{0,\,\frac{3}{4}\omega^2\right\}\subset\{\varepsilon_n\}_{n\in\mathbb{N}}$, where $\{\varepsilon_n\}_{n\in\mathbb{N}}$ are all the eigenvalues of $-\frac{d^2}{dt^2}+\omega^2-\frac{3}{2}\frac{\omega^2}{\{cosh[\frac{1}{2}\omega(t-\mathcal{T})]\}^2}$. In particular, 0 is the lowest eigenvalue of $-\frac{d^2}{dt^2}+\omega^2-\frac{3}{2}\frac{\omega^2}{\{cosh[\frac{1}{2}\omega(t-\mathcal{T})]\}^2}$.

Proof

- We follow [5] pp. 73 problem 5:
  - Rearrange the equation as $\left(\frac{d^2}{dt^2}+\varepsilon_n-\omega^2+\frac{\frac{3}{2}\omega^2}{\{cosh[\frac{1}{2}\omega(t-\mathcal{T})]\}^2}\right)y_n(t)=0$
  - Make a change of variables $t-\mathcal{T}\mapsto t$ to get:
  - $\left(\frac{d^2}{dt^2}+\varepsilon_n-\omega^2+\frac{\omega^2\frac{3}{2}}{\{cosh[\frac{1}{2}\omega t]\}^2}\right)y_n(t)=0$
  - Define $\xi(t):=tanh\left(\frac{1}{2}\omega t\right)$, $\epsilon:=\frac{2}{\omega}\sqrt{\omega^2-\varepsilon_n}$, $s:=2$
  - Then $\frac{dy}{dt}=\frac{d\xi}{dt}\frac{dy}{d\xi}$ and so $\frac{d^2y}{dt^2}=\frac{d}{dt}\left(\frac{d\xi}{dt}\frac{dy}{d\xi}\right)=\frac{d^2\xi}{dt^2}\frac{dy}{d\xi}+\left(\frac{d\xi}{dt}\right)^2\frac{d^2y}{d\xi^2}$
  - $\frac{d\xi}{dt}=\frac{1}{2}\omega\left[sech\left(\frac{1}{2}\omega t\right)\right]^2=\frac{1}{2}\omega\left(1-\xi^2\right)$ and $\frac{d^2\xi}{dt^2}=-\frac{1}{2}\omega^2\left(1-\xi^2\right)\xi$
  - So that $\frac{d^2y}{dt^2}=-\frac{1}{2}\omega^2\left(1-\xi^2\right)\xi\frac{dy}{d\xi}+\frac{1}{4}\omega^2\left(1-\xi^2\right)^2\frac{d^2y}{d\xi^2}=\frac{1}{2}\omega^2\left[-\left(1-\xi^2\right)\xi\frac{dy}{d\xi}+\frac{1}{2}\left(1-\xi^2\right)^2\frac{d^2y}{d\xi^2}\right]=$ $=\frac{1}{2}\omega^2\frac{d}{d\xi}\left[\left(1-\xi^2\right)\frac{dy}{d\xi}\right]$
  - $\epsilon=2-n$ and so $\frac{2}{\omega}\sqrt{\omega^2-\varepsilon_n}=2-n$ and so $\omega^2-\varepsilon_n=\frac{\omega^2}{4}\left(4-4n+n^2\right)=\omega^2-\omega^2n+\frac{1}{4}\omega^2n^2$ and so $\boxed{\varepsilon_n=\omega^2n\left(1-\frac{1}{4}n\right)}$ for $n\in\mathbb{N}$. But ϵ must be positive! so we obtain only two eigenvalues in this way, and the rest will be obtained differently.

    ■

2.3.1 Zero Modes–Collective Coordinates

We are in trouble, because the first eigenvalue is 0, and we have a term $(\varepsilon_0)^{-\frac{1}{2}}=\frac{1}{0}$ in our transition amplitude.

We identify the $y_0(t)$ eigenvector–the eigenvector corresponding to eigenvalue zero–as a "direction" in function space that leaves the action invariant. This is because $\left(-\frac{d^2}{dt^2}+\frac{\partial^2V\left[I_\mathcal{T},\dot{I}_\mathcal{T}\right]}{\partial x^2}\right)$ is actually $\frac{\delta^2S}{\delta\eta^2}$ where η is a quantum variation around the classical path $I_\mathcal{T}$. But we know of such a "direction" already: varying $I_\mathcal{T}\mapsto I_{\mathcal{T}+\Delta\mathcal{T}}$ leaves the action invariant. Thus y_0 must correspond to this shift in $\mathcal{T}$. Thus if $I_\mathcal{T}(t)+\Delta\mathcal{T}y_0(t)\propto I_{\mathcal{T}+\Delta\mathcal{T}}(t)$ then $y_0(t)\propto\frac{I_{\mathcal{T}+\Delta\mathcal{T}}(t)-I_\mathcal{T}(t)}{\Delta\mathcal{T}}\overset{\Delta\mathcal{T}\to0}{\to}-\frac{d}{d\mathcal{T}}I_\mathcal{T}(t)=\frac{d}{dt}I_\mathcal{T}(t)$. We must still normalize this vector to be able to use it: $\int_{-\infty}^{\infty}dty_0(t)^2\overset{!}{=}1$, but this computation we have already made, and found $\int_{-\infty}^{\infty}dt\left(\frac{d}{dt}I_\mathcal{T}(t)\right)^2=S_0$. Thus we find: $\boxed{y_0(t)=(S_0)^{-\frac{1}{2}}\frac{d}{dt}I_\mathcal{T}(t)}$.

So in the path integral which we separated in chapter one into an $|\mathbb{N}|$ dimensional integral over coefficients corresponding to eigenvectors, we must separate the zero mode, because it is in fact not Gaussian $\int_{\mathbb{R}} dc_0 e^0$. It is clear that c_0 actually corresponds to $\mathcal{T}$ and so we can swap this integral with an integral over $\mathcal{T}$. This is conventionally called the introduction of a *collective coordinate.* This fits well with our scheme as we already anticipated integration of the various $\mathcal{T}$ parameters.

2.3.1.1 Claim

$\int dc_0 = \sqrt{S_0} \int d\mathcal{T}$

Proof

- If c_0 changes by Δc_0, our path (the integration variable in the path integral) changes by $\Delta x(t) = y_0(t) \Delta c_0$.
- On the other hand, if $\mathcal{T}$ changes by $\Delta\mathcal{T}$, our path changes by $\Delta x(t) = \Delta I_{\mathcal{T}}(t) = \frac{\partial I_{\mathcal{T}}(t)}{\partial \mathcal{T}} \Delta\mathcal{T} = -\sqrt{S_0} y_0(t) \Delta\mathcal{T}$
- Because the change to $x(t)$ must be the same, we conclude that $dc_0 = \sqrt{S_0} d\mathcal{T}$. (The minus sign is irrelevant).

■

2.3.2 The Remaining Eigenvalues

- Define $\det_0$ (operator) := $\prod$ (eigenvalues of operator except the zero one).

Going back to our product of eigenvalues, we thus omit the zero eigenvalue and replace it with "preparation" for integration over the zero mode:

$$\left\{ \frac{det\left(-\frac{d^2}{dt^2} + \frac{\partial^2 V\left[I_{\mathcal{T}}, \dot{I}_{\mathcal{T}}\right]}{\partial x^2}\right)}{\det\left(-\frac{d^2}{dt^2} + \omega^2\right)} \right\}^{-\frac{1}{2}} = \sqrt{\frac{S_0}{2\pi}} d\mathcal{T} \left\{ \frac{det_0\left(-\frac{d^2}{dt^2} + \frac{\partial^2 V\left[I_{\mathcal{T}}, \dot{I}_{\mathcal{T}}\right]}{\partial x^2}\right)}{\det\left(-\frac{d^2}{dt^2} + \omega^2\right)} \right\}^{-\frac{1}{2}}$$

For normalization purposes we must also multiply by $\frac{1}{\sqrt{2\pi}}$ which is what we would have obtained from the Gaussian integral of the zero mode. Recall from chapter one that the eigenvalues of $-\frac{d^2}{dt^2} + \omega^2$ were $\left(\frac{\pi n}{T}\right)^2 + \omega^2 \overset{T\to\infty}{\longrightarrow} \omega^2$. So for every term we don't include in the product we should "compensate" by dividing by ω^2 so that all together we have:

$$\sqrt{\frac{S_0}{2\pi}} \omega d\mathcal{T} \left\{ \frac{det_0\left(-\frac{d^2}{dt^2} + \frac{\partial^2 V\left[I_{\mathcal{T}}, \dot{I}_{\mathcal{T}}\right]}{\partial x^2}\right)}{\omega^{-2}\det\left(-\frac{d^2}{dt^2} + \omega^2\right)} \right\}^{-\frac{1}{2}}$$

2.3.2.1 Claim

$$\frac{det_0\left(-\frac{d^2}{dt^2}+\frac{\partial^2 V\left[I_{\mathcal{T}},\, \dot{I}_{\mathcal{T}}\right]}{\partial x^2}\right)}{\omega^{-2}\det\left(-\frac{d^2}{dt^2}+\omega^2\right)} = \frac{3}{4}\times\frac{1}{9}$$

Proof

- The next eigenvalue after 0 is $\frac{3}{4}\omega^2$. But since we are also dividing by the harmonic oscillator eigenvalues in the limit $T\to\infty$, we must divide each eigenvalue by ω^2. Thus we get for the first nonzero eigenvalue a contribution of $\frac{3}{4}$.
- The $\frac{1}{9}$ factor is computed in [5] pp. 80 and also applied to our particular problem in [12].

■

Notes So all together we have $\boxed{\lim_{T\to\infty}\left\langle -a\,\middle|\, e^{-T\hat{H}}\,\middle|\, a\right\rangle_{\text{one-instanton}} = \lim_{T\to\infty}\left(\sqrt{\frac{\omega}{\pi}}e^{-\frac{\omega}{2}T}\right)D\omega d\mathcal{T}}$

where we define $\boxed{D := \sqrt{\frac{6}{\pi}}\sqrt{S_0}e^{-S_0}}$ to make the expressions shorter.

2.4 Dilute Instanton Gas

2.4.1 Energy Eigenvalues

Because we know that a correction of one-instanton to the harmonic oscillator entails a factor of $D\omega d\mathcal{T}$ to the transition amplitude, we may readily generalize that the contribution of n instantons is $(D\omega d\mathcal{T})^n$.

2.4.1.1 Claim

The contribution of an n-instanton to the transition amplitude is $(D\omega)^n\, d\mathcal{T}_1 \dots d\mathcal{T}_n$.

Proof We only need to explain the $\sqrt{\frac{6}{\pi}}^n$ term, because the power in the exponent follows from the fact that the action of the multi-instanton solution is merely n times the action of the single instanton. The other factors are byproducts of the collective coordinates, and so only $\sqrt{\frac{6}{\pi}}^n$ actually follows from computing the functional determinant. We leave this as an exercise to the reader.

■

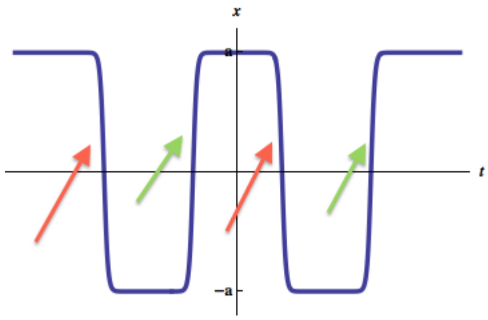

Figure 2.5: A 4-instanton. The green arrows point to instantons and red arrows point to anti-instantons in the sequence.

2.4.1.2 Conclusion

So we may finally write:

$$\begin{aligned}
\lim_{T\to\infty}\left\langle -a\left| e^{-T\hat{H}}\right| a\right\rangle_{n=1} &= \lim_{T\to\infty}\left(\sqrt{\frac{\omega}{\pi}}e^{-\frac{\omega}{2}T}\right)\sum_{n\in 2\mathbb{N}+1}\int_{-T/2}^{T/2}d\mathcal{T}_1\ldots\int_{\mathcal{T}_{n-1}}^{T/2}d\mathcal{T}_n\left[D\omega\right]^n \\
&= \lim_{T\to\infty}\left(\sqrt{\frac{\omega}{\pi}}e^{-\frac{\omega}{2}T}\right)\sum_{n\in 2\mathbb{N}+1}\left(D\omega\right)^n\frac{T^n}{n!} \\
&= \lim_{T\to\infty}\left(\sqrt{\frac{\omega}{\pi}}e^{-\frac{\omega}{2}T}\right) sinh\left[D\omega T\right] \\
&= \lim_{T\to\infty}\left(\sqrt{\frac{\omega}{\pi}}e^{-\frac{\omega}{2}T}\right)\frac{1}{2}\left[e^{D\omega T}-e^{-D\omega T}\right] \\
&= \lim_{T\to\infty}\frac{1}{2}\sqrt{\frac{\omega}{\pi}}\left[e^{-\frac{\omega}{2}T+D\omega T}-e^{-\frac{\omega}{2}T-D\omega T}\right] \\
&= \lim_{T\to\infty}\frac{1}{2}\sqrt{\frac{\omega}{\pi}}\left[e^{-\frac{\omega}{2}(1-2D)T}-e^{-\frac{\omega}{2}(1+2D)T}\right] \\
&= \lim_{T\to\infty}\frac{1}{2}\sqrt{\frac{\omega}{\pi}}\left[e^{-\frac{\omega}{2}\left(1-\sqrt{\frac{2\omega^3}{\pi\lambda}}e^{-\frac{\omega^3}{12\lambda}}\right)T}-e^{-\frac{\omega}{2}\left(1+\sqrt{\frac{2\omega^3}{\pi\lambda}}e^{-\frac{\omega^3}{12\lambda}}\right)T}\right]
\end{aligned}$$

And so we find exactly the same two lowest energy eigenvalues as the ordinary quantum mechanics techniques.

2.4.2 Symmetry of Ground State

Using exactly the same procedure, we can evaluate $\left\langle a\left| e^{-\hat{H}T}\right| a\right\rangle$. Now we have one classical solution which is the solution $x(t)=a$. But as above, we also must take into account multi-instanton approximate solutions, which take us back and forth. However, in contrast to before, now we need an *even* numbered multi-instanton.

$$\lim_{T\to\infty}\left\langle a\left|e^{-\hat{H}T}\right|a\right\rangle \approx \lim_{T\to\infty}\left(\sqrt{\frac{\omega}{\pi}}e^{-\frac{\omega}{2}T}\right)\sum_{n\in2\mathbb{N}}(D\omega)^n\frac{T^n}{n!}$$
$$= \lim_{T\to\infty}\left(\sqrt{\frac{\omega}{\pi}}e^{-\frac{\omega}{2}T}\right)cosh\left[D\omega T\right]$$
$$= \lim_{T\to\infty}\frac{1}{2}\sqrt{\frac{\omega}{\pi}}\left[e^{-\frac{\omega}{2}\left(1-\sqrt{\frac{2\omega^3}{\pi\lambda}}e^{-\frac{\omega^3}{12\lambda}}\right)T}+e^{-\frac{\omega}{2}\left(1+\sqrt{\frac{2\omega^3}{\pi\lambda}}e^{-\frac{\omega^3}{12\lambda}}\right)T}\right]$$

2.4.3 Wave Functions

From our analysis it is also possible to extract the energy eigenfunctions: $\psi_0(-a)\,\psi_0^*(a) = \psi_0(a)\,\psi_0^*(a) = \frac{1}{2}\sqrt{\frac{\omega}{\pi}} \Longrightarrow \boxed{\psi_0(a) = \psi_0(-a) = \left(\frac{\omega}{4\pi}\right)^{\frac{1}{4}}}$ and so it is appears that the wave function for the lowest state remains symmetric after all under the exchange of $\pm a$.

2.4.4 Validity of Dilute Instanton Approximation

Even though we are seemingly summing over an *arbitrarily large* number of instantons (and so the dilute gas approximation should break down at some point), in infinite sums of the form $\sum_{n\in\mathbb{N}}\frac{x^n}{n!}$, only the terms for which $n < x$ are non-negligible. That means for us only terms where $n < D\omega T$ are actually important in the infinite sum, where we recall $D \equiv \sqrt{\frac{6}{\pi}}\sqrt{S_0}e^{-S_0}$ from before.

For the dilute gas approximation to be valid, we have assumed that instantons don't "interact", that is, that one instanton event is finished well before another one starts: $|\mathcal{T}_i - \mathcal{T}_j| \gg \frac{1}{\omega}$. So we must make sure this condition holds, at least for all solutions with $n < D\omega T$. On average, we have $|\mathcal{T}_i - \mathcal{T}_j| = \frac{T}{n}$, and so, we only care about the "worst" case in which $|\mathcal{T}_i - \mathcal{T}_j| = \frac{T}{D\omega T} = \frac{1}{D}\times\frac{1}{\omega}$. So we find that in order for this approximation to hold we need $D \ll 1$, which means $\boxed{\sqrt{\frac{6}{\pi}}\sqrt{\frac{\omega^3}{12\lambda}}e^{-\frac{\omega^3}{12\lambda}} \ll 1}$. This will be true if $\lambda \ll 1$.

Chapter 3

The Periodic Well

Consider the same unit mass particle as before, but now under the influence of the potential $V[x] = \sum_{n\in\mathbb{Z}} v(x-na)$ where $v(x)$ is a single well inside $\left(-\frac{a}{2},\, \frac{a}{2}\right)$ and zero outside that interval
A practical example for a system that has such a potential is the sine Gordon in QFT or just a simple pendulum in QM.

The x-values of the minima (maxima) of the potential are the set $a\mathbb{Z}$.

3.1 Kronig-Penny Type Model

Using Bloch's theorem we know that the eigenvalues of this system will be divided into energy bands, each of which a continuum (indexed by $k \in \left(-\frac{\pi}{a},\, \frac{\pi}{a}\right)$) and that the energy eigenstates are also eigenstates of the translation operators by $a\mathbb{Z}$: $\hat{T}_{ma}\psi_k(x) \equiv \psi_k(x+ma) = e^{ikma}\psi_k(x)$. We shall try to obtain results using "instanton calculus" instead.

3.2 Generic Transition Amplitude

A single instanton $\frac{a}{2}\left[\tanh\left(\frac{\omega}{2}(t-\mathcal{T})\right)+2j+1\right]$ shifts site j to site $j+1$, anti-instantons $\frac{a}{2}\left[\tanh\left(-\frac{\omega}{2}(t-\mathcal{T})\right)+2j+1\right]$ shift from $j+1$ to j. Now an instanton is an event localized both in space *and* in time.

3.2.1 A sequence of single-instantons versus A single big instanton

If we want to move from j to $j+5$ we could think of two options:

1. Stringing together 5 1-instantons, $j \to j+1$, $j+1 \to j+2$, $\ldots$ Then the contribution is proportional to e^{-S_0} where $S_0 = 5\frac{2}{3}a^2\omega$ (as we computed in the double well).

2. Taking a single 5-instanton: $\frac{5a}{2}\left[\tanh\left(\frac{\omega}{2}(t-\mathcal{T})\right)+2j+1\right]$ contribution proportional to $e^{-S_0'}$ where $S_0' = \frac{2}{3}(5a)^2\omega$. Thus this is $\mathcal{O}\left(\left(e^{-S_0}\right)^5\right)$! We will only take into account stringing 1-instantons then.

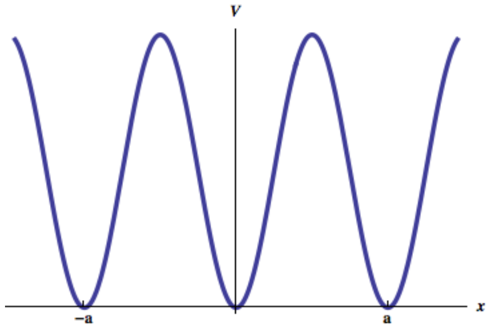

Figure 3.1: The periodic well.

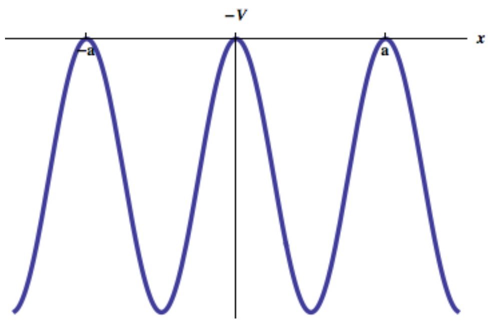

Figure 3.2: In Euclidean spacetime the potential is inverted.

So if we want to go from $j \to j'$ we assume any sequence of 1-instantons and anti 1-instantons are combined, just so that the number of instantons minus the number of anti-instantons is equal $j' - j$.

3.2.2 The Transition Amplitude

If we denote our minimum sites by j for $x = ja$ where $j \in \mathbb{Z}$, then we can write a transition amplitude as:

$$\lim_{T\to\infty} \left\langle j \left| e^{-\hat{H}T} \right| j' \right\rangle \approx \lim_{T\to\infty} \underbrace{\left(\sqrt{\frac{\omega}{\pi}} e^{-\frac{\omega}{2}T}\right)}_{\text{SHO}} \underbrace{\sum_{n\in\mathbb{N}} (D\omega)^n \frac{T^n}{n!}}_{n \text{ instantons}} \underbrace{\sum_{n'\in\mathbb{N}} (D\omega)^{n'} \frac{T^{n'}}{n'!}}_{n' \text{ anti-instantons}} \delta_{(j-j')-(n-n')}$$

What we have asserted with this statement is that we can have any number of instantons and anti-instantons (doesn't matter where) just as long as the total change in position $j - j'$ is equal to the total number of instantons n minus the total number of anti-instantons n'.

Next, write $\delta_{(j-j')-(n-n')} = \int_0^{2\pi} \frac{d\theta}{2\pi} e^{i\theta[(j-j')-(n-n')]}$ to get:

$$\begin{aligned}
\lim_{T\to\infty}\left\langle j\left|e^{-\hat{H}T}\right|j'\right\rangle &= \lim_{T\to\infty}\left(\sqrt{\frac{\omega}{\pi}}e^{-\frac{\omega}{2}T}\right)\sum_{n\in\mathbb{N}}[D\omega]^n\frac{T^n}{n!}\sum_{n'\in\mathbb{N}}[D\omega]^{n'}\frac{T^{n'}}{n'!}\int_0^{2\pi}\frac{d\theta}{2\pi}e^{i\theta[(j-j')-(n-n')]}\\
&= \lim_{T\to\infty}\left(\sqrt{\frac{\omega}{\pi}}e^{-\frac{\omega}{2}T}\right)\int_0^{2\pi}\frac{d\theta}{2\pi}e^{i\theta[(j-j')]}\sum_{n\in\mathbb{N}}\frac{1}{n!}\left[D\omega e^{-i\theta}T\right]^n\sum_{n'\in\mathbb{N}}\frac{1}{n'!}\left[D\omega e^{i\theta}T\right]^{n'}=\\
&= \lim_{T\to\infty}\left(\sqrt{\frac{\omega}{\pi}}e^{-\frac{\omega}{2}T}\right)\int_0^{2\pi}\frac{d\theta}{2\pi}e^{i\theta[(j-j')]}\exp\left[D\omega e^{-i\theta}T\right]\exp\left[D\omega e^{i\theta}T\right]\\
&= \lim_{T\to\infty}\left(\sqrt{\frac{\omega}{\pi}}e^{-\frac{\omega}{2}T}\right)\int_0^{2\pi}\frac{d\theta}{2\pi}e^{i\theta[(j-j')]}\exp\left[D\omega 2\cos(\theta)\,T\right]\\
&= \sqrt{\frac{\omega}{\pi}}\lim_{T\to\infty}\int_0^{2\pi}\frac{d\theta}{2\pi}e^{i\theta[(j-j')]}\exp\left[-\frac{\omega}{2}\left(1-4D\cos(\theta)\right)T\right]
\end{aligned}$$

3.2.3 The θ-Vacuum

We can read-off the energy eigenvalues from the previous expression readily:

$$E_\theta = \frac{\omega}{2} - D\omega 2\cos(\theta)$$

This is the same result we would obtain using Bloch's theorem. The barrier penetration coefficient enters in D via e^{-S_0}. This is already a "theta-vacuum" which will be important later in gauge theory: Even though we originally formulated a tunneling between $|j\rangle$ and $|j'\rangle$, it turns out that the actual vacua (by Bloch's theorem) have to be eigenstates of of the translation operator, and so, as we found, it is more natural to write $|\theta\rangle := \sum_n e^{-in\theta}|n\rangle$. Then

$$\begin{aligned}
\hat{T}_{ma}|\theta\rangle &= \sum_n e^{-ina\theta}\hat{T}_{ma}|n\rangle\\
&= \sum_n e^{-ina\theta}|n+m\rangle\\
&= e^{ima\theta}|\theta\rangle
\end{aligned}$$

Part II

Instantons in Quantum Field Theory

Chapter 4

Pure Yang-Mills Theory

In this chapter we finally get to field theory, in which the main goal will be to show that the vacuum has such a structure as to effectively add a CP-violating term to the Lagrangian. We follow the presentation in [4].

4.1 Gauge Theory

4.1.1 Gauge Groups and their Corresponding Lie Algebras

Let G be a compact Lie group, called *the gauge group*. Let $\{T^a\}_{a=1}^N$ be the generators of its corresponding Lie algebra $\mathfrak{g}$. Thus we have $\left[T^a, T^b\right] = f^{abc}T^c$ where f^{abc} are called *the structure constants* of $\mathfrak{g}$ and we employ the Einstein summation convention on the group indices (despite all group indices being superscript). For $SU(2)$ for example, $f^{abc} \equiv \varepsilon^{abc}$, the totally anti-symmetric tensor. If G is Abelian, for instance, for $U(1)$, the structure constants are $f^{abc} \equiv 0$.

We pick a representation of $\mathfrak{g}$ in which $tr\left(T^aT^b\right) \sim \delta^{ab}$. For example, for $SU(2)$, $T^a = -i\frac{\sigma^a}{2}$ where σ^a are the Pauli matrices. For $SU(3)$, $T^a = -i\frac{\lambda^a}{2}$ where λ^a are the Gell-Mann matrices.

4.1.1.1 Definition

Define the *Cartan inner product* between two generators so that we would have $\left(T^a, T^b\right) := \delta^{ab}$.

4.1.1.2 Claim

For $SU(2)$ in the representation of $\mathfrak{su}(2)$ specified above, $\left(T^a, T^b\right) = -2tr\left(T^aT^b\right)$.

Proof

$$\begin{aligned}
-2tr\left(T^aT^b\right) &= -2tr\left[\left(-i\frac{\sigma^a}{2}\right)\left(-i\frac{\sigma^b}{2}\right)\right] \\
&= \frac{1}{2}tr\left(i\varepsilon^{abc}\sigma^c+\delta^{ab}I\right) \\
&\overset{tr(\sigma^c)=0}{=} \frac{1}{2}\delta^{ab}tr\left(I\right) \\
&= \delta^{ab}
\end{aligned}$$

■

4.1.2 Gauge Fields

Consider a field theory of N real vector fields, $\{A_\mu{}^a\}_{a=1}^N$, where N is the dimension of $\mathfrak{g}$ above. Because we are interested in instanton solutions, we will work exclusively in Euclidean spacetime. Thus all spacetime indices, still in Greek letters, will be subscript, yet Einstein summation convention is still in effect.

For convenience work instead with one *matrix-valued* vector field $A_\mu := gA_\mu{}^aT^a$ where $g\in\mathbb{R}$ is a *coupling constant*. So A_μ takes values in $\mathfrak{g}$. Define the field-strength tensor $\boxed{F_{\mu\nu}:=\partial_{[\mu}A_{\nu]}+A_{[\mu}A_{\nu]}}$.

4.1.2.1 Note on Abelian Groups

If $G=U\left(1\right)$, this is the electromagnetic field-strength tensor because $U\left(1\right)$ is Abelian and for Abelian groups $[A_\mu,\,A_\nu]=0$.

4.1.2.2 The Action

Assume that the (Euclidean) action for this theory is given by $S\left[A_\mu\right]=\frac{1}{4g^2}\int_{\mathbb{R}^4}d^4x\left(F_{\mu\nu},\,F_{\mu\nu}\right)$. This is a reasonable assumption (that is, this is the most generic term to put in the action) taking into consideration certain constraints:

- Lorentz invariance.
- Gauge invariance (to be verified later on).
- The need for renormalizability (that is, we need the mass dimension of this term to obey a certain constraint).
- Naively assuming CP invariance of the theory (this will turn out to be a misguided assumption, and as a result, we will *add* another term to the Lagrangian).

Note: because we are in Euclidean spacetime, there is no distinction between lower and upper spacetime indices.

4.1.3 Gauge Transformations

Define a *local gauge transformation* as the following map on A_μ: $\boxed{A_\mu \mapsto VA_\mu V^{-1} + V\partial_\mu V^{-1}}$ for any $V \in G^{\mathbb{R}^4}$ (here V is function of spacetime, and should really be written $V(x)$, but this is not explicit in the notation in order to keep the expressions in manageable size).

4.1.3.1 Claim

Under such a map, $F_{\mu\nu} \mapsto VF_{\mu\nu}V^{-1}$.

Proof

$$\begin{aligned}
F_{\mu\nu} &\mapsto \partial_\mu A_\nu - \partial_\nu A_\mu + A_\mu A_\nu - A_\nu A_\mu \\
&= \partial_\mu \left(VA_\nu V^{-1} + V\partial_\nu V^{-1}\right) - \mu \leftrightarrow \nu \\
&\quad + \left(VA_\mu V^{-1} + V\partial_\mu V^{-1}\right)\left(VA_\nu V^{-1} + V\partial_\nu V^{-1}\right) - \mu \leftrightarrow \nu \\
&= \left(\partial_\mu V\right) A_\nu V^{-1} + V\left(\partial_\mu A_\nu\right) V^{-1} + VA_\nu \partial_\mu V^{-1} + \left(\partial_\mu V\right)\partial_\nu V^{-1} + V\partial_\mu\partial_\nu V^{-1} - \mu \leftrightarrow \nu \\
&\quad + VA_\mu A_\nu V^{-1} + VA_\mu \partial_\nu V^{-1} + V\left(\partial_\mu V^{-1}\right) VA_\nu V^{-1} + V\left(\partial_\mu V^{-1}\right) V\partial_\nu V^{-1} - \mu \leftrightarrow \nu \\
&\overset{*}{=} \underbrace{\left(\partial_\mu V\right) A_\nu V^{-1}}_{I} + V\left(\partial_\mu A_\nu\right) V^{-1} + VA_\nu\partial_\mu V^{-1} + \left(\partial_\mu V\right)\partial_\nu V^{-1} - \mu \leftrightarrow \nu \\
&\quad + VA_\mu A_\nu V^{-1} + VA_\mu\partial_\nu V^{-1} - \underbrace{\left(\partial_\mu V\right) A_\nu V^{-1}}_{I} - \left(\partial_\mu V\right)\partial_\nu V^{-1} - \mu \leftrightarrow \nu \\
&= V\left(\partial_\mu A_\nu\right) V^{-1} - \mu \leftrightarrow \nu \\
&\quad + VA_\mu A_\nu V^{-1} - \mu \leftrightarrow \nu \\
&\equiv VF_{\mu\nu}V^{-1}
\end{aligned}$$

where in $*$ we have used two facts:

1. $\partial_\mu \left(VV^{-1}\right)$
2. $\partial_{[\mu}\partial_{\nu]} = 0$

■

4.1.3.2 Claim

The action we defined above is gauge invariant

Proof $S\left[VA_\mu V^{-1} + V\partial_\mu V^{-1}\right] = \frac{1}{4g^2}\int_{\mathbb{R}^4} d^4x \left(VF_{\mu\nu}V^{-1},\, VF_{\mu\nu}V^{-1}\right)$. But this inner-product is proportional to the trace, which is cyclic, so we get $\frac{1}{4g^2}\int_{\mathbb{R}^4} d^4x \left(F_{\mu\nu},\, F_{\mu\nu}\right) \equiv S\left[A_\mu\right]$.
■

4.1.4 The Covariant Derivative

Define a *covariant derivative* for $F_{\mu\nu}$ by $D_\lambda F_{\mu\nu} := \partial_\lambda F_{\mu\nu} + \left[A_\lambda,\, F_{\mu\nu}\right]$.

4.1.4.1 Note

This reduces to the ordinary derivative when G is Abelian, because we have $[A_\lambda,\, F_{\mu\nu}] = 0$.

4.1.4.2 Claim

Then the equation of motion stemming from the action defined above is given by $\boxed{D_\mu F_{\mu\nu} = 0}$ (which is a generalized inhomogeneous source-free Maxwell's equation.)

Proof

- We have $\mathcal{L} = \frac{1}{4g^2}\left(F_{\mu\nu},\, F_{\mu\nu}\right)$.
- The equation of motion is given by $\partial_\beta \frac{\partial \mathcal{L}}{\partial(\partial_\beta {A_\alpha}^d)} - \frac{\partial\mathcal{L}}{\partial {A_\alpha}^d} = 0$ for all α in the spacetime indices ($\{1,\, 2,\, 3,\, 4\}$) and for all d in the group indices ($\{1,\, \dots,\, N\}$).
- $\frac{\partial\mathcal{L}}{\partial {A_\alpha}^d} = \frac{\partial}{\partial {A_\alpha}^d}\frac{1}{4g^2}\left(F_{\mu\nu},\, F_{\mu\nu}\right) = \frac{1}{4g^2}\left(\frac{\partial}{\partial {A_\alpha}^d}F_{\mu\nu},\, F_{\mu\nu}\right) + \frac{1}{4g^2}\left(F_{\mu\nu},\, \frac{\partial}{\partial {A_\alpha}^d}F_{\mu\nu}\right)$
- We compute one derivative first:

$$\begin{aligned}
\frac{\partial}{\partial {A_\alpha}^d}F_{\mu\nu} &= \frac{\partial}{\partial {A_\alpha}^d}\left[A_\mu,\, A_\nu\right] \\
&= \frac{\partial}{\partial {A_\alpha}^d}\left[g{A_\mu}^a T^a,\, g{A_\nu}^b T^b\right] \\
&= g^2 \frac{\partial}{\partial {A_\alpha}^d} {A_\mu}^a {A_\nu}^b f^{abc} T^c \\
&= g^2 f^{abc} T^c \left(\delta_{\alpha\mu}\delta^{da}{A_\nu}^b + {A_\mu}^a \delta_{\alpha\nu}\delta^{db}\right) \\
&= g^2\left(f^{dbc}\delta_{\alpha\mu}{A_\nu}^b + f^{adc}{A_\mu}^a\delta_{\alpha\nu}\right)T^c
\end{aligned}$$

- If we write $F_{\mu\nu} = g{F_{\mu\nu}}^c T^c$ where ${F_{\mu\nu}}^c \equiv \partial_\mu {A_\nu}^c - \partial_\nu {A_\mu}^c + g{A_\mu}^a {A_\nu}^b f^{abc}$ then we have for the complete derivative of the Lagrangian:

$$\begin{aligned}
\frac{\partial\mathcal{L}}{\partial {A_\alpha}^d} &= \frac{1}{4g^2}\left(\frac{\partial}{\partial {A_\alpha}^d}F_{\mu\nu},\, F_{\mu\nu}\right) + \frac{1}{4g^2}\left(F_{\mu\nu},\, \frac{\partial}{\partial {A_\alpha}^d}F_{\mu\nu}\right) \\
&= \frac{1}{4g^2}\left(g^2\left(f^{dbc}\delta_{\alpha\mu}{A_\nu}^b + f^{adc}{A_\mu}^a\delta_{\alpha\nu}\right)T^c,\, g{F_{\mu\nu}}^e T^e\right) + \frac{1}{4g^2}\left(F_{\mu\nu},\, \frac{\partial}{\partial {A_\alpha}^d}F_{\mu\nu}\right) \\
&= \frac{1}{2}g\left(f^{dbc}\delta_{\alpha\mu}{A_\nu}^b + f^{adc}{A_\mu}^a\delta_{\alpha\nu}\right){F_{\mu\nu}}^c \\
&= \frac{1}{2}g\left(f^{dbc}{A_\nu}^b {F_{\alpha\nu}}^c + f^{adc}{A_\mu}^a {F_{\mu\alpha}}^c\right) \\
&= \frac{1}{2}g\left(f^{dac}{A_\mu}^a {F_{\alpha\mu}}^c - f^{dac}{A_\mu}^a {F_{\mu\alpha}}^c\right) \\
&= g f^{dac}{A_\mu}^a {F_{\alpha\mu}}^c
\end{aligned}$$

- Next we need to compute $\frac{\partial\mathcal{L}}{\partial(\partial_\beta {A_\alpha}^d)} = \frac{1}{4g^2}\left(\frac{\partial}{\partial(\partial_\beta {A_\alpha}^d)}F_{\mu\nu},\, F_{\mu\nu}\right) + \frac{1}{4g^2}\left(F_{\mu\nu},\, \frac{\partial}{\partial(\partial_\beta {A_\alpha}^d)}F_{\mu\nu}\right)$.

- Computing only one derivative we have:

$$\begin{aligned}\frac{\partial}{\partial\left(\partial_\beta A_\alpha{}^d\right)} F_{\mu\nu} &= \frac{\partial}{\partial\left(\partial_\beta A_\alpha{}^d\right)}\left[gT^a\partial_\mu A_\nu{}^a - gT^a\partial_\nu A_\mu{}^a\right] \\ &= g\left[\delta_{\beta\mu}\delta_{\alpha\nu}\delta^{ad} - \delta_{\beta\nu}\delta_{\alpha\mu}\delta^{ad}\right]T^a \\ &= g\left[\delta_{\beta\mu}\delta_{\alpha\nu} - \delta_{\beta\nu}\delta_{\alpha\mu}\right]T^d\end{aligned}$$

- Thus we get for the complete derivative:

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\left(\partial_\beta A_\alpha{}^d\right)} &= \frac{1}{4g^2}\left(\frac{\partial}{\partial\left(\partial_\beta A_\alpha{}^d\right)}F_{\mu\nu},\, F_{\mu\nu}\right) + \frac{1}{4g^2}\left(F_{\mu\nu},\, \frac{\partial}{\partial\left(\partial_\beta A_\alpha{}^d\right)}F_{\mu\nu}\right) \\ &= \frac{1}{4g^2}\left(g\left[\delta_{\beta\mu}\delta_{\alpha\nu} - \delta_{\beta\nu}\delta_{\alpha\mu}\right]T^d,\, gF_{\mu\nu}{}^cT^c\right) + \frac{1}{4g^2}\left(F_{\mu\nu},\, \frac{\partial}{\partial\left(\partial_\beta A_\alpha{}^d\right)}F_{\mu\nu}\right) \\ &= \frac{1}{2}\left[\delta_{\beta\mu}\delta_{\alpha\nu} - \delta_{\beta\nu}\delta_{\alpha\mu}\right]F_{\mu\nu}{}^d \\ &= F_{\beta\alpha}{}^d\end{aligned}$$

- Thus we found the equation of motion is given by $\partial_\beta F_{\beta\alpha}{}^d - gf^{dac}A_\mu{}^aF_{\alpha\mu}{}^c = 0$.

- Multiply this equation by g and also by T^d. We will get N such equation for each d. Sum up all these equations to get: $\partial_\beta F_{\beta\alpha} - g^2 f^{dac}A_\mu{}^aF_{\alpha\mu}{}^cT^d = 0$.

- Using the definition of the structure constants $f^{dac}T^d = f^{acd}T^d = [T^a,\, T^c]$ we have:

$$\begin{aligned}\partial_\beta F_{\beta\alpha} - g^2 A_\mu{}^aF_{\alpha\mu}{}^c\left[T^a,\, T^c\right] &= 0 \\ \partial_\beta F_{\beta\alpha} - \left[gA_\mu{}^aT^a,\, gF_{\alpha\mu}{}^cT^c\right] &= 0 \\ \partial_\beta F_{\beta\alpha} - \left[A_\mu,\, F_{\alpha\mu}\right] &= 0 \\ \partial_\mu F_{\mu\alpha} + \left[A_\mu,\, F_{\mu\alpha}\right] &= 0 \\ D_\mu F_{\mu\alpha} &= 0\end{aligned}$$

■

4.1.4.3 Note about Abelian Groups

When G is Abelian, $[A_\nu,\, F_{\mu\nu}] = 0$ and the solution is $A_\mu = 0$. In the non-Abelian case, this equation is non-linear and non-trivial solutions may exist.

4.2 Finite Action

Just as for quantum mechanics, now for gauge field theory we are interested in the lowest energy eigenvalues. Thus we want to compute a path-integral. As we've seen in the case of quantum mechanics, it is thus worthwhile to know the classical solutions to the equations of motion and approximate the path-integral about those solutions.

Something that was implicit above should now be made clear: for this type of approximation (semi-classical approximation), classical solutions which have an infinite value for their corresponding action are not important, because their contribution to the path-integral is proportional to $e^{-S[sol.]}$ (however, in general, it is actually the solutions with *finite action* which form a set of measure zero in the space of all functions (and so should not contribute to the integral), and only gain significance in the context of the semi-classical approximation.)

4.2.0.4 Claim

In four spacetime dimensions if $S < \infty$ then $F_{\mu\nu}$ must decrease *more rapidly* than $\frac{1}{|x|^2}$.

Proof Otherwise $S \sim \int_{\mathbb{R}^4} d^4x \left(\frac{1}{r^2}\right)^2 \sim \int_0^\infty drr^3 \frac{1}{r^4} \sim \log(r)|_0^\infty \to \infty$.
■

4.2.1 Field Configurations of Finite Action

So we are interested in such field configurations so that in some series expansion of $\lim_{r\to\infty} F$ in powers of $\frac{1}{r}$, the first term in F is $\frac{1}{r^3}$. Naively, this means that in some series expansion of $\lim_{r\to\infty} A$ in powers of $\frac{1}{r}$, the first term in A is $\frac{1}{r^2}$, because $F \propto \partial_\mu A \propto \frac{1}{r^3}$.

4.2.1.1 Pure Gauge Configurations

However, it turns out that there is another possibility, which is more interesting. F could also be zero if A is some gauge transformation of zero (such a configuration is called a *pure gauge*). That is, $A_\mu = V\partial_\mu V^{-1}$ for some $V(x) \in G^{\mathbb{R}^4}$.

4.2.1.2 Claim

If A is a pure gauge then $F = 0$.

Proof

$$\begin{aligned}
F_{\mu\nu} &= \partial_{[\mu}\left(V\partial_{\nu]}V^{-1}\right) + \left[V\partial_\mu V^{-1}, V\partial_\nu V^{-1}\right] \\
&= V_{,\mu}V^{-1}{}_{,\nu} + VV^{-1}{}_{,\nu,\mu} - V_{,\nu}V^{-1}{}_{,\mu} - VV^{-1}{}_{,\mu,\nu} \\
&\quad +VV^{-1}{}_{,\mu}VV^{-1}{}_{,\nu} - VV^{-1}{}_{,\nu}VV^{-1}{}_{,\mu} \\
&\overset{(V_{,\mu,\nu}=V_{,\nu,\mu})}{=} V_{,\mu}V^{-1}{}_{,\nu} - V_{,\nu}V^{-1}{}_{,\mu} + VV^{-1}{}_{,\mu}VV^{-1}{}_{,\nu} - VV^{-1}{}_{,\nu}VV^{-1}{}_{,\mu} \\
&\overset{\left(\left(VV^{-1}\right)_{,\mu}=0\right)}{=} V_{,\mu}V^{-1}{}_{,\nu} - V_{,\nu}V^{-1}{}_{,\mu} - V_{,\mu}V^{-1}{}_{,\nu} + V_{,\nu}V^{-1}{}_{,\mu} \\
&= 0
\end{aligned}$$

■

Conclusion Thus for finite action we need such fields configurations so that $\lim_{|x|\to\infty} A_\mu \overset{!}{=} V\partial_\mu V^{-1} + \mathcal{O}\left(\frac{1}{|x|^2}\right)$ for some $V(x) \in G^{\mathbb{R}^4}$. However, because in this expression V is only evaluated for $|x| \to \infty$, we can conveniently think of $V(x) \in G^{S^3}$ instead of $V(x) \in G^{\mathbb{R}^4}$, where S^3, mathematically the 4-dimensional sphere with radius *one*, is for us homeomorphic to the

4-dimensional sphere with radius infinity. So every finite action field configuration is associated with an element $V(x) \in G^{S^3}$. But if two pure gauge field configurations are in the same gauge orbit, their corresponding V is *not* the same. That is, V, which characterizes a field configuration, is *not* gauge invariant.

4.2.1.3 Claim

Gauge transforming a pure gauge field configuration $V\partial_\mu V^{-1}$ with $U \in G$ transforms V into UV.

Proof If we perform a gauge transformation $A_\mu \mapsto UA_\mu U^{-1} + U\partial_\mu U^{-1}$ on a pure gauge configuration $V\partial_\mu V^{-1}$ we get:

$$\begin{aligned} V\partial_\mu V^{-1} &\mapsto UV\left(\partial_\mu V^{-1}\right)U^{-1} + U\partial_\mu U^{-1} \\ &= UV\partial_\mu\left(V^{-1}U^{-1}\right) - UVV^{-1}\partial_\mu U^{-1} + U\partial_\mu U^{-1} \\ &= UV\partial_\mu (UV)^{-1} \end{aligned}$$

■

Conclusion Thus effectively we have $V \mapsto UV$ instead of $V \mapsto V$, which is what we would expect from a gauge invariant object.

4.2.1.4 Claim

It is not possible, in general, to "gauge away" any pure gauge field configuration to zero by picking $U|_{r=\infty} = V^{-1}\big|_{r=\infty}$, thereby arranging that $A_\mu = \mathbb{1}\partial_\mu\mathbb{1}^{-1} = 0$.

Proof We assume all gauge transformations are continuous *throughout spacetime*, so $U|_{r=\infty}$ must a continuous deformation of $U|_{r=0} \in G^{S^3}$, which must be a constant (because it cannot depend on the angles). But all constant elements in G^{S^3} are continuously deformable into $\mathbb{1}$ (because we assume G is connected). So that we may only pick such $U|_{r=\infty}$ which are continuous deformations of $\mathbb{1}$. Thus, if V^{-1} is *not* continuously deformable into $\mathbb{1}$ (which in general could be the case) then we cannot pick $U = V^{-1}$.
■

Conclusion Even though we found $V \mapsto UV$ instead of $V \mapsto V$, U must be continously deformable to $\mathbb{1}$ and so: V and UV are continuously deformable into one another.

As a result, we find that the gauge-invariant object associated with $A_\mu = V\partial_\mu V^{-1}$ is not $V(x)$ per se but the class of all elements of G^{S^3} which are continuously deformable into $V(x)$.

4.3 Homotopy

Two maps that are continuously deformable into one another are *homotopic*. So to classify the gauge-invariant objects associated with finite-action field configurations, we need to find all homotopy classes in G^{S^3}. We also specialize to the case $G = SU(2)$ and so $\mathfrak{g} = \mathfrak{su}(2)$.

4.3.0.5 Claim

$SU(2) \simeq S^3$

Proof Write a general element of $SU(2)$ as $\begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix}$ where $\{a, b\} \subset \mathbb{C}$ and $|a|^2 + |b|^2 = 1$. This is indeed the most general form of an element in $Mat_{2\times 2}(\mathbb{C})$ which is unitary (is its own inverse) and has determinant 1 (to see that write a general 2×2 matrix with complex-number entries and solve the system of equations stemming from these two constraints). Define a map $\mathbb{C}^2 \to Mat_{2\times 2}(\mathbb{C})$ by $\begin{bmatrix} a \\ b \end{bmatrix} \mapsto \begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix}$. This is a homeomorphism onto its image (injective, continuous, and inverse is continuous). Thus $SU(2)$ can be identified with the vectors in $(a, b) \in \mathbb{C}^2$ which have $|a|^2 + |b|^2 = 1$, which is exactly the unit sphere in $\mathbb{C}^2 \simeq \mathbb{R}^4$–the 3-sphere.
■

Conclusion Using this homeomorphism we only need to think of homotopy classification of maps in $\left(S^3\right)^{S^3}$ instead of $SU(2)^{S^3}$, which is more convenient.

4.3.1 Third Homotopy Group of $SU(2)$ and the Winding Number

The classifications of all maps $S^3 \to S^3$ belongs in the field of algebraic topology [6]. For topologists, one main goal is to classify the spaces which constitute the *range* of these maps. Thus, in very crude terms, they decide if the topological spaces A and B are equivalent if the set of maps $S^n \to A$ is "equivalent" to the set of maps $S^n \to B$ for some $n \in \mathbb{N}$, where S^n is the n-sphere in $\mathbb{R}^n$. It turns out that these sets of maps (or rather equivalence classes of them–homotopy classes) form the mathematical structure of a group. This group is called *the n-th homotopy group of a space*. In order to decide if two groups are equivalent we have at our disposal the notion of *group isomorphism*. Algebraic topologists can then very clearly rule that A and B are *not* equivalent (homeomorphic, in topological jargon) if the corresponding n-th homotopy groups are not isomorphic (the converse is not in general true).

For our current purposes, the distinction between two topological spaces is not so important, but rather a "by product" of the hard work that algebraic topologists have made: the construction of the homotopy group. In order to give the n-th homotopy group the structure of a group, an equivalence relation is defined on the set of maps $S^n \to A$. Two maps $\{f, g\} \subset A^{S^n}$ are equivalent iff $\exists$ a point $x \in S^n$ and a *continuous* map $h \in A^{[0, 1]\times S^n}$, called a homotopy between f and g, such that:

1. $h(0, -) = f(-)$
2. $h(1, -) = g(-)$
3. $h(-, x) = f(x) = g(x)$ for some $x \in S^n$.

That is, h can be thought of as a continuous deformation or interpolation between the "path" f to the "path" g, where all the deformations are based at $f(x) = g(x)$. The group of classes of maps in A^{S^n} which *have* a homotopy between them is denoted by $\pi_n(A)$. The law of composition on this group is defined in a natural way by concatenating two paths (and taking the equivalence class of that), the identity is the constant path at a point (or all paths equivalent to that), and inverses are paths that go in reverse direction. Thus it is clear that, for instance, if $\pi_1(A) \simeq \{0\}$, that is, the trivial group, then all paths are homotopic to the constant point. What that means is that all paths can be continuously contracted into one point (or rather, a path that just goes through one point for its range). This is not always possible, but when it is, the space is called *simply connected*.

Back to our matter at hand, we are interested in computing $\pi_3\left(S^3\right)$.

4.3.1.1 Claim

$\pi_3\left(S^3\right) \simeq \mathbb{Z}$.

Proof

- We know that $\exists$ a group homomorphism $\pi_3\left(S^3\right) \to H_3\left(S^3\right)$ where H_3 is the third *homology* group with integer coefficients.
- According to Hurewicz' theorem [13], because S^3 is 2-connected, this map is an isomorphism.
- But $H_3\left(S^3\right) = \mathbb{Z}$ is easy to calculate.

■

Conclusion This integer in $\mathbb{Z}$ represents the number of times the 3-sphere wraps around itself (negative values for opposite orientation).

So the finite-action field configurations are "indexed" by $\mathbb{Z}$, in the sense that each A_μ obtains a label from $\mathbb{Z}$ and iff two finite-action field configurations have the same label they are homotopic.

This label is conventionally called *the winding number* of an element of $\left(S^3\right)^{S^3}$.

4.3.2 Standard Mappings of Integer Winding Numbers

Define the following reference maps which will serve us later:

1. $B(x)^{(0)} := \mathbb{1}$
2. $B(x)^{(1)} := \frac{1}{|x|}\left(x_4\mathbb{1} + i\vec{x}\cdot\vec{\sigma}\right)$ where $|x| \equiv \sqrt{(x_4)^2 + (x_1)^2 + (x_2)^2 + (x_3)^2}$. Observe how in this definition the group indices and the spacetime indices are mixed.

3. $B(x)^{(\nu)} := \left[B(x)^{(1)}\right]^{\nu}$ for any $\nu \in \mathbb{Z}$.

4.3.2.1 Claim

Using the homeomorphism we established between $SU(2) \simeq S^3$, $B(x)^{(1)}$ is actually the identity mapping between $S^3 \overset{B(x)^{(1)}}{\to} SU(2) \simeq S^3$.

Proof Plugging in the actual Pauli matrices, we have

$$\begin{aligned} B(x)^{(1)} &\equiv \frac{x_4}{|x|}\mathbb{1} + i\frac{\vec{x}}{|x|}\cdot\vec{\sigma} \\ &= \frac{1}{|x|}\begin{bmatrix} x_4 + ix_3 & ix_1 + x_2 \\ ix_1 - x_2 & x_4 - ix_3 \end{bmatrix} \end{aligned}$$

Now we use the homeomorphism from S^3 to $SU(2)$ which we established $\begin{bmatrix} a \\ b \end{bmatrix} \mapsto \begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix}$, to write $\frac{1}{|x|}\begin{bmatrix} x_4 + ix_3 & ix_1 + x_2 \\ ix_1 - x_2 & x_4 - ix_3 \end{bmatrix} \mapsto \frac{1}{|x|}\begin{bmatrix} x_4 + ix_3 \\ ix_1 + x_2 \end{bmatrix} \in \mathbb{C}^2$. Now use the homeomorphism between $\mathbb{C}^2 \simeq \mathbb{R}^4$ to write $\frac{1}{|x|}\begin{bmatrix} x_4 + ix_3 \\ ix_1 + x_2 \end{bmatrix} \mapsto \frac{1}{|x|}\begin{bmatrix} x_4 \\ x_3 \\ x_2 \\ x_1 \end{bmatrix}$. So clearly if we *started* in S^3 we ended up in *exactly* the same point.
■

4.3.2.2 Claim

$B(x)^{(\nu)} \in SU(2)^{S^3}$ for all $\nu \in \mathbb{Z}$.

Proof

- When $\nu = 0$ the claim is true.
- When $\nu = 1$:
  - $\left(\frac{x_4}{|x|}\mathbb{1} + i\frac{\vec{x}}{|x|}\cdot\vec{\sigma}\right)^{-1} = \frac{x_4}{|x|}\mathbb{1} - i\frac{\vec{x}}{|x|}\cdot\vec{\sigma}$ because $\left(\frac{x_4}{|x|}\mathbb{1} - i\frac{\vec{x}}{|x|}\cdot\vec{\sigma}\right)\left(\frac{x_4}{|x|}\mathbb{1} + i\frac{\vec{x}}{|x|}\cdot\vec{\sigma}\right) = \frac{(x_4)^2\mathbb{1} + (\vec{x}\vec{\sigma})^2}{|x|^2}$.
    * But

$$\begin{aligned} (\vec{x}\vec{\sigma})^2 &= \sum_{i,j=1}^{3} x_i x_j \sigma_i \sigma_j \\ &= \sum_{i,j=1}^{3} x_i x_j \left(\delta_{ij}\mathbb{1} + i\sum_k \varepsilon_{ijk}\sigma_k\right) \\ &= (\vec{x})^2\mathbb{1} + i\sum_{i,j,k=1}^{3} \varepsilon_{ijk} x_i x_j \sigma_k \end{aligned}$$

* $\varepsilon_{ijk}x_ix_j = 0$ because $\varepsilon_{ijk}x_ix_j \overset{\text{relabeling}_{i\longleftrightarrow j}}{=} \varepsilon_{jik}x_jx_i \overset{x_ix_j=x_jx_i}{=} \varepsilon_{jik}x_ix_j = -\varepsilon_{ijk}x_ix_j$.
* So we have $\left(\frac{x_4}{|x|}\mathbb{1} - i\frac{\vec{x}}{|x|}\cdot\vec{\sigma}\right)\left(\frac{x_4}{|x|}\mathbb{1} + i\frac{\vec{x}}{|x|}\cdot\vec{\sigma}\right) = \frac{(x_4)^2\mathbb{1}+(\vec{x})^2\mathbb{1}}{|x|^2} = \mathbb{1}$.
* This shows that $\frac{x_4}{|x|}\mathbb{1}+i\frac{\vec{x}}{|x|}\cdot\vec{\sigma}$ is unitary, because $\left(\frac{x_4}{|x|}\mathbb{1} + i\frac{\vec{x}}{|x|}\cdot\vec{\sigma}\right)^* = \frac{x_4}{|x|}\mathbb{1} - i\frac{\vec{x}}{|x|}\cdot\vec{\sigma}$ (recall that the Pauli matrices are Hermitian).

– To show that the determinant is really one:

$$\begin{aligned}\det\left(\frac{x_4}{|x|}\mathbb{1} + i\frac{\vec{x}}{|x|}\cdot\vec{\sigma}\right) &= \frac{1}{|x|^2}\det\left(\begin{bmatrix}x_4+ix_3 & ix_1+x_2\\ ix_1-x_2 & x_4-ix_3\end{bmatrix}\right)\\ &= \frac{1}{|x|^2}\left\{(x_4)^2+(x_3)^2-\left[-(x_1)^2-(x_2)^2\right]\right\}\\ &= 1\end{aligned}$$

– So we conclude that $\frac{x_4}{|x|}\mathbb{1} + i\frac{\vec{x}}{|x|}\cdot\vec{\sigma}$ is a bona fide element of $SU(2)$.

- We know that $SU(2)$ is a group. In particular, it is *closed* under multiplication and inverses. So $B(x)^{(\nu)} \equiv \left[B(x)^{(1)}\right]^{\nu}$ must lie in $SU(2)$ as well, for any $\nu \in \mathbb{Z}$.

■

Note With these standard mappings we can construct finite-action field configurations of arbitrary winding numbers.

4.3.3 Topological Charge

4.3.3.1 Definition

Define *the Cartan-Maurer Integral Invariant*,

$$\nu[V(x)] := \frac{1}{48\pi^2}\int_{S^3} d\theta_1 d\theta_2 d\theta_3 \sum_{i,j,k=1}^{3}\varepsilon_{ijk}\left(V\frac{\partial}{\partial\theta_i}V^{-1}, V\left(\frac{\partial}{\partial\theta_j}V^{-1}\right)V\frac{\partial}{\partial\theta_k}V^{-1}\right)$$

$\forall V(x) \in SU(2)^{S^3}$ where θ_1, θ_2 and θ_3 are angles that parametrize S^3 and $(\,,\,)$ is the Cartan inner product.

4.3.3.2 Claim

The definition does not depend on a particular choice of parametrization

Proof Follows directly from the fact that $\varepsilon_{ijk}\frac{\partial\theta'_l}{\partial\theta_i}\frac{\partial\theta'_m}{\partial\theta_j}\frac{\partial\theta'_n}{\partial\theta_k} = \det\left(\frac{\partial\theta'}{\partial\theta}\right)\varepsilon_{lmn}$.

■

4.3.3.3 Example

In particular, for the representation we chose of $SU(2)$ we have:

$$\nu[V(x)] = -\frac{1}{24\pi^2}\int_0^{\pi}d\theta_1\int_0^{\pi}d\theta_2\int_0^{2\pi}d\theta_3\varepsilon_{ijk}tr\left(V\left(\partial_iV^{-1}\right)V\left(\partial_jV^{-1}\right)V\partial_kV^{-1}\right)$$

4.3.3.4 Claim

$\nu\left[V\left(x\right)\right]=\nu\left[\tilde{V}\left(x\right)\right]$ if $V\left(x\right)$ is homotopic to $\tilde{V}\left(x\right)$.

Proof Suffice to show that $\nu\left[V\left(x\right)\right]$ is invariant under infinitesimal deformations to $V\left(x\right)$, because we can build continuous deformations from infinitesimal deformations.

- If $V\left(x\right)=\exp\left(i\lambda^{a}\left(x\right)T^{a}\right)$, then $\delta V=V\delta\lambda^{a}\left(x\right)T^{a}\equiv V\delta T$.
- Then
$$\begin{aligned}\delta\left(V\partial_{i}V^{-1}\right) & = \left(\delta V\right)\partial_{i}V^{-1}+V\partial_{i}\left(\delta V^{-1}\right)\\ & = V\left(\delta T\right)\partial_{i}V^{-1}+V\partial_{i}\left(-\delta TV^{-1}\right)\\ & = -V\left(\partial_{i}\delta T\right)V^{-1}\end{aligned}$$
.
- Because all three derivatives in $\nu\left[V\left(x\right)\right]$ make an equal contribution, we find that
$$\begin{aligned}\nu\left[\delta V\right] & \propto \int_{S^{3}}d\theta_{1}d\theta_{2}d\theta_{3}\varepsilon^{ijk}\left(V\partial_{i}V^{-1},\,V\left(\partial_{j}V^{-1}\right)V\left(\partial_{k}\delta T\right)V^{-1}\right)\\ & \propto \int_{S^{3}}d\theta_{1}d\theta_{2}d\theta_{3}\varepsilon^{ijk}\left(\partial_{i}V^{-1},\,-\left(\partial_{j}V\right)\left(\partial_{k}\delta T\right)\right)\end{aligned}$$
- If we make partial integration we get a term symmetric in ik or jk, and together with ε^{ijk}, $\nu\left[\delta V\right]=0$.

■

4.3.3.5 Claim

$\nu\left[B\left(x\right)^{(1)}\right]=1$

Proof

- Because the trace is cyclic, we have
$$\begin{aligned}\varepsilon_{ijk}tr\left(V\left(\partial_{i}V^{-1}\right)V\left(\partial_{j}V^{-1}\right)V\partial_{k}V^{-1}\right) & = 3tr\left[V\left(\partial_{1}V^{-1}\right)V\left(\partial_{2}V^{-1}\right)V\partial_{3}V^{-1}\right]\\ & \quad -3tr\left[V\left(\partial_{1}V^{-1}\right)V\left(\partial_{3}V^{-1}\right)V\partial_{2}V^{-1}\right]\end{aligned}$$
- We introduce the following parametrization of S^{3}: $\begin{bmatrix}x_{4}\\x_{3}\\x_{2}\\x_{1}\end{bmatrix}=\begin{bmatrix}\cos\left(\theta_{1}\right)\\\sin\left(\theta_{1}\right)\cos\left(\theta_{2}\right)\\\sin\left(\theta_{1}\right)\sin\left(\theta_{2}\right)\sin\left(\theta_{3}\right)\\\sin\left(\theta_{1}\right)\sin\left(\theta_{2}\right)\cos\left(\theta_{3}\right)\end{bmatrix}$ where $\left\{\theta_{1},\,\theta_{2}\right\}\subset\left[0,\,\pi\right]$ and $\theta_{3}\in\left[0,\,2\pi\right)$. For brevity introduce the following notation: $s_{j}:=\sin\left(\theta_{j}\right)$ and $c_{j}:=\cos\left(\theta_{j}\right)$ for all $j\in\left\{1,\,2,\,3\right\}$. So we have $\begin{bmatrix}x_{4}\\x_{3}\\x_{2}\\x_{1}\end{bmatrix}=\begin{bmatrix}c_{1}\\s_{1}c_{2}\\s_{1}s_{2}s_{3}\\s_{1}s_{2}c_{3}\end{bmatrix}$

- Then

$$\begin{aligned} B\left(\theta_1,\,\theta_2,\,\theta_3\right)^{(1)} &\equiv \frac{x_4}{|x|}\mathbb{1} + i\frac{\vec{x}}{|x|}\cdot\vec{\sigma} \\ &= \left(c_1\mathbb{1} + is_1s_2c_3\sigma_1 + s_1s_2s_3\sigma_2 + s_1c_2\sigma_3\right) \\ &= \begin{bmatrix} c_1 + is_1c_2 & is_1s_2c_3 + s_1s_2s_3 \\ is_1s_2c_3 - s_1s_2s_3 & c_1 - is_1c_2 \end{bmatrix} \end{aligned}$$

- Thus

$$\left[B\left(\theta_1,\,\theta_2,\,\theta_3\right)^{(1)}\right]^{-1} = \left(c_1\mathbb{1} - is_1s_2c_3\sigma_1 - s_1s_2s_3\sigma_2 - s_1c_2\sigma_3\right)$$

- Compute the derivatives:
  - For θ_1:

$$\partial_{\theta_1}\left\{\left[B\left(\theta_1,\,\theta_2,\,\theta_3\right)^{(1)}\right]^{-1}\right\} = \begin{bmatrix} -s_1 - ic_1c_2 & -ic_1s_2c_3 - c_1s_2s_3 \\ -ic_1s_2c_3 + c_1s_2s_3 & -s_1 + ic_1c_2 \end{bmatrix}$$

  - For θ_2:

$$\partial_{\theta_2}\left\{\left[B\left(\theta_1,\,\theta_2,\,\theta_3\right)^{(1)}\right]^{-1}\right\} = \begin{bmatrix} is_1s_2 & -is_1c_2c_3 - s_1c_2s_3 \\ -is_1c_2c_3 + s_1c_2s_3 & -is_1s_2 \end{bmatrix}$$

  - For θ_3:

$$\partial_{\theta_3}\left\{\left[B\left(\theta_1,\,\theta_2,\,\theta_3\right)^{(1)}\right]^{-1}\right\} = \begin{bmatrix} 0 & -s_1s_2c_3 + is_1s_2s_3 \\ s_1s_2c_3 + is_1s_2s_3 & 0 \end{bmatrix}$$

- Thus for the first term:

$$\begin{aligned} B^{(1)}\partial_{\theta_1}\left[B^{(1)}\right]^{-1} &= \begin{bmatrix} c_1 + is_1c_2 & is_1s_2c_3 + s_1s_2s_3 \\ is_1s_2c_3 - s_1s_2s_3 & c_1 - is_1c_2 \end{bmatrix}\begin{bmatrix} -s_1 - ic_1c_2 & -ic_1s_2c_3 - c_1s_2s_3 \\ -ic_1s_2c_3 + c_1s_2s_3 & -s_1 + ic_1c_2 \end{bmatrix} \\ &= \begin{bmatrix} -ic_2 & s_2\left(-ic_3 - s_3\right) \\ s_2\left(-ic_3 + s_3\right) & ic_2 \end{bmatrix} \end{aligned}$$

- For the second term:

$$\begin{aligned} B^{(1)}\partial_{\theta_2}\left[B^{(1)}\right]^{-1} &= \begin{bmatrix} c_1 + is_1c_2 & is_1s_2c_3 + s_1s_2s_3 \\ is_1s_2c_3 - s_1s_2s_3 & c_1 - is_1c_2 \end{bmatrix}\begin{bmatrix} is_1s_2 & -is_1c_2c_3 - s_1c_2s_3 \\ -is_1c_2c_3 + s_1c_2s_3 & -is_1s_2 \end{bmatrix} \\ &= \begin{bmatrix} ic_1s_1s_2 & s_1\left(s_1 - ic_1c_2\right)\left(c_3 - is_3\right) \\ -s_1\left(s_1 + ic_1c_2\right)\left(c_3 + is_3\right) & -ic_1s_1s_2 \end{bmatrix} \end{aligned}$$

- For the third term:

$$\begin{aligned} B^{(1)}\partial_{\theta_3}\left[B^{(1)}\right]^{-1} &= \begin{bmatrix} c_1 + is_1c_2 & is_1s_2c_3 + s_1s_2s_3 \\ is_1s_2c_3 - s_1s_2s_3 & c_1 - is_1c_2 \end{bmatrix}\begin{bmatrix} 0 & -s_1s_2c_3 + is_1s_2s_3 \\ s_1s_2c_3 + is_1s_2s_3 & 0 \end{bmatrix} \\ &= \begin{bmatrix} i\left(s_1\right)^2\left(s_2\right)^2 & -s_1s_2\left(is_1c_2 + c_1\right)\left(c_3 - is_3\right) \\ s_1s_2\left(-is_1c_2 + c_1\right)\left(c_3 + is_3\right) & -i\left(s_1\right)^2\left(s_2\right)^2 \end{bmatrix} \end{aligned}$$

- So we find

$$B^{(1)}\left(\partial_{\theta_1}\left[B^{(1)}\right]^{-1}\right)B^{(1)}\left(\partial_{\theta_2}\left[B^{(1)}\right]^{-1}\right)B^{(1)}\partial_{\theta_3}\left[B^{(1)}\right]^{-1} = \begin{bmatrix}-\left(s_1\right)^2 s_2 & 0\\ 0 & -\left(s_1\right)^2 s_2\end{bmatrix}$$

which has a trace of $-2\left(s_1\right)^2 s_2$.

- and

$$B^{(1)}\left(\partial_{\theta_1}\left[B^{(1)}\right]^{-1}\right)B^{(1)}\left(\partial_{\theta_3}\left[B^{(1)}\right]^{-1}\right)B^{(1)}\partial_{\theta_2}\left[B^{(1)}\right]^{-1} = \begin{bmatrix}\left(s_1\right)^2 s_2 & 0\\ 0 & \left(s_1\right)^2 s_2\end{bmatrix}$$

which has a trace of $2\left(s_1\right)^2 s_2$.

- Thus all together we find that the integrand is equal to $3\times\left[-2\left(s_1\right)^2 s_2\right]-3\times\left[2\left(s_1\right)^2 s_2\right]=-12\left(s_1\right)^2 s_2$.

- Plugging this into the Cartan-Maurer integral we find:

$$\begin{aligned}\nu\left[B\left(\theta_1,\,\theta_2,\,\theta_3\right)^{(1)}\right] &= -\frac{1}{24\pi^2}\int_0^{\pi}d\theta_1\int_0^{\pi}d\theta_2\int_0^{2\pi}d\theta_3\left[-12\left(s_1\right)^2 s_2\right]\\ &= \frac{1}{2\pi^2}\int_0^{\pi}d\theta_1\int_0^{\pi}d\theta_2\int_0^{2\pi}d\theta_3\left[\sin\left(\theta_1\right)\right]^2\sin\left(\theta_2\right)\\ &= 1\end{aligned}$$

■

4.3.3.6 Claim

$\nu\left[U\left(x\right)V\left(x\right)\right]=\nu\left[U\left(x\right)\right]+\nu\left[V\left(x\right)\right]$ for any $\left\{U\left(x\right),\,V\left(x\right)\right\}\subset G^{S^3}$.

Proof Plugging into the formula the product we have:

$$\begin{aligned}
\nu[UV] &= -\frac{1}{24\pi^2}\int_0^\pi d\theta_1\int_0^\pi d\theta_2\int_0^{2\pi} d\theta_3\varepsilon_{ijk}tr\left(UV\left(\partial_i(UV)^{-1}\right)UV\left(\partial_j(UV)^{-1}\right)UV\partial_k(UV)^{-1}\right)\\
&= -\frac{1}{24\pi^2}\int_0^\pi d\theta_1\int_0^\pi d\theta_2\int_0^{2\pi} d\theta_3\varepsilon_{ijk}tr\left(V\left(\partial_iV^{-1}\right)V\left(\partial_jV^{-1}\right)V\left(\partial_kV^{-1}\right)\right)\\
&\quad -\frac{1}{24\pi^2}\int_0^\pi d\theta_1\int_0^\pi d\theta_2\int_0^{2\pi} d\theta_3\varepsilon_{ijk}tr\left(UV\left(\partial_iV^{-1}\right)V\left(\partial_jV^{-1}\right)\left(\partial_kU^{-1}\right)\right)\\
&\quad -\frac{1}{24\pi^2}\int_0^\pi d\theta_1\int_0^\pi d\theta_2\int_0^{2\pi} d\theta_3\varepsilon_{ijk}tr\left(V\left(\partial_iV^{-1}\right)\left(\partial_jU^{-1}\right)UV\left(\partial_kV^{-1}\right)\right)\\
&\quad -\frac{1}{24\pi^2}\int_0^\pi d\theta_1\int_0^\pi d\theta_2\int_0^{2\pi} d\theta_3\varepsilon_{ijk}tr\left(UV\left(\partial_iV^{-1}\right)\left(\partial_jU^{-1}\right)U\left(\partial_kU^{-1}\right)\right)\\
&\quad -\frac{1}{24\pi^2}\int_0^\pi d\theta_1\int_0^\pi d\theta_2\int_0^{2\pi} d\theta_3\varepsilon_{ijk}tr\left(\left(\partial_iU^{-1}\right)UV\left(\partial_jV^{-1}\right)V\left(\partial_kV^{-1}\right)\right)\\
&\quad -\frac{1}{24\pi^2}\int_0^\pi d\theta_1\int_0^\pi d\theta_2\int_0^{2\pi} d\theta_3\varepsilon_{ijk}tr\left(U\left(\partial_iU^{-1}\right)UV\left(\left(\partial_jV^{-1}\right)\right)\left(\partial_kU^{-1}\right)\right)\\
&\quad -\frac{1}{24\pi^2}\int_0^\pi d\theta_1\int_0^\pi d\theta_2\int_0^{2\pi} d\theta_3\varepsilon_{ijk}tr\left(\left(\partial_iU^{-1}\right)U\left(\partial_jU^{-1}\right)UV\left(\left(\partial_kV^{-1}\right)\right)\right)\\
&\quad -\frac{1}{24\pi^2}\int_0^\pi d\theta_1\int_0^\pi d\theta_2\int_0^{2\pi} d\theta_3\varepsilon_{ijk}tr\left(U\left(\partial_iU^{-1}\right)U\left(\partial_jU^{-1}\right)U\left(\partial_kU^{-1}\right)\right)\\
&= \nu[V]+\nu[U]+\text{remainder}
\end{aligned}$$

The remainder is zero (using cyclicity of the trace and the fact we are summing on ε_{ijk}, so $ijk = jki$ etc.):

$$\begin{aligned}
\text{traced remainder} &= -U\left(\partial_iV\right)\left(\partial_jV^{-1}\right)\left(\partial_kU^{-1}\right)\\
&\quad -U\left(\partial_kV\right)\left(\partial_iV^{-1}\right)\left(\partial_jU^{-1}\right)\\
&\quad -U\left(\partial_jV\right)\left(\partial_kV^{-1}\right)\left(\partial_iU^{-1}\right)\\
&\quad -V\left(\partial_iV^{-1}\right)\left(\partial_jU^{-1}\right)\partial_kU\\
&\quad -V\left(\partial_jV^{-1}\right)\left(\partial_kU^{-1}\right)\partial_iU\\
&\quad -V\left(\partial_kV^{-1}\right)\left(\partial_iU^{-1}\right)\partial_jU\\
&= -3U\left(\partial_iV\right)\left(\partial_jV^{-1}\right)\partial_kU^{-1}-3V\left(\partial_iV^{-1}\right)\left(\partial_jU^{-1}\right)\partial_kU\\
&= -3U\left(\partial_iV\right)\left(\partial_jV^{-1}\right)\partial_kU^{-1}-3V\left(\partial_jV^{-1}\right)\left(\partial_kU^{-1}\right)\partial_iU\\
&\quad -3U\left(\partial_iV\right)\left(\partial_jV^{-1}\right)\partial_kU^{-1}-3\left(\partial_iU\right)V\left(\partial_jV^{-1}\right)\left(\partial_kU^{-1}\right)\\
&\overset{ij/ik\text{ symmetric terms vanish}}{=} -3\partial_i\left[UV\left(\partial_jV^{-1}\right)\partial_kU^{-1}\right]
\end{aligned}$$

Integrating this and using the fundamental theorem of calculus, we cancel the derivative with the integral and get the difference of $UV\left(\partial_jV^{-1}\right)\partial_kU^{-1}$ between the two end points of the range of θ_i. Because we expect $UV\left(\partial_jV^{-1}\right)\partial_kU^{-1}$ to be continuous, the value on the two endpoints must be exactly the same and we get all together zero.

■

Conclusion Then we have $\nu\left[B\left(x\right)^{(n)}\right]=n\ \forall n\in\mathbb{Z}$. This follows from the fact that the winding number of a constant map is clearly zero, so $0=\nu\left[\mathbb{1}\right]=\nu\left[\left(B^{(1)}\right)^{-1}B^{(1)}\right]\equiv\nu\left[B^{(-1)}B^{(1)}\right]=\nu\left[B^{(-1)}\right]+\nu\left[B^{(1)}\right]$ so $\nu\left[B^{(-1)}\right]=-1$. Then we can get all the other integers using the above formula.

4.3.3.7 Claim

$B\left(x\right)^{(n_1)}$ is homotopic to $B\left(x\right)^{(n_2)}$ iff $n_1=n_2$.

Proof

- $\boxed{\Longrightarrow}$
  - We assume that $B\left(x\right)^{(n_1)}$ is homotopic to $B\left(x\right)^{(n_2)}$.
  - We have shown that the Cartan-Maurer integral invariant is stable under homotopies, so $\nu\left[B\left(x\right)^{(n_1)}\right]=\nu\left[B\left(x\right)^{(n_2)}\right]$.
  - But we have just proved before that $\nu\left[B\left(x\right)^{(n)}\right]=n\,\forall n\in\mathbb{Z}$, thus it follows that $n_1=n_2$.
- $\boxed{\Longleftarrow}$
  - Because $n_1=n_2$ we have that $\nu\left[B\left(x\right)^{(n_1)}\right]=\nu\left[B\left(x\right)^{(n_2)}\right]$.
  - But ν is stable under homotopies, so that means that $B\left(x\right)^{(n_1)}$ must be homotopic to $B\left(x\right)^{(n_2)}$.

■

4.3.3.8 Claim

$\forall\,V\left(x\right)\in SU\left(2\right)^{S^3}\ \exists\,n\in\mathbb{Z}$ such that $B\left(x\right)^{(n)}$ is homotopic to $V\left(x\right)$.

Proof Define $n:=\nu\left[V\left(x\right)\right]\in\mathbb{Z}$. Then by construction $B\left(x\right)^{(n)}$ is homotopic to $V\left(x\right)$, as they have the same winding number.
■

4.3.3.9 Claim

$\nu\left[V\left(x\right)\right]=\frac{1}{32\pi^2}\int_{\mathbb{R}^4}d^4x\left(F,\,\tilde{F}\right)$ where $\tilde{F}_{\mu\nu}\equiv\frac{1}{2}\varepsilon_{\mu\nu\lambda\sigma}F_{\lambda\sigma}$ is the Hodge dual of $F_{\mu\nu}$, and F is a field strength associated with some field configuration such that at $r\to\infty$, $A_\mu=V\partial_\mu V^{-1}$.

Note This formula is important because it gives us a way to compute the winding number of a *field configuration* rather than of a gauge group element which corresponds to a *pure gauge field configuration*. With this formula at hand, we can compute the winding number for *any* field configuration. For finite action field configurations, we are guaranteed the result would be some integer.

Proof

- Define $G_\mu := 2\varepsilon_{\mu\nu\lambda\sigma}\left(A_\nu,\, \partial_\lambda A_\sigma + \frac{2}{3}A_\lambda A_\sigma\right)$, conventionally called *the Chern-Simons current.*

- Note that we can also write $G_\mu = \varepsilon_{\mu\nu\lambda\sigma}\left(A_\nu,\, F_{\lambda\sigma}\right) - \frac{2}{3}\varepsilon_{\mu\nu\lambda\sigma}\left(A_\nu,\, A_\lambda A_\sigma\right)$.
  To see this:

$$\begin{aligned}
\varepsilon_{\mu\nu\lambda\sigma}\left(F_{\lambda\sigma} - \frac{2}{3}A_\lambda A_\sigma\right) &= \varepsilon_{\mu\nu\lambda\sigma}\left(A_{\sigma,\,\lambda} - A_{\lambda,\,\sigma} - A_\sigma A_\lambda + \frac{1}{3}A_\lambda A_\sigma\right) \\
&\overset{\varepsilon_{\mu\nu\lambda\sigma}\text{ is A.S.}}{=} \varepsilon_{\mu\nu\lambda\sigma}\left(A_{\sigma,\,\lambda} + A_{\sigma,\,\lambda} + A_\lambda A_\sigma + \frac{1}{3}A_\lambda A_\sigma\right) \\
&= \varepsilon_{\mu\nu\lambda\sigma}\left(2A_{\sigma,\,\lambda} + \frac{4}{3}A_\lambda A_\sigma\right)
\end{aligned}$$

  - If we calculate $\partial_\mu G_\mu$ we would obtain $\partial_\mu G_\mu = \left(F_{\mu\nu},\, \tilde{F}_{\mu\nu}\right)$:

$$\begin{aligned}
\partial_\mu G_\mu &= 2\varepsilon_{\mu\nu\lambda\sigma}\partial_\mu\left(A_\nu,\, \partial_\lambda A_\sigma + \frac{2}{3}A_\lambda A_\sigma\right) \\
&= -4\varepsilon_{\mu\nu\lambda\sigma}\partial_\mu tr\left(A_\nu\partial_\lambda A_\sigma + \frac{2}{3}A_\nu A_\lambda A_\sigma\right) \\
&= -4\varepsilon_{\mu\nu\lambda\sigma} tr\left[\partial_\mu\left(A_\nu\partial_\lambda A_\sigma\right) + \frac{2}{3}\partial_\mu\left(A_\nu A_\lambda A_\sigma\right)\right] \\
&= -4\varepsilon_{\mu\nu\lambda\sigma} tr\left[\left(\partial_\mu A_\nu\right)\partial_\lambda A_\sigma + \underbrace{\partial_{\mu\lambda}A_\sigma}_{\text{zero}}\right] \\
&\quad -\frac{8}{3}\varepsilon_{\mu\nu\lambda\sigma} tr\left[\left(\partial_\mu A_\nu\right)A_\lambda A_\sigma + A_\nu\left(\partial_\mu A_\lambda\right)A_\sigma + A_\nu A_\lambda\partial_\mu A_\sigma\right] \\
&\overset{\text{trace is cyclic}}{=} -4\varepsilon_{\mu\nu\lambda\sigma} tr\left[\left(\partial_\mu A_\nu\right)\partial_\lambda A_\sigma\right] \\
&\quad -\frac{8}{3}\varepsilon_{\mu\nu\lambda\sigma} tr\left[\left(\partial_\mu A_\nu\right)A_\lambda A_\sigma + \left(\partial_\mu A_\lambda\right)A_\sigma A_\nu + \left(\partial_\mu A_\sigma\right)A_\nu A_\lambda\right] \\
&= -4\varepsilon_{\mu\nu\lambda\sigma} tr\left[\left(\partial_\mu A_\nu\right)\partial_\lambda A_\sigma\right] - 8\varepsilon_{\mu\nu\lambda\sigma} tr\left[\left(\partial_\mu A_\nu\right)A_\lambda A_\sigma\right] \\
&= 2\varepsilon_{\mu\nu\lambda\sigma}\left[\left(A_{\nu,\,\mu},\, A_{\sigma,\,\lambda}\right) + 2\left(A_{\nu,\,\mu},\, A_\lambda A_\sigma\right)\right] \\
&= 2\varepsilon_{\mu\nu\lambda\sigma}\left[\left(A_{\nu,\,\mu},\, A_{\sigma,\,\lambda}\right) + \left(A_{\nu,\,\mu},\, A_\lambda A_\sigma\right) + \left(A_\mu A_\nu,\, A_{\sigma,\,\lambda}\right)\right] \\
&= 2\varepsilon_{\mu\nu\lambda\sigma}\left[\left(A_{\nu,\,\mu},\, A_{\sigma,\,\lambda}\right) + \left(A_{\nu,\,\mu},\, A_\lambda A_\sigma\right) + \left(A_\mu A_\nu,\, A_{\sigma,\,\lambda}\right) + \underbrace{\left(A_\mu A_\nu,\, A_\lambda A_\sigma\right)}_{\text{zero by cycl.}}\right] \\
&= 2\varepsilon_{\mu\nu\lambda\sigma}\left(A_{\nu,\,\mu} + A_\mu A_\nu,\, A_{\sigma,\,\lambda} + A_\lambda A_\sigma\right) \\
&= \frac{1}{2}\varepsilon_{\mu\nu\lambda\sigma}\left(A_{\nu,\,\mu} - A_{\mu,\,\nu} + A_\mu A_\nu - A_\nu A_\mu,\, A_{\sigma,\,\lambda} - A_{\lambda,\,\sigma} + A_\lambda A_\sigma - A_\sigma A_\lambda\right) \\
&= \frac{1}{2}\varepsilon_{\mu\nu\lambda\sigma}\left(F_{\mu\nu},\, F_{\lambda\sigma}\right) \\
&= \left(F_{\mu\nu},\, \tilde{F}_{\mu\nu}\right)
\end{aligned}$$

- Finally we have

$$\begin{aligned}
\int_{\mathbb{R}^4} d^4x \left(F,\, \tilde{F}\right) &= \int_{\mathbb{R}^4} d^4x \partial_\mu G_\mu \\
&\overset{\text{Stokes'}}{=} \int_{S^3 \text{ at infinity}} d^3 S \hat{r}_\mu G_\mu \\
&= \int_{S^3} d^3 S \hat{r}_\mu \left[\varepsilon_{\mu\nu\lambda\sigma} \left(A_\nu,\, F_{\lambda\sigma}\right) - \frac{2}{3}\varepsilon_{\mu\nu\lambda\sigma} \left(A_\nu,\, A_\lambda A_\sigma\right)\right] \\
&= \underbrace{\int_{S^3} d^3 S \hat{r}_\mu \left[\varepsilon_{\mu\nu\lambda\sigma} \left(A_\nu,\, F_{\lambda\sigma}\right)\right]}_{F\to 0 \text{ as } r\to\infty \text{ so this term is } 0} \\
&\quad -\frac{2}{3}\int_{S^3} d^3 S \hat{r}_\mu \varepsilon_{\mu\nu\lambda\sigma} \left(A_\nu,\, A_\lambda A_\sigma\right)
\end{aligned}$$

- Because the last term is evaluated at $r \to \infty$, we can safely assume that A is a pure gauge field configuration. Assume $V \in G^{S^3}$ is the element associated with A: $A_\mu = V \partial_\mu V^{-1}$. So we have:

$$\begin{aligned}
\int_{\mathbb{R}^4} d^4x \left(F,\, \tilde{F}\right) &= -\frac{2}{3}\int_{S^3} d^3 S \hat{r}_\mu \varepsilon_{\mu\nu\lambda\sigma} \left(V\left(\partial_\nu V^{-1}\right),\, V\left(\partial_\lambda V^{-1}\right) V \left(\partial_\sigma V^{-1}\right)\right) \\
&\cdots \\
&= \frac{2}{3}\int_0^\pi d\theta_1 \int_0^\pi d\theta_2 \int_0^{2\pi} d\theta_3 \varepsilon_{ijk} \left(V\left(\partial_i V^{-1}\right),\, V\left(\partial_j V^{-1}\right) V \partial_k V^{-1}\right) \\
&= -\frac{4}{3}\int_0^\pi d\theta_1 \int_0^\pi d\theta_2 \int_0^{2\pi} d\theta_3 \varepsilon_{ijk} tr \left(V\left(\partial_i V^{-1}\right) V\left(\partial_j V^{-1}\right) V \partial_k V^{-1}\right) \\
&= 32\pi^2 \nu\left[V\right]
\end{aligned}$$

■

4.3.4 The Bogomol'nyi Bound and (anti-) Self Dual Field Strengths

4.3.4.1 Claim

$S\left[A_\mu\right] \geq \frac{8\pi^2}{g^2}\left|\nu\left[A_\mu\right]\right|$ and equality is obtained when $F_{\mu\nu} = \pm\tilde{F}_{\mu\nu}$.

Proof 1

- $\widetilde{F}_{\mu\nu}\widetilde{F}_{\mu\nu} = \frac{1}{4}F_{\rho\lambda}\,\frac{1}{2}F_{\rho'\lambda'}\,2!\,2!\,\delta^{\rho\lambda}_{\rho'\lambda'} = F_{\mu\nu}F_{\mu\nu}$
- $\int d^4x \left(F_{\mu\nu},\, F_{\mu\nu}\right) = \left\{\int d^4x \left(F_{\mu\nu},\, F_{\mu\nu}\right) \int d^4x \left(\tilde{F}_{\mu\nu},\, \tilde{F}_{\mu\nu}\right)\right\}^{\frac{1}{2}}$
- But by the Schwartz inequality, $\left|\int d^4x \left(F_{\mu\nu},\, \tilde{F}_{\mu\nu}\right)\right|^2 \leq \int d^4x \left(F_{\mu\nu},\, F_{\mu\nu}\right) \int d^4x \left(\tilde{F}_{\mu\nu},\, \tilde{F}_{\mu\nu}\right)$.
- So we have that $\int d^4x \left(F_{\mu\nu},\, F_{\mu\nu}\right) \geq \left|\int d^4x Tr\left(F_{\mu\nu},\, \tilde{F}_{\mu\nu}\right)\right|$
- Thus $4g^2 S\left[A_\mu\right] \geq 32\pi^2 \left|\nu\left[A_\mu\right]\right|$.
- The Schwartz inequality is an equality iff $F_{\mu\nu} = \pm\tilde{F}_{\mu\nu}$.

■

Proof 2

- We have

$$\begin{aligned}
S[A_\mu] &\equiv \frac{1}{4g^2}\int_{\mathbb{R}^4} d^4x\,(F_{\mu\nu},\,F_{\mu\nu}) \\
&\equiv -2\frac{1}{4g^2}\int_{\mathbb{R}^4} d^4x\,tr\,[F_{\mu\nu}F_{\mu\nu}] \\
&= -2\frac{1}{4g^2}\int_{\mathbb{R}^4} d^4x\,tr\left[\pm F_{\mu\nu}\tilde{F}_{\mu\nu}+\frac{1}{2}F_{\mu\nu}F_{\mu\nu}+\frac{1}{2}\tilde{F}_{\mu\nu}\tilde{F}_{\mu\nu}\mp F_{\mu\nu}\tilde{F}_{\mu\nu}\right] \\
&= -2\frac{1}{4g^2}\int_{\mathbb{R}^4} d^4x\,tr\left[\pm F_{\mu\nu}\tilde{F}_{\mu\nu}+\frac{1}{2}\left(F_{\mu\nu}\mp\tilde{F}_{\mu\nu}\right)^2\right] \\
&= \pm\frac{1}{4g^2}\int_{\mathbb{R}^4} d^4x\left(F_{\mu\nu},\,\tilde{F}_{\mu\nu}\right)-\frac{1}{8g^2}\int_{\mathbb{R}^4} d^4x\,tr\left[\left(F_{\mu\nu}\mp\tilde{F}_{\mu\nu}\right)^2\right] \\
&= \pm\frac{8\pi^2}{g^2}\nu[A_\mu]-\frac{1}{4g^2}\int_{\mathbb{R}^4} d^4x\,tr\left[\left(F_{\mu\nu}\mp\tilde{F}_{\mu\nu}\right)^2\right]
\end{aligned}$$

- So when $F=\tilde{F}$ we have the extremal value of the action as $\frac{8\pi^2}{g^2}\nu[A_\mu]$, and when $F=-\tilde{F}$ we have the extremal value as $-\frac{8\pi^2}{g^2}\nu[A_\mu]$.

■

4.3.4.2 Conclusion

We conclude that if we found such field configurations for which $\boxed{F_{\mu\nu}=\pm\tilde{F}_{\mu\nu}}$, they would actually *solve the equations of motion*, because such configurations indeed extremize the action. This is good because this equation is a first order differential equation compared with the second order EoM. We also know what is the value of the action when it is minimal: $\boxed{S_0 := \frac{8\pi^2}{g^2}}$ (for nontrivial solutions).

4.4 The BPST Instanton

In 1975 Belavin, Polyakov, Schwarz and Tyupkin suggested the following solution to the (anti-) self-dual field strength equation in [3].

4.4.0.3 Claim

The following family of field strengths, parametrized by $\rho\in\mathbb{R}$:

$$A_\mu = \frac{|x|^2}{|x|^2+\rho^2}B(x)^{(1)}\partial_\mu\left\{\left[B(x)^{(1)}\right]^{-1}\right\}$$

where$B(x)^{(1)}\in SU(2)^{\mathbb{R}^4}$ is as defined above, fulfill the anti-self-dual condition.

Notes

- This solution is called *the BPST instanton* (of winding number 1).
- Since [3] has been published, solutions of higher winding number to the self-dual equation have been found, but they are not so useful in finding the vacua structure as explained in the discussion of the periodic well detailing the difference between 5 single-instantons versus one 5-instanton. Further discussion can be found in [4].
- ρ is called the size of the instanton. The existence of solutions of arbitrary sizes is a necessary consequence of the scale invariance of the classical field theory.
- Anti-instanton is obtained by replacing $B(x)^{(1)}$ with $B(x)^{(-1)}$. The anti-instanton will fulfill the self-dual condition whereas the instanton fulfills the self-dual condition.
- Observe how in the limit $|x| \to \infty$, this solution is indeed a pure gauge, and as we computed already (using the formula for the winding number in terms of F (which is in turn given in terms of A)), its winding number is thus exactly 1.
- We shall denote these solutions as $A_{inst,\,\rho,\,\mu}(x)$.
- We can also define a shifted instanton as $A_{inst,\,\rho,\,x_0,\,\mu}(x) := A_{inst,\,\rho,\,\mu}(x - x_0)$ which we can think of as merely having a "center" in spacetime at x_0 instead of 0.

Proof

- For a 4×4 anti-symmetric tensor, the self-dual condition can be formulated as three conditions:
  - For the 01 component we have

$$\begin{aligned} F_{01} &= \frac{1}{2}\varepsilon_{01\mu\nu}F_{\mu\nu} \\ &= \frac{1}{2}\varepsilon_{0123}F_{23} + \frac{1}{2}\varepsilon_{0132}F_{32} \\ &= F_{23} \end{aligned}$$

  - For the 02 component we have $F_{02} = -F_{13}$
  - For the 03 component we have $F_{03} = F_{12}$
  - All other conditions are redundant due to anti-symmetry of F, so all in all, in order to verify that $F_{\mu\nu}$ is self-dual, we must make sure that the following condition holds:

$$\boxed{\begin{cases} F_{01} = & F_{23} \\ F_{02} = & -F_{13} \\ F_{03} = & F_{12} \end{cases}}$$

- Note: this means that a self-dual 4×4 anti-symmetric tensor has merely 3 real parameters.

- We compute $F_{\mu\nu}$ step by the step:
  - Define $f_{\rho}(r) := \frac{2}{r+\rho^2}\ \forall r\in\mathbb{R}$.
  - Note that $2f_{\rho}'(r)+[f_{\rho}(r)]^2=0$:

$$2f_{\rho}'(r)+[f_{\rho}(r)]^2 \quad = \quad 2(-1)\frac{2}{(r+\rho^2)^2}+\frac{4}{(r+\rho^2)^2}=0$$

  - Define $s_{\mu} := \begin{bmatrix}\mathbb{1}\\ -i\vec{\sigma}\end{bmatrix}$.
  - Then $B(x)^{(1)}=\hat{x}_{\mu}s_{\mu}^{\dagger}$ and $\left[B(x)^{(1)}\right]^{-1}=\hat{x}_{\mu}s_{\mu}$.
  - And so we may write the BPST instanton as $A_{\mu}=f_{\rho}\left(|x|^2\right)\frac{1}{2}|x|^2\hat{x}_{\alpha}s_{\alpha}^{\dagger}\partial_{\mu}[\hat{x}_{\beta}s_{\beta}]$.
  - Observe that $\partial_{\mu}\hat{x}_{\nu}=\frac{1}{|x|}(\delta_{\mu\nu}-\hat{x}_{\mu}\hat{x}_{\nu})$.
  - So the BPST instanton is

$$\begin{aligned}A_{\mu} &= f_{\rho}\left(|x|^2\right)\frac{1}{2}|x|^2\hat{x}_{\alpha}s_{\alpha}^{\dagger}\partial_{\mu}[\hat{x}_{\beta}s_{\beta}]\\ &= f_{\rho}\left(|x|^2\right)\frac{1}{2}|x|^2\hat{x}_{\alpha}s_{\alpha}^{\dagger}s_{\beta}\frac{1}{|x|}(\delta_{\mu\beta}-\hat{x}_{\mu}\hat{x}_{\beta})\\ &= f_{\rho}\left(|x|^2\right)\frac{1}{2}|x|\left(\hat{x}_{\alpha}s_{\alpha}^{\dagger}s_{\mu}-\hat{x}_{\alpha}\hat{x}_{\mu}\hat{x}_{\beta}s_{\alpha}^{\dagger}s_{\beta}\right)\end{aligned}$$

  - Define $s_{\mu\nu} := \frac{1}{4i}\left(s_{\mu}s_{\nu}^{\dagger}-s_{\nu}s_{\mu}^{\dagger}\right)$ and $\bar{s}_{\mu\nu} := \frac{1}{4i}\left(s_{\mu}^{\dagger}s_{\nu}-s_{\nu}^{\dagger}s_{\mu}\right)$. These are closely related to $SO(4)$ and also to the 't Hooft symbols, as we shall see.
    * Note that these objects are anti-symmetric:
      · $s_{\nu\mu}=\frac{1}{4i}\left(s_{\nu}s_{\mu}^{\dagger}-s_{\mu}s_{\nu}^{\dagger}\right)=-s_{\mu\nu}$
      · $\bar{s}_{\nu\mu}=\frac{1}{4i}\left(s_{\nu}^{\dagger}s_{\mu}-s_{\mu}^{\dagger}s_{\nu}\right)=-\bar{s}_{\mu\nu}$
    * Note that $\bar{s}_{\mu\nu}=(-1)^{\delta_{0\mu}+\delta_{0\nu}}s_{\mu\nu}$:
      · When $\mu=0$ and $\nu\neq 0$, $\bar{s}_{0\nu}=\frac{1}{4i}\left(s_{0}^{\dagger}s_{\nu}-s_{\nu}^{\dagger}s_{0}\right)=\frac{1}{4i}(-i\sigma_{\nu}-i\sigma_{\nu})=-\frac{1}{4i}\left(s_{0}s_{\nu}^{\dagger}-s_{\nu}s_{0}^{\dagger}\right)=-s_{0\nu}$.
      · When $\mu\neq 0$ and $\nu=0$, $\bar{s}_{\mu 0}=-\bar{s}_{0\mu}=+s_{0\mu}=-s_{\mu 0}$.
      · When $\mu\neq 0$ and $\nu\neq 0$, $\bar{s}_{\mu\nu}=\frac{1}{4i}\left(s_{\mu}^{\dagger}s_{\nu}-s_{\nu}^{\dagger}s_{\mu}\right)=\frac{1}{4i}\left(s_{\mu}s_{\nu}^{\dagger}-s_{\nu}s_{\mu}^{\dagger}\right)=s_{\mu\nu}$.
    * Note that s is self dual:

· First condition for s:

$$\begin{aligned}
s_{01} &\equiv \frac{1}{4i}\left(s_0 s_1^\dagger - s_1 s_0^\dagger\right) \\
&= \frac{1}{4i}\left(\mathbb{1} i\sigma_1 - (-i\sigma_1)\,\mathbb{1}^\dagger\right) \\
&= \frac{1}{2}\sigma_1 \\
&= \frac{1}{4i}\left[\sigma_2,\, \sigma_3\right] \\
&= \frac{1}{4i}\left((-i\sigma_2)\, i\sigma_3 - (-i\sigma_3)\, i\sigma_2\right) \\
&= \frac{1}{4i}\left(s_2 s_3^\dagger - s_3 s_2^\dagger\right) \\
&\equiv s_{23}
\end{aligned}$$

· Second condition for s:

$$\begin{aligned}
s_{02} &\equiv \frac{1}{4i}\left(s_0 s_2^\dagger - s_2 s_0^\dagger\right) \\
&= \frac{1}{4i}\left(\mathbb{1} i\sigma_2 - (-i\sigma_2)\,\mathbb{1}^\dagger\right) \\
&= \frac{1}{2}\sigma_2 \\
&= -\frac{1}{4i}\left[\sigma_1,\, \sigma_3\right] \\
&= -\frac{1}{4i}\left((-i\sigma_1)\, i\sigma_3 - (-i\sigma_3)\, i\sigma_1\right) \\
&= -\frac{1}{4i}\left(s_1 s_3^\dagger - s_3 s_1^\dagger\right) \\
&\equiv -s_{13}
\end{aligned}$$

· Third condition for s:

$$\begin{aligned}
s_{03} &\equiv \frac{1}{4i}\left(s_0 s_3^\dagger - s_3 s_0^\dagger\right) \\
&= \frac{1}{4i}\left(\mathbb{1} i\sigma_3 - (-i\sigma_3)\,\mathbb{1}^\dagger\right) \\
&= \frac{1}{2}\sigma_3 \\
&= \frac{1}{4i}\left[\sigma_1,\, \sigma_2\right] \\
&= \frac{1}{4i}\left((-i\sigma_1)\, i\sigma_2 - (-i\sigma_1)\, i\sigma_2\right) \\
&= \frac{1}{4i}\left(s_1 s_2^\dagger - s_1 s_2^\dagger\right) \\
&\equiv s_{12}
\end{aligned}$$

$*$ and $\bar{s}$ is anti-self-dual:

· First condition for $\bar{s}$:

$$\begin{aligned}
\bar{s}_{01} &\equiv \frac{1}{4i}\left(s_0^\dagger s_1 - s_1^\dagger s_0\right) \\
&= \frac{1}{4i}\left(\mathbb{1}^\dagger\left(-i\sigma_1\right) - \left(i\sigma_1\right)\mathbb{1}\right) \\
&= -\frac{1}{2}\sigma_1 \\
&= -\frac{1}{4i}\left[\sigma_2,\, \sigma_3\right] \\
&= -\frac{1}{4i}\left(i\sigma_2\left(-i\sigma_3\right) - i\sigma_3\left(-i\sigma_2\right)\right) \\
&= -\frac{1}{4i}\left(s_2^\dagger s_3 - s_3^\dagger s_2\right) \\
&\equiv -\bar{s}_{23}
\end{aligned}$$

· Second condition for $\bar{s}$:

$$\begin{aligned}
\bar{s}_{02} &\equiv \frac{1}{4i}\left(s_0^\dagger s_2 - s_2^\dagger s_0\right) \\
&= \frac{1}{4i}\left(\mathbb{1}^\dagger\left(-i\sigma_2\right) - \left(i\sigma_2\right)\mathbb{1}\right) \\
&= -\frac{1}{2}\sigma_2 \\
&= \frac{1}{4i}\left[\sigma_1,\, \sigma_3\right] \\
&= \frac{1}{4i}\left(i\sigma_1\left(-i\sigma_3\right) - i\sigma_3\left(-i\sigma_1\right)\right) \\
&= \frac{1}{4i}\left(s_1^\dagger s_3 - s_3^\dagger s_1\right) \\
&\equiv \bar{s}_{13}
\end{aligned}$$

· Third condition for $\bar{s}$:

$$\begin{aligned}
\bar{s}_{03} &\equiv \frac{1}{4i}\left(s_0^\dagger s_3 - s_3^\dagger s_0\right) \\
&= \frac{1}{4i}\left(\mathbb{1}^\dagger\left(-i\sigma_3\right) - \left(i\sigma_3\right)\mathbb{1}\right) \\
&= -\frac{1}{2}\sigma_3 \\
&= -\frac{1}{4i}\left[\sigma_1,\, \sigma_2\right] \\
&= -\frac{1}{4i}\left(i\sigma_1\left(-i\sigma_2\right) - i\sigma_1\left(-i\sigma_2\right)\right) \\
&= -\frac{1}{4i}\left(s_1^\dagger s_2 - s_1^\dagger s_2\right) \\
&\equiv -\bar{s}_{12}
\end{aligned}$$

$*$ Note that $s_\mu s_\nu^\dagger = \delta_{\mu\nu}\mathbb{1} + 2is_{\mu\nu}$:

1. $s_\mu s_\mu^\dagger = \mathbb{1}$ because if $\mu \neq 0$ then $\sigma_i^2 = \mathbb{1}$, for all $\mu \in \{0,\, 1,\, 2,\, 3\}$.
2. $s_0 s_j^\dagger = i\sigma_j = \frac{1}{2}\left(i\sigma_j + i\sigma_j\right) = \frac{1}{2}\left(s_0 s_j^\dagger - s_j s_0^\dagger\right) = 2is_{0j}$ for all $j \in \{1,\, 2,\, 3\}$.

3. $s_j s_0^\dagger = -i\sigma_j = \frac{1}{2}(-i\sigma_j - i\sigma_j) = \frac{1}{2}\left(s_j s_0^\dagger - s_0 s_j^\dagger\right) = 2is_{j0}$ for all $j \in \{1, 2, 3\}$.
4. $s_i s_j^\dagger = \sigma_i\sigma_j = i\varepsilon_{ijk}\sigma_k = \frac{1}{2}(\varepsilon_{ijk} - \varepsilon_{jik})\sigma_k = \frac{1}{2}(\sigma_i\sigma_j - \sigma_j\sigma_i) = \frac{1}{2}\left(s_i s_j^\dagger - s_j s_i^\dagger\right) = 2is_{ij}$ for all $\{i, j\} \subset \{1, 2, 3\}$ such that $i \neq j$.

* and that $s_\mu^\dagger s_\nu = \delta_{\mu\nu}\mathbb{1} + 2i\bar{s}_{\mu\nu}$:
  1. $s_\mu^\dagger s_\mu = \mathbb{1}$ because if $\mu \neq 0$ then $\sigma_i^2 = \mathbb{1}$, for all $\mu \in \{0, 1, 2, 3\}$.
  2. $s_0^\dagger s_j = -i\sigma_j = \frac{1}{2}(-i\sigma_j - i\sigma_j) = \frac{1}{2}\left(s_0^\dagger s_j - s_j^\dagger s_0\right) = 2i\bar{s}_{0j}$ for all $j \in \{1, 2, 3\}$.
  3. $s_j^\dagger s_0 = i\sigma_j = \frac{1}{2}(i\sigma_j + i\sigma_j) = \frac{1}{2}\left(s_j^\dagger s_0 - s_0^\dagger s_j\right) = 2i\bar{s}_{j0}$ for all $j \in \{1, 2, 3\}$.
  4. $s_i^\dagger s_j = \sigma_i\sigma_j = \frac{1}{2}(\sigma_i\sigma_j - \sigma_j\sigma_i) = \frac{1}{2}\left(s_i^\dagger s_j - s_j^\dagger s_i\right) = 2i\bar{s}_{ij}$for all $\{i, j\} \subset \{1, 2, 3\}$ such that $i \neq j$.
* Similarly we find that $s_\mu s_\nu = \delta_{\mu\nu}\mathbb{1} - 2i\bar{s}_{\mu\nu}(-1)^{\delta_{\mu 0}}$:
  1. $s_\mu s_\mu = \mathbb{1}$ because if $\mu \neq 0$ then $\sigma_i^2 = \mathbb{1}$, for all $\mu \in \{0, 1, 2, 3\}$.
  2. $s_0 s_j = -i\sigma_j = \frac{1}{2}(-i\sigma_j - i\sigma_j) = \frac{1}{2}\left(s_0^\dagger s_j - s_j^\dagger s_0\right) = 2i\bar{s}_{0j}$ for all $j \in \{1, 2, 3\}$.
  3. $s_j s_0 = -i\sigma_j = -\frac{1}{2}(i\sigma_j + i\sigma_j) = -\frac{1}{2}\left(s_j^\dagger s_0 - s_0^\dagger s_j\right) = -2i\bar{s}_{j0}$ for all $j \in \{1, 2, 3\}$.
  4. $s_i s_j = -\sigma_i\sigma_j = -\frac{1}{2}(\sigma_i\sigma_j - \sigma_j\sigma_i) = -\frac{1}{2}\left(s_i^\dagger s_j - s_j^\dagger s_i\right) = -2i\bar{s}_{ij}$ for all $\{i, j\} \subset \{1, 2, 3\}$ such that $i \neq j$.
* and that $s_\mu^\dagger s_\nu^\dagger = \delta_{\mu\nu}\mathbb{1} - 2is_{\mu\nu}(-1)^{\delta_{\mu 0}}$:
  1. $s_\mu s_\mu = \mathbb{1}$ because if $\mu \neq 0$ then $\sigma_i^2 = \mathbb{1}$, for all $\mu \in \{0, 1, 2, 3\}$.
  2. $s_0^\dagger s_j^\dagger = i\sigma_j = \frac{1}{2}(i\sigma_j + i\sigma_j) = \frac{1}{2}\left(s_0 s_j^\dagger - s_j s_0^\dagger\right) = 2is_{0j}$ for all $j \in \{1, 2, 3\}$.
  3. $s_j^\dagger s_0^\dagger = i\sigma_j = -\frac{1}{2}(-i\sigma_j - i\sigma_j) = -\frac{1}{2}\left(s_j s_0^\dagger - s_0 s_j^\dagger\right) = -2is_{j0}$ for all $j \in \{1, 2, 3\}$.
  4. $s_i^\dagger s_j^\dagger = -\sigma_i\sigma_j = -\frac{1}{2}(\sigma_i\sigma_j - \sigma_j\sigma_i) = -\frac{1}{2}\left(s_i s_j^\dagger - s_j s_i^\dagger\right) = -2is_{ij}$ for all $\{i, j\} \subset \{1, 2, 3\}$ such that $i \neq j$.
* Finally we claim that $s_{\mu\nu} = \sum_{a=1}^{3}\frac{1}{2}\eta_{a\mu\nu}\sigma_a$ for all $\{\mu, \nu\} \subset \{0, 1, 2, 3\}$ where $\eta_{a\mu\nu}$ is the 't Hooft symbol, defined as $\eta_{a\mu\nu} \equiv \delta_{\mu 0}\delta_{\nu a} - \delta_{\nu 0}\delta_{\mu a} + \varepsilon_{0a\mu\nu}$:

$$\begin{aligned}\frac{1}{2}\eta_{a\mu\nu}\sigma_a &= \frac{1}{2}(\delta_{\mu 0}\delta_{\nu a} - \delta_{\nu 0}\delta_{\mu a} + \varepsilon_{0a\mu\nu})\sigma_a \\ &= \frac{1}{2}(\delta_{\mu 0}\sigma_\nu - \delta_{\nu 0}\sigma_\mu + \varepsilon_{0a\mu\nu}\sigma_a)\end{aligned}$$

  · If $\mu = \nu$ we get 0, which is good as $s_{00} = 0$.
  · If $\mu = 0$ and $\nu \neq 0$ we get $\frac{1}{2}\sigma_\nu = \frac{1}{4i}(i\sigma_\nu + i\sigma_\nu) = \frac{1}{4i}\left(s_0 s_\nu^\dagger - s_\nu s_0^\dagger\right) = s_{0\nu}$

· If $\mu \neq 0$ and $\nu = 0$ we get $-\frac{1}{2}\sigma_\mu = -s_{0\mu} = s_{\mu 0}$.

· If $\mu \neq 0$ and $\nu \neq 0$ and $\mu \neq \nu$ we get $\frac{1}{2}\varepsilon_{0a\mu\nu}\sigma_a = \frac{1}{2i}\sigma_\mu\sigma_\nu = s_{\mu\nu}$.

– Using s and $\bar{s}$ we find that the BPST instanton is equal to:

$$
\begin{aligned}
A_\mu &= f_\rho\left(|x|^2\right)\frac{1}{2}|x|\left(\hat{x}_\alpha s^\dagger_\alpha s_\mu - \hat{x}_\alpha\hat{x}_\mu\hat{x}_\beta s^\dagger_\alpha s_\beta\right)\\
&= f_\rho\left(|x|^2\right)\frac{1}{2}|x|\left(\hat{x}_\alpha\left(\delta_{\alpha\mu}\mathbb{1} + 2i\bar{s}_{\alpha\mu}\right) - \hat{x}_\alpha\hat{x}_\mu\hat{x}_\beta\left(\delta_{\alpha\beta}\mathbb{1} + 2i\bar{s}_{\alpha\beta}\right)\right)\\
&= f_\rho\left(|x|^2\right)\frac{1}{2}|x|\left(2i\hat{x}_\alpha\bar{s}_{\alpha\mu} - 2i\hat{x}_\mu\underbrace{\hat{x}_\alpha\hat{x}_\beta\bar{s}_{\alpha\beta}}_{0\text{ by anti-symmetry}}\right)\\
&= f_\rho\left(|x|^2\right)ix_\alpha\bar{s}_{\alpha\mu}
\end{aligned}
$$

– Then

$$
\begin{aligned}
\partial_\mu A_\nu &= \partial_\mu\left[f_\rho\left(|x|^2\right)x_\alpha i\bar{s}_{\alpha\nu}\right]\\
&= f'_\rho\left(|x|^2\right)\left(\partial_\mu|x|^2\right)x_\alpha i\bar{s}_{\alpha\nu} + f_\rho\left(|x|^2\right)i\bar{s}_{\mu\nu}\\
&= 2f'_\rho\left(|x|^2\right)x_\mu x_\alpha i\bar{s}_{\alpha\nu} + f_\rho\left(|x|^2\right)i\bar{s}_{\mu\nu}\\
&= \left[2f'_\rho\left(|x|^2\right)x_\mu x_\alpha + f_\rho\left(|x|^2\right)\delta_{\alpha\mu}\right]i\bar{s}_{\alpha\nu}
\end{aligned}
$$

- Thus we find:

$$
\begin{aligned}
F_{\mu\nu} &\equiv \partial_\mu A_\nu - \partial_\nu A_\mu + A_\mu A_\nu - A_\nu A_\mu\\
&= \left[2f'_\rho\left(|x|^2\right)x_\mu x_\alpha + f_\rho\left(|x|^2\right)\delta_{\alpha\mu}\right]i\bar{s}_{\alpha\nu}\\
&\quad -\left[2f'_\rho\left(|x|^2\right)x_\nu x_\alpha + f_\rho\left(|x|^2\right)\delta_{\alpha\nu}\right]i\bar{s}_{\alpha\mu}\\
&\quad +f_\rho\left(|x|^2\right)ix_\alpha\bar{s}_{\alpha\mu}f_\rho\left(|x|^2\right)ix_\beta\bar{s}_{\beta\nu}\\
&\quad -f_\rho\left(|x|^2\right)ix_\beta\bar{s}_{\beta\nu}f_\rho\left(|x|^2\right)ix_\alpha\bar{s}_{\alpha\mu}\\
&= 2f'_\rho\left(|x|^2\right)2x_\alpha i\bar{s}_{\alpha[\nu}x_{\mu]}\\
&\quad +\left[f_\rho\left(|x|^2\right)\right]^2 x_\alpha x_\beta\left[i\bar{s}_{\alpha\mu},\, i\bar{s}_{\beta\nu}\right]\\
&\quad +f_\rho\left(|x|^2\right)2i\bar{s}_{\mu\nu}
\end{aligned}
$$

- Using the 't Hooft symbols (the rules of which can be found in the appendix of [10]) we can easily compute the awkward commutator. First for s:

$$\begin{aligned}
[is_{\alpha\mu},\, is_{\beta\nu}] &= [-is_{\alpha\mu},\, -is_{\beta\nu}] \\
&= \left[\eta_{a\alpha\mu}\frac{\sigma_a}{2i},\, \eta_{b\beta\nu}\frac{\sigma_b}{2i}\right] \\
&= -\frac{1}{4}\eta_{a\alpha\mu}\eta_{b\beta\nu}2i\varepsilon_{abc}\sigma_c \\
&= \frac{1}{2i}\sigma_c\left(\varepsilon_{cab}\eta_{a\alpha\mu}\eta_{b\beta\nu}\right) \\
&= \frac{1}{2i}\sigma_c\left(\delta_{\alpha\beta}\eta_{c\mu\nu} - \delta_{\alpha\nu}\eta_{c\mu\beta} - \delta_{\mu\beta}\eta_{c\alpha\nu} + \delta_{\mu\nu}\eta_{c\alpha\beta}\right) \\
&= \left(-i\delta_{\alpha\beta}s_{\mu\nu} + i\delta_{\alpha\nu}s_{\mu\beta} + i\delta_{\mu\beta}s_{\alpha\nu} - i\delta_{\mu\nu}s_{\alpha\beta}\right)
\end{aligned}$$

- And so for our actual expression we have:

$$\begin{aligned}
[i\bar{s}_{\alpha\mu},\, i\bar{s}_{\beta\nu}] &= \left[i\left(-1\right)^{\delta_{\alpha 0}+\delta_{\mu 0}} s_{\alpha\mu},\, i\left(-1\right)^{\delta_{\beta 0}+\delta_{\nu 0}} s_{\beta\nu}\right] \\
& \quad \left(-1\right)^{\delta_{\alpha 0}+\delta_{\mu 0}+\delta_{\beta 0}+\delta_{\nu 0}}\left[is_{\alpha\mu},\, is_{\beta\nu}\right] \\
&= \left(-1\right)^{\delta_{\alpha 0}+\delta_{\mu 0}+\delta_{\beta 0}+\delta_{\nu 0}}\left(-i\delta_{\alpha\beta}s_{\mu\nu} + i\delta_{\alpha\nu}s_{\mu\beta} + i\delta_{\mu\beta}s_{\alpha\nu} - i\delta_{\mu\nu}s_{\alpha\beta}\right) \\
&= -i\left(-1\right)^{\delta_{\alpha 0}+\delta_{\beta 0}}\delta_{\alpha\beta}\bar{s}_{\mu\nu} + i\left(-1\right)^{\delta_{\alpha 0}+\delta_{\nu 0}}\delta_{\alpha\nu}\bar{s}_{\mu\beta} \\
& \quad +i\left(-1\right)^{\delta_{\mu 0}+\delta_{\beta 0}}\delta_{\mu\beta}\bar{s}_{\alpha\nu} - i\left(-1\right)^{\delta_{\mu 0}+\delta_{\nu 0}}\delta_{\mu\nu}s_{\alpha\beta}
\end{aligned}$$

- When we sum these with $x_\alpha x_\beta$ we get:

$$\begin{aligned}
x_\alpha x_\beta\left[i\bar{s}_{\alpha\mu},\, i\bar{s}_{\beta\nu}\right] &= -i\left|x\right|^2\bar{s}_{\mu\nu} + ix_\nu x_\beta\bar{s}_{\mu\beta} + x_\alpha x_\mu i\bar{s}_{\alpha\nu} \\
&= -i\left|x\right|^2\bar{s}_{\mu\nu} - i2x_\alpha\bar{s}_{\alpha[\mu}x_{\nu]}
\end{aligned}$$

- So we find that the field strength tensor is:

$$\begin{aligned}
F_{\mu\nu} &= 2f'_\rho\left(|x|^2\right)2x_\alpha i\bar{s}_{\alpha[\nu}x_{\mu]} + f_\rho\left(|x|^2\right)2i\bar{s}_{\mu\nu} \\
& \quad +\left[f_\rho\left(|x|^2\right)\right]^2\left(-i\left|x\right|^2\bar{s}_{\mu\nu} - i2x_\alpha\bar{s}_{\alpha[\mu}x_{\nu]}\right) \\
&= 2ix_\alpha\bar{s}_{\alpha[\nu}x_{\mu]}\underbrace{\left[2f'_\rho\left(|x|^2\right) + \left[f_\rho\left(|x|^2\right)\right]^2\right]}_{\text{zero}} \\
& \quad -i\bar{s}_{\mu\nu}\left[|x|^2\left[f_\rho\left(|x|^2\right)\right]^2 - 2f_\rho\left(|x|^2\right)\right] \\
&= -i\bar{s}_{\mu\nu}\left[|x|^2\left(\frac{2}{|x|^2+\rho^2}\right)^2 - 2\frac{2}{r+\rho^2}\right] \\
&= 4i\frac{\rho^2}{\left(|x|^2+\rho^2\right)^2}\bar{s}_{\mu\nu}
\end{aligned}$$

- In conclusion, we found that $F_{\mu\nu}$ has the same tensor structure as $\bar{s}_{\mu\nu}$, which we proved earlier is anti-self-dual, and so, we have just proven that $\boxed{F_{\mu\nu} = -\tilde{F}_{\mu\nu}}$.

■

4.4.1 Collective Coordinates

In the double well, we had one "collective coordinate", $\mathcal{T}$, a parameter of the instanton solution which the action was invariant under. It was crucial to identify that coordinate in order to compute the path integral. In our case, *there are* 8 *different collective coordinates* for the instanton solutions we have just found.

4.4.1.1 Claim

$S\left[A_{inst,\,\rho_1,\,\mu}\right] = S\left[A_{inst,\,\rho_2,\,\mu}\right]$ (scale invariance). This gives us 1 collective coordinate.

Note The classical Yang-Mills Lagrangian is scale-invariant, as $\nexists$ a dimensionful parameter.

Proof The minimal action which we found, $S_0 := \frac{8\pi^2}{g^2}$, is independent of ρ. $\blacksquare$

4.4.1.2 Claim

$S\left[A_{inst,\,\rho,\,x_0,\,\mu}\right] = S\left[A_{inst,\,\rho,\,\mu}\right]$ for all $x_0 \in \mathbb{R}^4$ (translation invariance). This gives us 4 collective coordinates.

Proof Any volume integral must be translation invariant. $\blacksquare$

4.4.1.3 Claim

The action is invariant a *global* $SU(2)$ gauge transformation on $A_{inst,\,\rho,\,x_0\,\mu}$. This gives us 3 collective coordinates (as many as there are generators of $SU(2)$).

Proof If we perform a global constant transformation on $A_{inst,\,\rho,\,x_0\,\mu}$ we would get: $A_{inst,\,\rho,\,x_0\,\mu} \mapsto V A_{inst,\,\rho,\,x_0\,\mu} V^{-1} + \underbrace{V \partial_\mu V^{-1}}_{0 \text{ because } V \text{ is const.}}$, where $V \in SU(2)$.
Taking any other element, $U \in SU(2)$, because $A_{inst,\,\rho,\,x_0\,\mu}$ is asymptotically a pure gauge, we may also write:

$$\begin{aligned} V A_\mu(x) V^{-1} &= f_\rho\left(|x|^2\right) \frac{1}{2} |x|^2 V B(x)^{(1)} \partial_\mu \left[B(x)^{(1)}\right]^{-1} V^{-1} \\ &= f_\rho\left(|x|^2\right) \frac{1}{2} |x|^2 V B(x)^{(1)} U^{-1} \partial_\mu \left[V B(x)^{(1)} U^{-1}\right]^{-1} \end{aligned}$$

And so we see that really what we have by this constant gauge transformation of V on A is $B(x)^{(1)} \mapsto V B(x)^{(1)} U^{-1}$.

Now, due to the isomorphism between $SO(4) \simeq SU(2) \otimes SU(2)$, under a general rotation of the instanton, we have $B(x)^{(1)} \mapsto V B(x)^{(1)} U^{-1}$ where $\{V, U\} \subset SU(2)$ and are determined by the particular rotation of $\mathbb{R}^4$ we pick. So we could pick exactly the right rotation of $\Lambda \in SO(4)$ so that $B(\Lambda x)^{(1)} = V B(x)^{(1)} U^{-1}$, and thus, effectively, by redefining our chart on $\mathbb{R}^4$ undo this constant gauge transformation, and get back $A_{inst,\,\rho,\,x_0\,\mu}$.
$\blacksquare$

4.4.1.4 Claim

$\nexists$ other invariants for the action (via an index theorem by Atiyah, Ward in [1] which determines the dimension of the modulo space of $SU(2)$ as exactly 8).

4.4.1.5 Conclusion

As before, we would need to integrate over these coordinates directly, when computing the contribution of these solutions to the path integral.

4.5 Finding the Vacuum

Following [4],we work in an axial gauge, in which $A_3 \overset{!}{=} 0$. (Then the path-integral formulation is equivalent to the canonical quantization, and there is no need for ghost fields or extra conditions on the space of states.)

We work in a spacetime box of spatial volume V from time $-\frac{T}{2}$ to time $\frac{T}{2}$. Eventually we will send $V \to \infty$ and $T \to \infty$. We employ boundary conditions on the three-dimensional boundary of the box at times $\pm\frac{T}{2}$ such that the tangential term (to the surface of the box) of A_μ is *constant*. Then the surface term of δS will be zero.

This constant, however, is not arbitrary. It must obey the following conditions to maintain consistency:

1. $A_3 \overset{!}{=} 0$ gauge must be respected.
2. At infinity we should have finite-action field configurations. Since only the tangential component of A_μ determines the winding number (...), this means that spacetime will be filled with field configurations of a definite winding number.

As $V \to \infty$, the definiteness of the winding number (which follows from the finiteness of the action) is the only specific feature that remains of the boundary conditions. So in the path-integral we can forget about the boundary conditions and simply add a delta function for field configurations that have a definite integer winding number. In this way we will clearly obtain only finite action field configurations:
$P(V, T, n) := \mathcal{N} \int \mathcal{D}A_1 \mathcal{D}A_2 \mathcal{D}A_4 e^{-S[A_\mu]} \delta(n - \nu[A_\mu])$ for some $n \in \mathbb{Z}$.

4.5.0.6 Claim

For large T_1 and T_2, $P(V, T_1 + T_2, n) = \sum_{n_1+n_2=n} P(V, T_1, n_1) P(V, T_2, n_2)$

Proof Follows from the expression $\nu[A_\mu] = \frac{1}{32\pi^2} \int d^4x \left(F, \tilde{F}\right)$, the winding number as a *local* density.
■

4.5.1 The θ-Vacua

This composition law is not what we would expect from a transition matrix element that has a contribution from only a single energy eigenstate. In order to get the composition law we want, $e^{-E_i(T_1+T_2)}$, we make a Fourier transform of $P(V,\,T,\,n)$:

$$\begin{aligned}\tilde{P}(V,\,T,\,\theta) &:= \sum_{n\in\mathbb{Z}} e^{in\theta} P(V,\,T,\,n) \\ &= \mathcal{N}\int \mathcal{D}A_\mu e^{-S[A_\mu]} e^{i\nu[A_\mu]\theta}\end{aligned}$$

As a result of the previous composition law now we have $\tilde{P}(V,\,T_1+T_2,\,\theta) = \tilde{P}(V,\,T_1,\,\theta)\,\tilde{P}(V,\,T_2,\,\theta)$. So $\tilde{P}$ must be proportional to $\left\langle e^{-\hat{H}T}\right\rangle$ in *some* energy eigenstate. We naturally label these eigenstates with θ and as before call them the θ-vacua:

$$\left\langle \theta \left| e^{-\hat{H}T} \right| \theta \right\rangle = \mathcal{N}' \int \mathcal{D}A_\mu e^{-S[A_\mu]} e^{i\nu[A_\mu]\theta} \tag{4.1}$$

The conclusion is that our theory is split into disconnected sectors labelled by θ, each with its own vacuum. Naively, we could have obtained the same result by merely postulating an extra term in the Lagrangian proportional to $\nu[A_\mu] \sim \int d^4x \left(F,\,\tilde{F}\right)$. This term was, in fact, only rejected to begin with because it violates CP ($\left(F,\,\tilde{F}\right) \sim \vec{E}\cdot\vec{B}$ and $\vec{B}$ doesn't change sign under P), but otherwise it is just as good as $\int d^4x\,(F,\,F)$. In addition, we found it is a total divergence, and so should have no effect on the EoMs. But there seems to be an effect to it none the less, which is not classical.

4.5.2 Dilute Instanton Gas

Just as in the periodic well, we build approximate solutions which consist of n instantons and n' anti-instantons, where their centers x_0 are integrated over. We then sum over all such possible configurations:

$$\begin{aligned}\left\langle \theta \left| e^{-\hat{H}T} \right| \theta \right\rangle &\propto \sum_{(n,\,n')\in\mathbb{N}^2} \frac{1}{n!}\left[\left(Ke^{-S_0}\right)VT\right]^n \frac{1}{n'!}\left[\left(Ke^{-S_0}\right)VT\right]^{n'} e^{i(n-n')\theta} \\ &= \exp\left\{2KVTe^{-S_0}\cos(\theta)\right\}\end{aligned}$$

where $S_0 = \frac{8\pi^2}{g^2}$, V is volume and K is some constant which can be computed by calculating the infinite product of eigenvalues of a corresponding differential operator (see [12]). In general K will contain an infrared "embarrassment"–a divergence–but fortunately it only diverges when we assume our approximation is not valid. From this we can read off the energy of $|\theta\rangle$:

$$\boxed{\frac{E(\theta)}{V} = -2Ke^{-S_0}\cos(\theta)}$$

Because the energy and also the vacuum expectation value depend non-trivially on θ we must conclude that all the θ states are in fact distinct!

4.5.3 Other Gauge Groups

4.5.3.1 Claim

Every simple Lie group contains a subgroup isomorphic to $SU(2)$.

- For example, $\begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is a general element of such a subgroup of $SU(3)$ where $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SU(2)$.

4.5.3.2 Notes

$\exists$ a theorem due to Raoul Bott saying that if G is any simple Lie group, $H \leq G$ such that $H \simeq SU(2)$, then any element of G^{S^3} is homotopic with some element of H^{S^3}. Then we can consider the same BPST solutions we have considered where $A_\mu = a \cdot {A_\mu}^{(H)} + b \cdot {A_\mu}^{(G \backslash H)}$ and we would take $b = 0$.

- For example, for $SU(3)$, which has 8 gauge fields, the following is an instanton solution: three of the fields are just as the $SU(2)$ instanton, and the remaining five are zero.
- This is the only instanton solution of $SU(3)$ with winding number 1.
- Then there would be 12, and not 8 collective coordinates (the action is invariant under a global $SU(3)$ transformation, so 8 coordinates instead of the 3 of $SU(2)$, but one of the generators commutes with $SU(2)$, so we are left with total of 7 collective coordinates from the global $SU(3)$ transformation)

Chapter 5

The Strong CP Probelm

As we have seen, the instanton solution effectively creates a θ term proportional to $\theta \int d^4x \left(F,\, \tilde{F}\right)$ in the Lagrangian, which, if $\theta \neq 0$, violates CP.

Using the CPT theorem we conclude that T is violated.

According to experimental measurements using the electric dipole moment of the neutron ($d_n \approx \theta e \frac{m_\pi^2}{m_n^2}$), the amount of T violation corresponds to $\theta < 10^{-5}$ ([14]) or even $\theta < 10^{-9}$. This raises a fine-tuning question, which is known as the *strong CP problem*. To explain this fine tuning we need to go beyond the standard model.

One couldn't just throw away the instanton concept, because it does solve the $U(1)$ problem: The non-observation in experiments of a $U(1)$ axial symmetry which is expected in QCD. It was thought to come from spontaneous symmetry breaking, but no corresponding Goldstone Boson was found. Finally it was explained by 't Hooft in [11] that this symmetry is anomalous and the instanton solution fits perfectly to explain how. So we definitely need the instantons.

5.1 Peccei–Quinn theory

Following [7], $\exists$ three approaches to explaining the value of θ:

1. Unconventional dynamics.
2. Spontaneously broken CP.
3. An additional chiral symmetry.

Peccei employs the third approach.

This chiral symmetry can arise from assuming $m_u = 0$ (up quark) which is inconsistent with experimental data. But if that were the case, we could perform a global chiral rotation $\psi_f \mapsto e^{i\alpha_f \gamma_5}\psi_f$. The change in the path integral measure introduces a term proportional to $\exp\left\{-\frac{ig^2}{16\pi^2}\alpha_f \int d^4x \left(F,\, \tilde{F}\right)\right\}$. So by picking α_f properly we could eliminate the θ term that comes from the instantons. However, as has been said, when we do this, we introduce a phase to the mass of the f Fermion, $m_f \mapsto e^{-i\alpha_f} m_f$. This could have only worked if we had one quark which is massless.

Alternatively the chiral symmetry can arise from an additional global $U(1)$ chiral symmetry, $U(1)_{PQ}$. This symmetry is then to be spontaneously broken. Its introduction into the theory replaces the static θ term in the effective Lagrangian with a dynamical CP conserving field which has come to be known as *the axion*–the Goldstone Boson of the broken $U(1)_{PQ}$ symmetry. This symmetry could not have been exact because the axion cannot be exactly massless.

Thus the θ-term in our Lagrangian becomes:

$$\mathcal{L} = \frac{\theta}{32\pi^2}\left(F,\,\tilde{F}\right) \mapsto \frac{1}{2}\left(\partial_\mu a\right)\left(\partial^\mu a\right) + \frac{\frac{a}{M}+\theta}{32\pi^2}\left(F,\,\tilde{F}\right)$$

where a is the dynamical axion field and M is the mass scale at which it appears. Then by an opportune shift in the axion field $a \mapsto a - \theta M$ we can get rid of the θ term. If the axion is very light ($\sim 1eV$), the cut-off scale at which it appears is very low.

We can also add interaction terms for the axion with the quarks, for instance, for the up quark, a term of the form $-i\frac{f_u}{M}\left(\partial_\mu a\right)\bar{u}\gamma_5\gamma^\mu u$ where f_u is the coupling constant for the interaction. Following the very same procedure of chiral perturbation theory we can construct an effective Lagrangian for the pion-axion interaction.

Bibliography

[1] M. F. Atiyah, N. J. Hitchin, and I. M. Singer. Self-duality in four-dimensional Riemannian geometry. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 362(1711):425–461, 1978.

[2] Jacques Bair, Piotr Blaszczyk, Robert Ely, Valérie Henry, Vladimir Kanovei, Karin U Katz, Mikhail G Katz, Semen S Kutateladze, Thomas McGaffey, David M Schaps, et al. Is mathematical history written by the victors? *arXiv preprint arXiv:1306.5973*, 2013.

[3] A. A. Belavin, A. Polyakov, A. S. M. Schwartz, and Yu. S. Tyupkin. Pseudoparticle solutions of the Yang-Mills equations. *Physics Letters B*, 59(1):85–87, 10 1975.

[4] Sidney Coleman. *Aspects of symmetry : selected Erice lectures of Sidney Coleman.* Cambridge University Press, Cambridge England New York, 1988.

[5] L. D. Landau and L. M. Lifshitz. *Quantum Mechanics Non-Relativistic Theory, Third Edition: Volume 3.* Butterworth-Heinemann, 3 edition, January 1981.

[6] James Munkres. *Topology.* Pearson, 2 edition, January 2000.

[7] R.D. Peccei. The Strong CP problem and axions. *Lect.Notes Phys.*, 741:3–17, 2008.

[8] Michael E. Peskin and Dan V. Schroeder. *An Introduction To Quantum Field Theory (Frontiers in Physics)*, chapter 9. Functional Methods. Westview Press, 1995.

[9] Jun John Sakurai. *Modern Quantum Mechanics.* Addison-Wesley Publishing Company, Reading, MA, Revised edition, 1994.

[10] Gerard 't Hooft. Computation of the Quantum Effects Due to a Four-Dimensional Pseudoparticle. *Phys.Rev.*, D14:3432–3450, 1976.

[11] Gerard 't Hooft. How Instantons Solve the U(1) Problem. *Phys.Rept.*, 142:357–387, 1986.

[12] A.I. Vainshtein, Valentin I. Zakharov, V.A. Novikov, and Mikhail A. Shifman. ABC's of Instantons. *Sov.Phys.Usp.*, 25:195, 1982.

[13] George W. Whitehead. *Elements of homotopy theory / George W. Whitehead*. Springer-Verlag New York, 1978.

[14] F. Wilczek. Problem of strong P and T invariance in the presence of instantons. *Phys. Rev. Lett.*, 40:279–282, Jan 1978.