

Faculty of physics

Advanced Quantum Field Theory 2

Notes of lectures by Prof. Martin Beneke, Ph.D.

Academic year 2025/2026
Winter semester

Author:

Simon Bukovšek

Last change: February 17, 2026

Contents

1 Supersymmetry **4**
- 1.1 Algebraic basics and supersymmetry multiplets . . . 4
  - 1.1.1 Algebra of supersymmetry . . . 6
  - 1.1.2 SUSY multiplets, one-particle states . . . 8
- 1.2 Supersymmetric QFT ($\mathcal{N} = 1$) . . . 9
  - 1.2.1 Wess–Zumino model . . . 9
- 1.3 Supersymmetric gauge theory . . . 12
- 1.4 Supersymmetry breaking . . . 14
- 1.5 Minimally supersymmetric Standard model . . . 15

2 Spin-2 fields **21**
- 2.1 The massive spin-2 field . . . 21
- 2.2 The massless spin-2 field . . . 23
- 2.3 GR as an effective field theory of gravitation . . . 24
  - 2.3.1 Classical general relativity . . . 24
  - 2.3.2 Quantization of weak gravitational fields . . . 26
  - 2.3.3 Interaction with matter . . . 28
  - 2.3.4 Renormalization . . . 30
- 2.4 van Dam–Veltman–Zakharov (vDVZ) discontinuity . . . 32
  - 2.4.1 Tree level and 1-loop considerations . . . 32
  - 2.4.2 Significance of the vDVZ discontinuity . . . 34

3 QFT at finite temperature **38**
- 3.1 Non-interacting gas . . . 39
  - 3.1.1 Path representation of the partition function . . . 39
  - 3.1.2 Free, real scalar field . . . 40
  - 3.1.3 Free charged scalar and Bose–Einstein condensation . . . 42
  - 3.1.4 Kubo–Martin–Schwinger relation . . . 44
  - 3.1.5 Fermions . . . 44
- 3.2 Interactions . . . 45
  - 3.2.1 Propagators, Feynman rules and renormalization . . . 45
  - 3.2.2 $\mathcal{O}(\lambda)$ correction to $\ln(Z)$ and the two-point function . . . 47
- 3.3 IR divergences and resummation . . . 49
  - 3.3.1 Effective field theory for the zero mode . . . 51
- 3.4 Fermions, sums and loops . . . 52
- 3.5 QCD at finite temperature . . . 52

3.5.1 Path integral representation . . . 52
3.5.2 Gluon self-energy . . . 54
3.5.3 Electrostatic and magnetic QCD . . . 57
3.6 Phase transitions and symmetry restoration at finite temperature . . . 59
3.7 Real-time formalism . . . 61
3.7.1 Linear response . . . 61
3.7.2 Green's functions and spectral function . . . 62
3.7.3 Free theory . . . 65
3.7.4 Schwinger–Keldysh (closed-time-path) formalism . . . 65
3.8 Particle production/decay in a medium . . . 68
3.9 The hard-thermal loop (HTL) approximation . . . 69
3.10 Out-of-equilibrium kinetic theory from QFT . . . 71

4 Topological field configurations **74**
4.1 Solitons in one (space) dimension (the kink) . . . 74
4.1.1 The kink solution . . . 75
4.1.2 Quantum corrections . . . 78
4.1.3 Derrick's theorem . . . 79
4.2 Localized, static, topological, finite energy, classical field configurations in more than one space dimension . . . 81
4.2.1 Vacuum manifold and finite-energy solutions . . . 81
4.2.2 Homotopy . . . 82
4.2.3 Vortices in $d = 1 + 2$ dimensions . . . 83
4.2.4 Magnetic monopoles in $d = 3 + 1$ . . . 85
4.3 Vacuum structure of Yang–Mills theory . . . 87
4.4 Instantons and tunneling in QM . . . 90
4.4.1 Instantons . . . 90
4.4.2 Calculation of the determinant . . . 92
4.4.3 Double-well potential . . . 93
4.4.4 Periodic potential . . . 95
4.5 False vacuum decay and bounces . . . 96
4.6 The Yang–Mills instanton . . . 97
4.6.1 The instanton for $n - m = 1$ . . . 97
4.6.2 Evaluation of the fluctuation determinant K . . . 99
4.7 Instantons, fermions and chiral anomaly . . . 101
4.8 Baryon and lepton number violation in the SM . . . 106

5 Large orders and large N **109**
5.1 Asymptotic series and Borel summation . . . 110
5.2 Large order behaviour in quantum mechanics . . . 114
5.2.1 Anharmonic oscillator . . . 114
5.2.2 Degenerate ground state – the double well potential . . . 118
5.3 Extension to quantum field theory . . . 122
5.3.1 ϕ^4 theory . . . 122
5.3.2 Yang–Mills theory, QCD . . . 123
5.4 Large-N expansion on two-dimensional QFTs . . . 123
5.4.1 Gross–Neveu model . . . 123
5.4.2 $1/N$ expansion, non-perturbative . . . 126
5.4.3 Two-dimensional theories with scalar fields . . . 129

5.5 Large-N QCD 129
5.5.1 Four dimensions 130
5.5.2 Two dimensions 132
5.6 Operator product expansion and renormalons 136
5.6.1 Operator product expansion 136
5.6.2 Bubble graphs and renormalons 138
5.6.3 OPE as a transseries 139

Index **141**

Disclaimer: these notes are not official and were not endorsed by professor. The content is almost entirely based on the lectures, but occasionally the notes were modified or extended slightly to improve clarity or add extra information. In particular, the mistakes are almost surely mine. If you find any, please let me know at simon.bukovsek@tum.de or Simon.Bukovsek@campus.lmu.de.

Chapter 1

Supersymmetry

Motivation

1. It is a generalization of the concept of symmetry.
2. There might be phenomenological benefits;
   (a) absence of quadratic UV divergences, no hierarchy problem,
   (b) a possibility for unification SU(3) × SU(2) × U(1),
   (c) a natural candidate for dark matter (with assumption).
3. Local supersymmetry automatically leads to local version of Lorentz symmetry, which is in a sense equivalent to the general coordinate invariance (supergravity – a non-renormalizable quantum field theory for $E \ll M_{\mathrm{Pl}}$).

1.1 Algebraic basics and supersymmetry multiplets

Let G be a symmetry group of a theory. This means that the variation of the Lagrangian under the symmetry transformation is zero $\delta_g S = 0$ (or a total derivative). The symmetry group has a representation $g \to \mathcal{U}(g) = e^{i\varepsilon^A \theta^A}$ with real parameters ε^A and generators θ^A. There is an important statement, called the Coleman–Mandula theorem (1967).

Theorem 1.1 (*Coleman–Mandula*)**.** *If G is a symmetry group of a relativistic QFT, and it contains the Poincaré group as a subgroup $P \subset G$, then it can be written as a direct product of the Poincaré group and an internal symmetry group $G = P \times G_{internal}$.*

This theorem comes with many important consequences.

1. The fields will always have a Lorentz index α and an internal index separately i, $\phi_{\alpha,i}(x)$. There is no symmetry that would mix these two indices. The transformation $\mathcal{U}(g)\phi(c)\mathcal{U}^{-1}(g)$ changes α and i separately.
2. Particle states are classified by irreducible representations of the symmetry group $G \supset P$. This means that the Casimir operators of the Poincaré group P (*Pauli–Lubanski vector* and

mass) are also Casimir operators of G. The representations of G will be labelled by mass ($P^2 = P_\mu P^\mu$) and spin/helicity ($W^2 = W_\mu W^\mu$, $W_\mu = \varepsilon_{\mu\nu\rho\sigma} P^\nu J^{\rho\sigma}/2$). Particle multiplets will contain particles with the same mass and spin.

3. The infinitesimal transformation of the field is written as a commutator

$$\delta\phi = i[\varepsilon^A Q^A, \phi] = i\varepsilon^A T^A \phi$$

where Q^A are the generators of the symmetry group in the Hilbert space (operators) and T^A are the generators in the representation acting on the fields. The generators Q^A fulfil the Lie algebra commutation relations. Coleman–Mandula theorem implies

$$[P^\mu, B^A] = [J^{\mu\nu}, B^A] = 0$$

for any generators of the internal symmetries B^A. Consequently, B^A are scalars under Lorentz transformations and for any state $|\psi\rangle$, $B^A|\psi\rangle$ has the same mass, spin, four momenta, and spin orientation as $|\psi\rangle$.

The idea of *supersymmetry* came quite soon after the Coleman–Mandula theorem – between years 1972 and 1974. Supersymmetry relates particles with different spin (so it has to be a space-time symmetry). Let's see how we can generalize the concept of symmetry:

$$\mathcal{U}(g) = e^{i\varepsilon^A Q^A}$$

where ε^A are *Grassmann numbers* and hence Q^A must be a fermionic operator. This means that the total algebra becomes *graded*; it is made of bosonic and fermionic generators $L = L_\mathrm{B} \oplus L_\mathrm{F}$. For $B^A \in L_\mathrm{B}$ and $F^A \in L_\mathrm{F}$ we have

$$[B^A, B^B] \in L_\mathrm{B}, \quad \{F^A, F^B\} \in L_\mathrm{B}, \quad [B^A, F^B] \in L_\mathrm{F}.$$

The variation of the fermionic generator F^i with respect to the group element g is

$$\delta_g F^i = i[\varepsilon^A B^A, F^i] = i\varepsilon^A T^A_{ij} F^j.$$

It follows that F^i transforms in a representation of the internal symmetry group. One finds (by Coleman–Mandula theorem), F^i carries a certain spin, and we can write them as Q^i_α where now i refers to the internal symmetry and α is an index of Lorentz symmetry.

Claim 1.2. *The fermionic generators can only transform in the* $(0, 1/2)$ *or* $(1/2, 0)$ *representation of the Lorentz group. Thus,* $\alpha = 1, 2$ *is a Weyl spinor index.*

Proof. Suppose Q_α is the (A, B) representation of the Lorentz group. Then $\overline{Q} = \varepsilon Q^*$ is in (B, A) (now ε represents the antisymmetric tensor). The anti-commutator $\{Q, \overline{Q}\}$, when decomposed into a direct sum of irreducible representations, contains also $(A+B, A+B)$ which must be a linear combination of bosonic generators $P^\mu \in (1/2, 1/2)$ or $J^{\mu\nu} \in (1, 0) \oplus (0, 1)$ or indeed $B^A \in (0, 0)$. So $A + B \leq 1/2$ and the only two fermionic possibilities are $(0, 1/2)$ and $(1/2, 0)$. □

Spinor index notation

To better keep track of the handedness of the spinors, we employ the dotted spinor notation:

$$\psi_\mathrm{L} \leftrightarrow \psi_\alpha, \quad \psi_\mathrm{R} \leftrightarrow \overline{\psi}^{\dot{\alpha}}.$$

With this notation we also have

$$\varepsilon^{\alpha\beta} = \varepsilon^{\dot\alpha\dot\beta} = -i\sigma^2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \varepsilon_{\alpha\beta} = \varepsilon_{\dot\alpha\dot\beta} = i\sigma^2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

This antisymmetric tensor can be used to raise and lower spinor indices:

$$\psi^\alpha = \varepsilon^{\alpha\beta}\psi_\beta, \quad \overline{\psi}_{\dot\alpha} = \varepsilon_{\dot\alpha\dot\beta}\overline{\psi}^{\dot\beta}.$$

Moreover, $(\psi^\alpha)^* = \varepsilon^{\alpha\beta}\psi^*_\beta$ is a right-handed spinor if ψ_α is left-handed. A few more useful identities are

$$\overline{\psi}^{\dot\alpha} = (\psi^\alpha)^*, \quad \psi_\alpha = (\overline{\psi}_{\dot\alpha})^*,$$
$$\psi^\dagger_{\rm R} \leftrightarrow \psi^\alpha, \quad \psi^\dagger_{\rm L} \leftrightarrow \overline{\psi}_{\dot\alpha}.$$

We use the summation convention $\phi\,\psi = \phi^\alpha\psi_\alpha$ (up-down) and $\overline{\phi}\,\overline{\psi} = \overline{\phi}_{\dot\alpha}\overline{\psi}^{\dot\alpha}$ (down-up). Be careful with the Hermitian conjugation: $(\phi\psi)^\dagger = \overline{\psi}\overline{\phi}$. The order when using indices is very important, but without indices it doesn't matter:

$$\phi\psi = \phi^\alpha\psi_\alpha = \phi_\alpha\varepsilon_{\alpha\beta}\psi^\beta = -\psi^\alpha\varepsilon_{\alpha\beta}\phi^\beta = \psi_\alpha\varepsilon^{\beta\alpha}\phi_\beta = \psi^\beta\phi_\beta = \psi\phi$$

The two Lorentz scalars that we can form are $\phi\psi = \phi^\dagger_{\rm R}\psi_{\rm L}$ and $\overline{\phi\psi} = \phi^\dagger_{\rm L}\psi_{\rm R}$. The Lorentz vectors are

$$\phi^\dagger_{\rm L}\overline{\sigma}^\mu\psi_{\rm L} = \overline{\phi}\overline{\sigma}^\mu\psi, \quad \psi^\dagger_{\rm R}\sigma^\mu\psi_{\rm R} = \phi\sigma^\mu\overline{\psi}.$$

The two sorts of spinors are

$$\text{Dirac spinor: } \begin{pmatrix}\psi_{\rm L}\\ \psi_{\rm R}\end{pmatrix} = \begin{pmatrix}\psi_\alpha\\ \overline{\chi}^{\dot\alpha}\end{pmatrix}, \qquad \text{Majorana spinor: } \begin{pmatrix}\psi_\alpha\\ \overline{\chi}^{\dot\alpha}\end{pmatrix} = \begin{pmatrix}\psi_\alpha\\ \varepsilon^{\alpha\beta}\chi^*_\beta\end{pmatrix}.$$

A free Lagrangian for the Dirac spinor looks like

$$\mathcal{L} = \overline{\psi}i\overline{\sigma}^\mu\partial_\mu\psi + \chi i\sigma^\mu\partial_\mu\overline{\chi} - m(\chi\psi + \overline{\psi}\overline{\chi}).$$

1.1.1 Algebra of supersymmetry

There are bosonic generators we all know and love from previous courses $P_\mu, J_{\mu\nu}, B^A$, and now we add the fermionic generators $Q^i_\alpha, \overline{Q}^i_{\dot\alpha}$ for $i = 1, 2, \ldots, ,\mathcal{N}$. Let's introduce the *central charges* $Z_{ij} = a^A_{ij}B^A$ where a^A_{ij} are antisymmetric in i, j. The algebra can be split into four parts.

(I) Poincaré algebra for P_μ and $J_{\mu\nu}$ (the commutation relations between P^μ and $J^{\mu\nu}$ we used in QFT 1).

(II) The commutation of the bosonic internal symmetry generators: it forms a Lie algebra $[B^A, B^B] = if^{ABC}B^C$, and it commutes with Lorentz operators $[P_\mu, B^A] = [J_{\mu\nu}, B^A]$=0.

(III) The supersymmetric algebra tells us how the fermionic generators commute and anti-commute with other operators:

$$[Q^i_\alpha, P_\mu] = [\overline{Q}^i_\alpha, P_\mu] = 0,$$
$$[Q^i_\alpha, J_{\mu\nu}] = (\sigma_{\mu\nu})_\alpha{}^\beta Q^i_\beta, \quad \sigma^{\mu\nu} = \frac{1}{4}(\sigma^\mu\overline{\sigma}^\nu - \sigma^\nu\overline{\sigma}^\mu),$$
$$[Q^i_\alpha, B^A] = (b^A)^i{}_j Q^j_\alpha, \quad [\overline{Q}^{i\dot\alpha}, B^A] = -(b^{A\dagger})^i{}_j\overline{Q}^{j\dot\alpha},$$
$$\{Q^i_\alpha, \overline{Q}^j_{\dot\beta}\} = 2\delta^{ij}\sigma^\mu_{\alpha\dot\beta}P_\mu, \quad \{Q^i_\alpha, Q^j_\beta\} = \varepsilon_{\alpha\beta}Z^{ij}, \quad \{\overline{Q}^i_{\dot\alpha}, \overline{Q}^j_{\dot\beta}\} = -\varepsilon_{\dot\alpha\dot\beta}Z^{ij}.$$

(IV) Lastly, the central operators commute with everything: $[Z^{ij}, \text{any operator}] = 0$.

Showing the relations above is quite lengthy, so it is one of the exercises (see homework). One has to mostly use the Coleman–Mandula theorem and supersymmetric versions of Jacobi identities (super Jacobi identities). A *Lie superalgebra* is a $\mathbb{Z}/2\mathbb{Z}$-graded non-associative algebra $L = L_0 \oplus L_1$ with a bilinear operation $[\cdot,\cdot] : L \times L \to L$ such that for homogeneous elements $x, y, z \in L$ it holds

- grading: $[L_i, L_j] \subseteq L_{i+j}$,
- super skew-symmetry: $[x, y] = -(-1)^{|x||y|}[y, x]$,
- super Jacobi identity

$$(-1)^{|x||z|}[x,[y,z]] + (-1)^{|y||x|}[y,[z,x]] + (-1)^{|z||y|}[z,[x,y]] = 0.$$

The degree $|x|$ of an element x is 0 if $x \in L_0$ and 1 if $x \in L_1$. In more concrete terms, we use regular commutators for bosonic-bosonic and bosonic-fermionic generators, and anti-commutators for fermionic-fermionic generators. Bosonic generators are in L_0 (and have degree 0) and fermionic generators are in L_1 (and have degree 1).

Remark.

1. If $\mathcal{N} > 1$ we call it the *extended supersymmetry*. For $\mathcal{N} = 1$ there are no central charges.
2. Contracting the anti-commutator with another σ matrix gives

$$(\overline{\sigma}_\nu)^{\dot\beta\alpha}\{Q^i_\alpha, \overline{Q}^j_{\dot\beta}\} = 4\delta^{ij}P_\nu,$$

   from which we can conclude the positivity of the vacuum energy

$$\begin{aligned} E_{\text{vac}} = \langle\Omega|P_0|\Omega\rangle &= \frac{1}{4\mathcal{N}}\overline{\sigma}_0^{\dot\beta\alpha}\delta^{ij}\,\langle\Omega|\{Q^i_\alpha, \overline{Q}^j_{\dot\beta}\}|\Omega\rangle \\ &= \frac{1}{4\mathcal{N}}\,\langle\Omega|\{Q^i_\alpha, Q^{i*}_\alpha\}|\Omega\rangle \\ &= \frac{1}{4\mathcal{N}}\sum_{i,\alpha}\left(|Q^i_\alpha\,|\Omega\rangle\,|^2 + |Q^{i*}_\alpha\,|\Omega\rangle\,|^2\right)^2 \geq 0. \end{aligned}$$

   If $E_{\text{vac}} > 0$ it implies that SUSY is spontaneously broken, since, for at least one Q or Q^*, $Q^i_\alpha\,|\Omega\rangle \neq 0$. If $E_{\text{vac}} = 0$, then SUSY is unbroken.
3. Since $[Q, P] = [\overline{Q}, P] = 0$, the operator P^2 is still a Casimir, so all states in supersymmetric multiplet have equal mass. But since $[Q, J] \neq 0$, the Pauli–Lubanski operator is not Casimir, so particles within a multiplet can have different spins.
4. SUSY multiplet contains equal numbers of bosonic and fermionic states. This can be seen from the operator $(-1)^{N_\text{F}} = (-1)^{2s}$ where N_F is the fermion number operator. This has eigenvalue $+1$ for bosons and -1 for fermions. Using this operator, we can calculate the difference of bosons and fermions by summing over all particles in the multiplet with momentum p^μ

$$p^\mu(n_\text{B} - n_\text{F}) = p^\mu\,\text{tr}((-1)^{N_\text{F}})\sum_{|n\rangle}\langle n|(-1)^{N_\text{F}}P^\mu|n\rangle = \frac{1}{4\mathcal{N}}\overline{\sigma}^{\mu,\dot\beta\alpha}\sum_{|n\rangle}\langle n|(-1)^{N_\text{F}}\{Q^i_\alpha, \overline{Q}^i_{\dot\beta}\}|n\rangle = 0.$$

The last equality holds because we can cleverly rearrange the order of the operators inside the bra-ket $\langle n|(-1)^{N_F}\overline{Q}^i_{\dot\beta}Q^i_\alpha|n\rangle$ into $-\langle n|(-1)^{N_F}Q^i_\alpha\overline{Q}^i_{\dot\beta}|n\rangle$ as follows:

$$\sum_n \langle n|(-1)^{N_F}\overline{Q}^i_{\dot\beta}Q^i_\alpha|n\rangle = \sum_m \langle m|Q^i_\alpha(-1)^{N_F}\overline{Q}^i_{\dot\beta}|m\rangle = -\sum_n \langle n|(-1)^{N_F}Q^i_\alpha\overline{Q}^i_{\dot\beta}|n\rangle\,.$$

We inserted $\sum_m |m\rangle\langle m|$ between $\overline{Q}$ and Q in the first step, and then we used the anti-commutation property of the $(-1)^{N_F}$ operator to exchange the order with Q. In the same step, we also relabelled $m \to n$.

1.1.2 SUSY multiplets, one-particle states

Case 1: massless particles, $m = 0$.

In the case $\mathcal{N} = 1$ the story is simple. The states can be classified by helicity λ and momentum p. It holds

$$P^\mu|p,\lambda\rangle = p^\mu|p,\lambda\rangle\,, \quad W^\mu|p,\lambda\rangle = \lambda p^\mu|p,\lambda\rangle\,.$$

Without loss of generality we can assume $Q_\alpha|p,\lambda\rangle = 0$ since otherwise we could replace $|p,\lambda\rangle \to Q_1Q_2|p,\lambda\rangle = |p,\lambda\rangle'$. We also choose a reference momentum $p^\mu = E(1,0,0,1)$. If we now act with $\overline{Q}_{\dot 1}$ to the state $|p,\lambda\rangle$ we get

$$|\overline{Q}_{\dot 1}|p,\lambda\rangle|^2 = \langle p,\lambda|Q_1\overline{Q}_{\dot 1}|p,\lambda\rangle = \langle p,\lambda|\{Q_1,\overline{Q}_{\dot 1}\}|p,\lambda\rangle = 2(\sigma^\mu)_{1\dot 1}p_\mu = 2(\sigma^0-\sigma^3)_{1\dot 1}E = 0.$$

By acting with *Pauli–Lubanski vector* on the state $\overline{Q}_{\dot 2}|p,\lambda\rangle$ it can be shown, that it has helicity $\lambda + 1/2$ (we didn't show the proof). In particular,

$$|p,\lambda+1/2\rangle = \frac{1}{\sqrt{4E}}\overline{Q}_{\dot 2}|p,\lambda\rangle\,.$$

It can also be shown that acting with $Q_1, \overline{Q}_{\dot 1}$ and $\overline{Q}_{\dot 2}$ on $|p,\lambda+1/2\rangle$ gives zero and acting by Q_2 gives back $|p,\lambda\rangle$. This means that a supersymmetric multiplet contains particles with helicities λ and $\lambda + 1/2$ (a fermion and a boson) and CPT invariance requires also the negative helicities, $\pm\lambda$, $\pm(\lambda+1/2)$.

With this in mind, we can construct particle multiplets. The relevant ones are

- the *chiral multiplet* which is made of

  $$\text{complex scalar (twice } \lambda = 0) + \text{Weyl fermion } (\lambda = \pm 1/2),$$

- the *vector multiplet* (or gauge multiplet) which is made of

  $$\text{a gauge boson } (\lambda = \pm 1) + \text{Majorana-Weyl fermion } (\lambda = \pm 1/2),$$

- and the *graviton multiplet* which is made of

  $$\text{graviton } (\lambda = \pm 2) + \text{gravitino } (\lambda = \pm 3/2).$$

For $\mathcal{N} > 1$ we need to introduce N fermionic creation and annihilation operators, so there are $2^\mathcal{N}$ states and

$$\binom{\mathcal{N}}{n} \text{ states with helicity } \lambda + n/2 \text{ for } n = 0, 1, \ldots, \mathcal{N}.$$

If we want to have a renormalizable gauge theory we need the helicities to be $|\lambda| \leq 1$. This restricts us to $\mathcal{N} \leq 4$. If we only want to include gravity (which is not renormalizable) we can have $\mathcal{N} \leq 8$ ($|\lambda| \leq 2$). For any $\mathcal{N} > 8$ we cannot form a consistent local interacting QFT as it would require particles coupling with conserved currents with spin higher than 2 (which don't exist).

For $\mathcal{N} > 1$ fermionic states must either be neutral (Majorana) or come in pairs transforming in complex conjugate representations with respect to the internal symmetry ("vector like"). In the Standard Model, there are only chiral fermions, so only $\mathcal{N} = 1$ SUSY is possible.

Case 2: massive particles, $m > 0$.

Here, we will restrict ourselves to $\mathcal{N} = 1$. The multiplet contains $4 \cdot (2s + 1)$ states, where $s = 0, 1/2, 1, \ldots$ is the usual spin. In particular, there are two states with spin s and one state with spin $s \pm 1/2$. The exception is $s = 0$, where we only have two states with $s = 0$ and one state with $s = 1/2$ (Higgs/Higgsino multiplet).

1.2 Supersymmetric QFT ($\mathcal{N} = 1$)

The field lives on a Grassmannian manifold and its (unitary) transformation reads

$$\Psi'(x) = \mathcal{U}(\varepsilon)\Psi(x)\mathcal{U}^{-1}(\varepsilon), \quad \mathcal{U}(\varepsilon) = e^{i(\varepsilon Q + \overline{\varepsilon}\overline{Q})}.$$

The infinitesimal transformation is

$$\delta_\varepsilon \Psi(x) = i[\varepsilon Q + \overline{\varepsilon}\overline{Q}, \Psi(x)].$$

Consequently,

$$[\varepsilon Q, \overline{\eta}\overline{Q}] = 2\varepsilon\sigma^\mu\overline{\eta}P_\mu.$$

On the fields, the commutator of the two transformations acts as a translation

$$(\delta_\eta\delta_\varepsilon - \delta_\varepsilon\delta_\eta)\Psi(x) = 2(\varepsilon\sigma^\mu\overline{\eta} - \eta\sigma^\mu\overline{\varepsilon}) \underbrace{[P_\mu, \Psi(x)]}_{-i\partial_\mu\Psi(x)}.$$

This follows from the Jacobi identity. The action of the field theory has to be invariant under SUSY transformations $\delta_\varepsilon S = 0$.

1.2.1 Wess–Zumino model

The simplest SUSY model is the *Wess–Zumino model* which is a field theory for chiral (or $s = 0$, $m^2 \neq 0$) multiplet: 1 complex scalar ϕ and 1 Weyl fermion ψ. The free theory Lagrangian is

$$\mathcal{L} = \partial^\mu\phi^*\partial_\mu\phi + \overline{\psi}i\overline{\sigma}^\mu\partial_\mu\psi.$$

We need to find $\delta_\varepsilon\phi$ and $\delta_\varepsilon\psi$ so that $\delta_\varepsilon\mathcal{L}$ is a total derivative. This can only work "on-shell", i.e. when the equations of motion are satisfied ($\overline{\sigma}^\mu\partial_\mu\psi = 0$). Let's first count the degrees of freedom on- and off-shell.

Field	DoF on-shell	DoF off-shell
ϕ	2	2
ψ	2	4
F	0	2

From this observation we can conclude that we need to introduce 1 complex auxiliary field F (no kinetic term, no dynamics) so that the new Lagrangian is

$$\mathcal{L} = \partial^\mu \phi^* \partial_\mu \phi + \overline{\psi} i \overline{\sigma}^\mu \partial_\mu \psi + F^* F.$$

We want to find the most general form for δ_ε on ϕ, ψ, F. First, we will start with a general ansatz, which we will constrain later:

$$\begin{aligned}
\delta_\varepsilon \phi &= a\varepsilon\psi + b\overline{\varepsilon}\overline{\psi}, \\
\delta_\varepsilon \psi_\alpha &= c\sigma^\mu_{\alpha\dot{\beta}} \overline{\varepsilon}^{\dot{\beta}} \partial_\mu \phi + d\sigma^\mu_{\alpha\dot{\beta}} \overline{\varepsilon}^{\dot{\beta}} \partial_\mu \phi^* + f\varepsilon_\alpha F + g\varepsilon_\alpha F^*, \\
\delta_\varepsilon F &= -ik\overline{\varepsilon}\overline{\sigma}^\mu \partial_\mu \psi - il\partial_\mu \overline{\psi}\overline{\sigma}^\mu \varepsilon.
\end{aligned}$$

From the dimensional analysis we have $[\varepsilon] = -1/2$, $[Q] = 1/2$ and $[F] = 2$. We used these facts to write down the ansatz above. When we determine the coefficients we require two things:

- the invariance of the action $\delta_\varepsilon S = 0$ and
- the derivations have to close the algebra, i.e. $\delta_\varepsilon \delta_\eta - \delta_\eta \delta_\varepsilon \propto \partial_\mu$.

The solution is

$$\begin{aligned}
\delta_\varepsilon \phi &= \sqrt{2}\varepsilon\psi \\
\delta_\varepsilon \psi &= -i\sqrt{2}\sigma^\mu \overline{\varepsilon} \partial_\mu \phi + \sqrt{2}\varepsilon F \\
\delta_\varepsilon F &= -i\sqrt{2}\overline{\varepsilon}\overline{\sigma}^\mu \partial_\mu \psi.
\end{aligned}$$

The choice of coefficients is unique up to rescaling $a \to \lambda a$, $c \to \lambda^{-1} c$. For the supersymmetric interactions we generalize the Lagrangian to n chiral multiplets (ψ_i, ϕ_i, F_i) $\mathcal{L} = \mathcal{L}_{\text{free}} + \mathcal{L}_{\text{int}}$. The most general renormalizable interaction is

$$\mathcal{L}_{\text{int}} = -\frac{1}{2} A_{ij}(\phi_k)\psi_i\psi_j + B_i(\phi_k)F_i + C(\phi_k) + \text{h.c.}$$

The notation implies summation over spinor indices $\psi_i\psi_j = \psi_i^\alpha \psi_{j,\alpha}$. The functions A_{ij}, B_i, C are analytic in ϕ_i. Up to mass dimension 4, we have the following options:

$$A\psi\psi \supset \phi\psi\psi, \psi\psi, \phi^*\psi\psi,$$

$$B_i \supset \phi_i, \phi_i^*, \ldots$$

$$C \supset \text{polynomials in } \phi, \phi^* \text{ up to dimension 4.}$$

In the following, we will scale $\sqrt{2}\varepsilon \to \varepsilon$. The variation of the function C is

$$\delta_\varepsilon C = \frac{\partial C}{\partial \phi_k}\varepsilon\psi_k + \frac{\partial C}{\partial \phi_k^*}\overline{\varepsilon}\overline{\psi}_k.$$

This cannot be cancelled by any other variation of the other terms, so $\delta_\varepsilon C = 0$. Hence, C cannot contain ϕ^* or ϕ, and we can forget about any constants. Hence, without loss of generality we can write $C = 0$. The other variations are a bit more interesting:

$$\delta_\varepsilon \left(-\frac{1}{2} A_{ij}(\phi_k)\psi_i\psi_j \right) = -\frac{1}{2}\frac{\partial A_{ij}}{\partial \phi_k}(\varepsilon\psi_k)(\psi_i\psi_j) - \frac{1}{2}\frac{\partial A_{ij}}{\partial \phi_k^*}(\overline{\varepsilon}\overline{\psi}_k)(\psi_i\psi_j) - \frac{1}{2}A_{ij}(\delta_\varepsilon\psi_i\,\psi_j + \psi_i\delta_\varepsilon\psi_j).$$

Using the *Fierz identity*

$$(\varepsilon\psi_k)(\psi_i\psi_j) + (\varepsilon\psi_i)(\psi_j\psi_k) + (\varepsilon\psi_j)(\psi_i\psi_k) = 0$$

we can see that $\partial A_{ij}/\partial\phi_k$ is totally symmetric in i, j, k. As there is no Fierz identity for the second term, we must have

$$\frac{\partial A_{ij}}{\partial\phi_k^*} = 0.$$

Consequently, A_{ij} is independent of ϕ^*, i.e., it is holomorphic function of ϕ_k. We introduce the *superpotential* $W(\phi_k)$ with

$$A_{ij} = \frac{\partial W}{\partial\phi_i\partial\phi_j}.$$

Now we are only left with the last term in the variation of A_{ij}, and we hope it can cancel with the variation of $B_i(\phi_k)F_i$. This will hold if

$$B_i(\phi_k) = \frac{\partial W(\phi_k)}{\partial\phi_i}$$

and then $\delta_\varepsilon\mathcal{L}$ = total derivative. The most general interacting Lagrangian given all the constraints is

$$\mathcal{L}_{\text{int}} = -\frac{1}{2}\frac{\partial W}{\partial\phi_i\partial\phi_j}\psi_i\psi_j + \frac{\partial W}{\partial\phi_i}F_i + \text{h.c.}$$

Next, we eliminate the auxiliary fields F_i using their equations of motion

$$\frac{\partial\mathcal{L}}{\partial F_i^*} = F_i + \frac{\partial W^*}{\partial\phi_i^*} = 0.$$

We get

$$\mathcal{L}_{\text{WZ}} = \partial^\mu\phi_i^*\partial_\mu\phi_i + \overline{\psi}_i i\overline{\sigma}^\mu\partial_\mu\psi_i - \frac{1}{2}(W_{ij}\psi_i\psi_j + W_{ij}^*\overline{\psi}_i\overline{\psi}_j) - W_iW_i^*,$$

where

$$W_{ij} = \frac{\partial W}{\partial\phi_i\partial\phi_j}, \quad W_i = \frac{\partial W}{\partial\phi_i}.$$

The last term is a scalar potential $V(\phi_i, \phi_i^*) = W_iW_i^*$. The most general renormalizable superpotential is

$$W = \frac{1}{2}M_{ij}\phi_i\phi_j + \frac{1}{6}y_{ijk}\phi_i\phi_j\phi_k$$

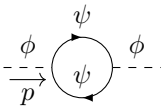
where both M_{ij} and y_{ijk} are totally symmetric. We can see that the mass dimension of W is 3. It is straightforward to take the derivatives of the potential. This will give us all possible interactions and mass terms:

$$-\frac{1}{2}W_{ij}\psi_i\psi_j + \text{h.c.} = -\underbrace{\frac{1}{2}M_{ij}\psi_i\psi_j}_{\text{Fermion mass matrix}} - \underbrace{\frac{1}{2}y_{ijk}\phi_k\psi_i\psi_j}_{\text{Yukawa interaction}} + \text{h.c.}$$

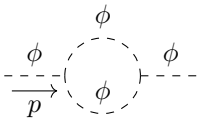
$$W_iW_i^* = M_{ij}M_{ik}^*\phi_i\phi_j + \left(\frac{1}{2}M_{ij}y_{ikl}^*\phi_j\phi_k^*\phi_l^* + \text{h.c.}\right) + \frac{1}{4}y_{ijk}y_{imn}^*\phi_j\phi_k\phi_m^*\phi_n^*.$$

From these expressions we can see that the mass of the fermion and the scalar are *the same* (all supersymmetric partners have the same mass). Moreover, from the last expression we can see

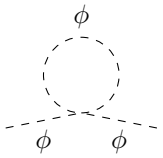
that the Yukawa interaction is related by the scalar self-interaction! To see the last statement more clearly, let's calculate the cancellation of the quadratic divergences in the scalar propagator.



$$= -\frac{iy}{2}\frac{iy^*}{2}\int\frac{\mathrm{d}^4k}{(2\pi)^4}\frac{\mathrm{tr}(i(\not{k}+m)(1+\gamma_5)i(\not{p}+\not{k}+m)(1-\gamma_5))}{4(k^2-m^2)((p+k)^2-m^2)} \sim -\frac{yy^*}{4}\int\frac{\mathrm{d}^4k}{(2\pi)^4}\frac{1}{k^2},$$



$$= \text{(no quadratic divergence)},$$



$$= (-i)\frac{yy^*}{4}\int\frac{\mathrm{d}^4k}{(2\pi)^4}\frac{i}{k^2-m^2} \sim \frac{yy^*}{4}\int\frac{\mathrm{d}^4k}{(2\pi)^4}\frac{1}{k^2}.$$

Evidently, the quadratic divergences cancel out exactly at one loop, and in fact to all orders of perturbation theory.

1.3 Supersymmetric gauge theory

We will consider the QFT for the vector multiplet (vector + Majorana–Weyl fermion). The gauge field is A_μ^A and the Weyl spinor (the *gaugino*) is λ_α^A. There is also a real auxiliary field D^A.

Field	DoF on-shell	DoF off-shell
A	2	3
λ	2	4
D	0	1

Gauge transformations of fields are

$$\begin{aligned}\delta A_\mu^A &= f^{ABC}\Lambda^C A_\mu^B + \frac{1}{g}\partial_\mu\Lambda^A,\\ \delta\lambda_\alpha^A &= f^{ABC}\Lambda^C\lambda_\alpha^B,\\ \delta D^A &= f^{ABC}\Lambda^C D^B.\end{aligned}$$

The Lagrangian reads

$$\mathcal{L} = -\frac{1}{4}G_{\mu\nu}^A G^{A,\mu\nu} + i\overline{\lambda}^A\overline{\sigma}^\mu(\mathrm{D}_\mu\lambda)^A + \frac{1}{2}D^A D^A[+\xi D].$$

The last term is only possible for U(1) gauge group and is called the *Fayet–Iliopoulos term*. The covariant derivative is defined as usual

$$(\mathrm{D}_\mu\lambda)^A = \partial_\mu\lambda^A + gf^{ABC}A_\mu^B\lambda^C$$

since the gaugino transforms in the adjoint representation. For non-Abelian theories there can also be θ-term, which only contributes non-perturbatively and is thoroughly discussed in Chapter 4. The SUSY transformations are

$$\begin{aligned}\delta_\varepsilon A_\mu^A &= i(\varepsilon\sigma_\mu\overline{\lambda}^A + \lambda^A\sigma_\mu\overline{\varepsilon})\\ \delta_\varepsilon\lambda_\alpha^A &= \frac{1}{2}(\sigma^\mu\overline{\sigma}^\nu\varepsilon)_\alpha G_{\mu\nu}^A + i\varepsilon_\alpha D^A\\ \delta_\varepsilon D^A &= \varepsilon\sigma^\mu(\mathrm{D}_\mu\overline{\lambda})^A - (\mathrm{D}_\mu\lambda)^A\sigma^\mu\overline{\varepsilon}.\end{aligned}$$

As we can see, the transformations are linear in ε. For non-Abelian gauge groups, this is a non-linear transformation in the fields.

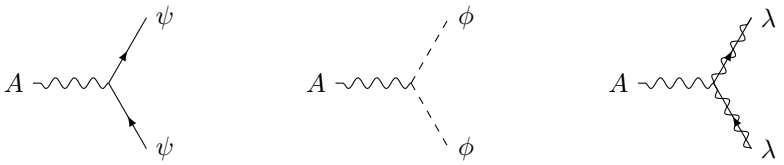
When we couple the gauge field to matter $X = \{\phi, \psi\}$, it transforms in some representation of the gauge group:

$$\delta X_\alpha = i\Lambda^A T_{\alpha\beta}^A X_\beta.$$

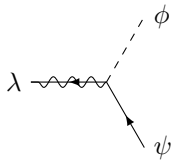
Combining gauge Lagrangian with Wess–Zumino Lagrangian yields a gauge invariant theory, but not supersymmetric. Supersymmetric and gauge invariant Lagrangian can be derived using the *superfield formalism* (see exercises), the result reads

$$\mathcal{L} = \mathcal{L}_{\mathrm{WZ}}|_{\partial_\mu\to\mathrm{D}_\mu} + \mathcal{L}_{\mathrm{gauge}} + ig\sqrt{2}(\phi^\dagger T^A\psi^\alpha\lambda_\alpha^A - \overline{\psi}_{\dot\alpha}T^A\phi\lambda^{A,\dot\alpha}) + g\phi^\dagger T^A\phi D^A.$$

It is the most general supersymmetric renormalizable gauge-invariant Lagrangian. Aside from the regular gauge interactions



we also get the extra term $\phi\psi \to \lambda$.



In order to get the SUSY invariance, the supersymmetric variation of F, $\delta_\varepsilon F$, has to be modified (maybe replacing regular derivative by covariant derivative).

Remark.

1. If the gauge group is bigger $G = G_1 \otimes \cdots \otimes G_n$, the covariant derivative becomes

$$\mathrm{D}^\mu\phi_{a_1,\ldots,a_n} = \partial^\mu\phi_{a_1,\ldots,a_n} - ig_1 T_{a_1b_1}^{A_1}A^{\mu,A_1}\phi_{b_1,a_2,\ldots,a_n} + \ldots$$

   The Lagrangian contains the terms

$$\mathcal{L} \supset \sum_{i=1}^n ig_i\sqrt{2}(\phi^\dagger_{a_1,\ldots,a_n}T^{A_i}_{a_ib_i}\psi^\alpha_{a_1,\ldots,b_i,\ldots,a_n}\lambda^{A_i}_{i\alpha} + \mathrm{h.c.}) + g_i\phi^\dagger_{a_1,\ldots,a_n}T^{A_i}_{a_ib_i}\phi_{a_1,\ldots,b_i,\ldots,a_n}D_i^{A_i}.$$

2. The auxiliary fields can be derived from the equations of motion:

$$F_i = -\left(\frac{\partial W}{\partial \phi_i}\right)^* = -W_i^*, \quad D^A = -\xi - g\phi^\dagger T^A \phi.$$

The term with ξ can only exist for U(1) symmetry, and thus it doesn't have any indices. The scalar potential is

$$V = -F_i^* F_i - (W_i F_i + \text{h.c.}) - \frac{1}{2} D^A D^A - g\phi^\dagger D^A T^A \phi - \xi D = |W_i|^2 + \frac{1}{2}|\xi + g\phi^\dagger T^A \phi|^2.$$

The first term is scalar self-interaction related to Yukawa couplings, and the second term is the scalar self-interaction related to the gauge coupling.

3. Spontaneous symmetry breaking of gauge symmetry: $G \to H$. The number of massive gauge bosons is $\dim G - \dim H$. When SUSY is unbroken (and only gauge symmetry is broken), the states organize into multiplets. For every broken dimension, we get one massive multiplet, which contains

   - 1 massive gauge boson,
   - 2 massive Weyl fermions (gaugino and Higgsino),
   - 1 real scalar field (regular Higgs field)

   After the breaking, we are still left with some residual gauge symmetry H:

   - $\dim H$ massless gauge bosons and gauginos,
   - $N - (\dim G - \dim H)$ complex scalars and Weyl fermions (N is the number of chiral multiplets, if ϕ is initially in adjoint representation of SU(n), this would mean $N = n^2 - 1$).

1.4 Supersymmetry breaking

SUSY breaking (spontaneous or explicit) can also occur (*superhiggs mechanism*). Indeed, it has to occur as degenerate fermion/boson pairs are not observed in nature. In the spontaneous case, we have some operator $\mathcal{O}$, such that

$$\langle \Omega | \delta_\varepsilon \mathcal{O} | \Omega \rangle \neq 0.$$

Poincaré symmetry forbids non-scalar fields to have non-zero vacuum expectation values which already eliminates a lot of options. We have two options:

1. F-term breaking: $\langle \Omega | \delta_\varepsilon \Psi_i | \Omega \rangle \propto \langle \Omega | F_i | \Omega \rangle \neq 0$,
2. D-term breaking: $\langle \Omega | \delta_\varepsilon \lambda^A | \Omega \rangle \propto \langle \Omega | D^A | \Omega \rangle \neq 0$.

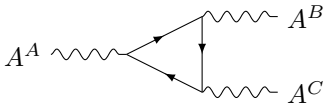
The SUSY breaking lifts mass degeneracy between superpartners, but there is still some relation between their masses (1979)

$$\sum_J (-1)^{2J} (2J+1) m_J^2 = 0$$

where J is the spin and (see Weinberg QFT III, 27.5.10):

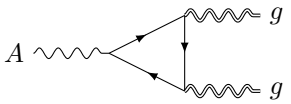
$$\sum_{\substack{\text{real} \\ \text{scalars}}} m_i^2 - 2 \sum_{\substack{\text{chiral} \\ \text{fermions}}} m_i^2 + 3 \sum_{\text{vectors}} m_i^2 = -2 \langle \Omega | D^A | \Omega \rangle \operatorname{tr} T^A + \mathcal{O}(g^2).$$

The relation holds at tree level and comes from the scalar self-interaction related to the gauge coupling. The trace is zero for all non-Abelian gauge groups, because non-Abelian generators are traceless.



$\sim \mathrm{tr}(T^A\{T^B, T^C\})$

If the theory is anomaly free, then the trace vanishes also for the U(1) symmetry, and the relation above simplifies to give zero on the right-hand side. For example, the traces of the U(1) have to be zero in order to cancel the gravitational contributions to anomaly via U(1)-graviton-graviton anomaly



$\sim \mathrm{tr}(T^A)\varepsilon^{\mu\nu\rho\sigma}R_{\mu\nu\kappa\lambda}R_{\rho\sigma}{}^{\kappa\lambda}.$

Of course, the unbroken conservation of charge, colour, and baryon and lepton numbers prevents the mass matrices from having elements linking particles with different values of these quantum numbers, so all these results hold separately for each set of conserved quantum numbers. For the lepton flavour number, we have

$$2m_{\nu_e}^2 + 4m_e^2 = \text{sum of } m_i^2 \text{ for particles with lepton number.}$$

Of course, this relation is not satisfied in nature (otherwise we would detect these other particles), so SUSY breaking cannot occur via F or D term if either belonged to the Standard Model. A possible solution is to have a hidden sector where SUSY is broken and then communicate this breaking to the visible sector (the Standard Model) via the electroweak force. For energies $E \ll M_s \approx \sqrt{\langle\Omega|F|\Omega\rangle}$ the SUSY is broken explicitly:

1. scalar mass terms: $-m_{ij}^2\phi_i^*\phi_j$, the scalar masses would not coincide with the fermion masses,
2. gaugino mass terms: $M_A\lambda^A\lambda^A$,
3. non-holomorphic trilinear scalar interactions: $A_{ijk}\phi_i\phi_j\phi_k$.

Terms like $\psi\overline{\lambda}$ or similar are forbidden by gauge invariance (which we want to keep unbroken in this case).

If SUSY is broken as fermionic symmetry, the new particles that appear are massless fermions called *goldstinos* (analogy to Goldstone bosons). A supersymmetric theory of gravity can be produced by considering SUSY as a local symmetry (which brings gauge fermions and gauge bosons). The spontaneous breaking of local SUSY gives mass to the spin-3/2 *gravitino* (2 DoF from massless case + 2 DoF from goldstino = 4 DoF of massive gravitino). The estimate for the gravitino mass is

$$m_{3/2} = \frac{\langle\Omega|F|\Omega\rangle}{\sqrt{3}M_{\mathrm{Pl}}}.$$

The gauge bosons (gravitons) remain massless and can mediate long-range gravitational interactions.

1.5 Minimally supersymmetric Standard model

Minimally supersymmetric Standard Model (MSSM) is the SUSY extension of the Standard Model. The field content extends each SM particle into a super-multiplet. Firstly, every gauge boson gets a gaugino superpartner:

- SU(3): g (gluon), $\tilde{g}$ (*gluino*),
- SU(2)$_L$: W (W boson), $\tilde{W}$ (*wino*),
- U(1)$_Y$: B (B boson), $\tilde{B}$ (*bino*).

The fermions can be represented by left-handed fermion multiplets and the complex conjugates of right-handed singlets, which also transform as left-handed ($u^c_{\mathrm{R}} = \varepsilon u^*_{\mathrm{R}}$).

	spin-1/2	spin-0	SU(3)	SU(2)	U(1)$_Y$
quarks, *squarks*	$Q = \begin{pmatrix} u_{\mathrm{L}} \\ d_{\mathrm{L}} \end{pmatrix}$	$\tilde{Q} = \begin{pmatrix} \tilde{u}_{\mathrm{L}} \\ \tilde{d}_{\mathrm{L}} \end{pmatrix}$	3	2	+1/6
	$U = u^c_{\mathrm{R}}$	$\tilde{U} = \tilde{u}^*_{\mathrm{R}}$	$\overline{3}$	1	−2/3
	$D = d^c_{\mathrm{R}}$	$\tilde{D} = \tilde{d}^*_{\mathrm{R}}$	$\overline{3}$	1	+1/3
leptons, *sleptons*	$L = \begin{pmatrix} \nu_{\mathrm{L}} \\ e_{\mathrm{L}} \end{pmatrix}$	$\tilde{L} = \begin{pmatrix} \tilde{\nu}_{\mathrm{L}} \\ \tilde{e}_{\mathrm{L}} \end{pmatrix}$	1	2	−1/2
	$E = e^c_{\mathrm{R}}$	$\tilde{E} = \tilde{e}^*_{\mathrm{R}}$	1	1	+1
Higgs, *Higgsinos*	$\tilde{H}_u = \begin{pmatrix} \tilde{H}^+_u \\ \tilde{H}^0_u \end{pmatrix}$	$H_u = \begin{pmatrix} H^+_u \\ H^0_u \end{pmatrix}$	1	2	+1/2
	$\tilde{H}_d = \begin{pmatrix} \tilde{H}^0_d \\ \tilde{H}^-_d \end{pmatrix}$	$H_d = \begin{pmatrix} H^0_d \\ H^-_d \end{pmatrix}$	1	2	−1/2

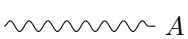
We can see that the content of the MSSM is more than doubled. Why do we have two Higgs doublets? A term in the SM is

$$\mathcal{L}_{\mathrm{SM}} \supset -\lambda^D \overline{Q} \phi D_{\mathrm{R}} - \lambda^U \overline{Q} \tilde{\phi} U_{\mathrm{R}} + \mathrm{h.c.}$$

It contains $\tilde{\phi} = \varepsilon \phi^*$. However, this cannot be extended to a SUSY invariant term since the superpotential has to be holomorphic. In order to write down the proper MSSM Higgs-fermion interactions, we use the following notation:

$$H^i_d \varepsilon_{ij} Q^j D = H_d Q D, \quad H^i_u \varepsilon_{ij} Q^j U = H_u Q U, \quad H_u H_d = \varepsilon_{ij} H^i_u H^j_d, \quad H^\dagger_u H_u = H^{\dagger i}_u H^i_u.$$

The Feynman propagators are the following:

- Gauge bosons (regardless whether Abelian or not):  A
- Gauginos: λ
- Fermions: ψ
- Sfermions: ϕ
- Higgs bosons: H
- Higgsinos: $\tilde{H}$

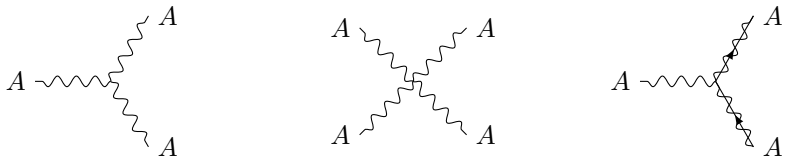
The Lagrangian in the MSSM is

$$\mathcal{L}_{\text{MSSM}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{chiral, gauge}} + \mathcal{L}_{\text{superpotential}} + \mathcal{L}_{\text{soft breaking}}.$$

- The gauge Lagrangian:

$$\mathcal{L}_{\text{gauge}} = \sum_{\text{U(1),SU(2),SU(3)}} \left(-\frac{1}{4} G^A_{\mu\nu} G^{\mu\nu,A} + i\overline{\lambda}^A \overline{\sigma}^\mu \mathrm{D}_\mu \lambda^A + \frac{1}{2} D^A D^A \right).$$

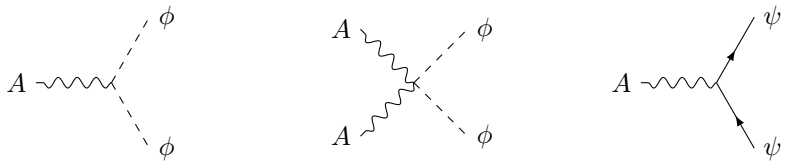
  It includes the gauge boson interaction and gaugino interactions.



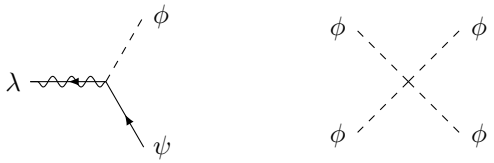
- The chiral Lagrangian

$$\mathcal{L}_{\text{chiral, gauge}} = \sum_{\text{scalars}} (\mathrm{D}_\mu \phi)^\dagger (\mathrm{D}^\mu \phi) + \sum_{\text{fermions}} \overline{\psi} i \overline{\sigma}^\mu \mathrm{D}_\mu \psi + \sum_I F_I^\dagger F_I + \\ + \sum_{\text{U(1),SU(2),SU(3)}} \sum_{\substack{\text{chiral} \\ \text{multiplets}}} \left[i g_i \sqrt{2} (\phi^\dagger T_i^A \psi \lambda^A - \overline{\psi} T_i^A \phi \overline{\lambda}^A) + g_i \phi^\dagger T_i^A \phi D^A \right],$$

  where $I = \{Q, U, D, L, E, H_u, H_d\}$. This includes the sfermion to gauge boson interactions and fermion to gauge boson interactions



  as well as the gaugino–fermion–sfermion interactions and scalar 4-point interactions.



  All interactions contain only gauge couplings g_i.

- The superpotential Lagrangian

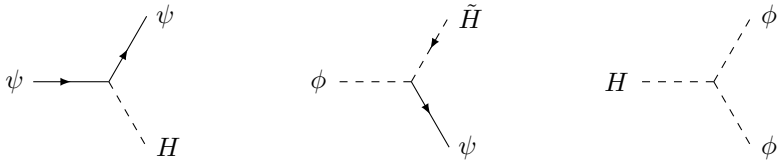
$$\mathcal{L}_{\text{superpotential}} = -\frac{1}{2} \frac{\partial^2 W}{\partial \phi_i \partial \phi_j} \psi_i \psi_j + \frac{\partial W}{\partial \phi_i} F_i + \text{h.c.} = \mathcal{L}_R + \mathcal{L}_{\not{R}}.$$

  We split the superpotential into R-parity conserving and R-parity violating parts $W = W_R + W_{\not{R}}$:

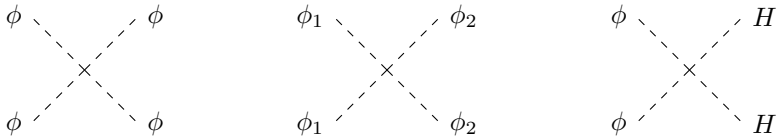
$$W_R = -y^u_{IJ} H_u \tilde{Q}_I \tilde{Q}_J + y^d_{IJ} H_d \tilde{Q}_I \tilde{Q}_J + y^E_{IJ} H_d \tilde{L}_I \tilde{E}_J + \mu H_u H_d,$$

$$W_{\not R} = \lambda'^{E}_{IJK}\tilde{L}_I\tilde{L}_J\tilde{E}_K + \lambda'^{D}_{IJK}\tilde{L}_I\tilde{Q}_J\tilde{D}_K + \lambda_{IJK}\tilde{U}_I\tilde{U}_J\tilde{U}_K + \mu'_I\tilde{L}_I H_u.$$

The R-parity is a discrete symmetry, that acts trivially on the SM particles ($\phi \to \phi$) and non-trivially on their superpartners ($\tilde{\phi} \to -\tilde{\phi}$). If the R-parity holds, the Lagrangian can only contain terms with even number of superpartners. The superpotential includes Yukawa type Higgs coupling to quarks and leptons, a mass term for Higgs and Higgsino, sfermion–fermion–Higgsino interactions and Higgs–sfermion–sfermion interactions

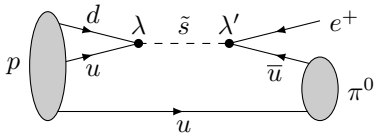


as well as scalar 4-point interactions (4 squarks, 4 sleptons, 2 squarks with 2 sleptons, 2 squarks with 2 Higgs fields and 2 sleptons with 2 Higgs fields).



Importantly, we don't get Higgs 4-point interaction from the superpotential, only from the gauge interactions.

The R-parity violating part includes lepton and baryon number violating interactions, which leads to proton decay.



The proton decay time can be estimated as

$$\tau_{\text{proton}} \approx 6 \cdot 10^{-13}\,\text{s}\left(\frac{m_Q}{1\,\text{TeV}}\right)^4 \frac{1}{(\lambda\lambda')^2}.$$

The experimental bound is $\tau_{\text{proton}} \gtrsim 1 \cdot 10^{41}\,\text{s}$, which gives $\lambda\lambda' < 10^{-28}$. If the R-parity is conserved ($W_{\not R} = 0$) only the interactions with even number of superpartners can exist, and the lightest supersymmetric particle (LSP) is stable. This particle should be neutral and colourless and is hence a good dark matter candidate. There are two options: sneutrino $\tilde{\nu}$ and neutralino (a combination of Higgsino, Wino and Bino).

- The soft SUSY breaking Lagrangian

$$\mathcal{L}_{\text{breaking}} = -\frac{1}{2}\sum_{\substack{\text{U(1),SU(2),}\\ \text{SU(3)}}} m_i\lambda_i^A\lambda_i^A + \sum_{\Phi} m_{\Phi,IJ}\Phi_I^\dagger\Phi_J - m_u^2 H_u^\dagger H_u - m_d^2 H_d^\dagger H_d + B\mu H_u H_d +$$
$$- (A^u y^u)_{IJ} H_u\tilde{Q}_I\tilde{Q}_J + (A^d y^d)_{IJ} H_d\tilde{Q}_I\tilde{Q}_J + (A^E y^E)_{IJ} H_d\tilde{L}_I\tilde{E}_J + \text{h.c.}$$

where $\Phi \in \{\tilde{Q}, \tilde{L}, \tilde{U}, \tilde{D}, \tilde{E}\}$, B is coupling with mass dimension 1 and I, J are generation indices.

We have introduced a lot of new parameters in the MSSM. The Standard Model has 17 couplings and 1 CP-violating phase (18 in total). The MSSM has 79 couplings and 44 CP-violating phases (123 in total). It is thus rich in phenomenology.

Let's take a look at one of the processes: the electroweak symmetry breaking. In the SM, it happens via

$$\mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_Y \to \mathrm{U}(1)_\mathrm{em}.$$

To avoid $\mathrm{SU}(3)_\mathrm{c}$ breaking, squarks cannot get vacuum expectation values. It turns out that neither can sleptons get vacuum expectation values, because otherwise the $\mathrm{SU}(2) \times \mathrm{U}(1)$ would break differently. Hence, we demand only electroweak symmetry breaking, which restricts parameters of the Lagrangian, but it begs a question how does this happen in a "natural" way. We won't answer this question here, but just outline the procedure. The Higgs potential is split into three parts:

$$V = V_D + V_F + V_{\text{SUSY breaking}}.$$

- Gauge D-terms:

$$V_D = \frac{1}{2}\left(\underbrace{D^A D^A}_{\mathrm{SU}(2)} + \underbrace{D^2}_{\mathrm{U}(1)}\right) = \frac{g^2}{2}\left(H_u^\dagger T^A H_u + H_d^\dagger T^A H_d\right)^2 + \frac{g'^2}{2}\left(\frac{1}{2}H_u^\dagger H_u + \frac{1}{2}H_d^\dagger H_d\right)$$
$$= \frac{g^2}{2}|H_u^\dagger H_d|^2 + \frac{g^2+g'^2}{8}(H_u^\dagger H_u - H_d^\dagger H_d)^2.$$

- F-terms:

$$V_F = \sum_i \left|\frac{\partial W}{\partial H_{u,i}}\right|^2 + \left|\frac{\partial W}{\partial H_{d,i}}\right|^2 = |\mu|^2(H_u^\dagger H_u + H_d^\dagger H_d).$$

- SUSY breaking terms:

$$V_{\text{SUSY breaking}} = m_u^2 H_u^\dagger H_u + m_d^2 H_d^\dagger H_d + (B\mu H_u H_d + \text{h.c.}).$$

Without loss of generality, we can choose real and positive vacuum expectation values for the neutral components of the Higgs doublets:

$$H_u = \begin{pmatrix} \hat{H}_u^+ \\ v_u + \hat{H}_u^0 \end{pmatrix}, \quad H_d = \begin{pmatrix} v_d + \hat{H}_d^0 \\ \hat{H}_d^- \end{pmatrix}.$$

We define the parameters $v^2 = v_u^2 + v_d^2$ and $\tan\beta = v_u/v_d$. The W and Z masses are the same as in the SM

$$m_W = \frac{g}{\sqrt{2}}v, \quad m_Z = \frac{\sqrt{g^2+g'^2}}{\sqrt{2}}v.$$

When we insert the expansion into the potential and expand to the second order in the fields,

we get

$$
\begin{aligned}
V = \text{const.} &+ \begin{pmatrix} \text{Im}\,\hat{H}_u^0 & \text{Im}\,\hat{H}_d^0 \end{pmatrix} \mathsf{M}_1 \begin{pmatrix} \text{Im}\,\hat{H}_u^0 \\ \text{Im}\,\hat{H}_d^0 \end{pmatrix} + \\
&+ \begin{pmatrix} \text{Re}\,\hat{H}_u^0 & \text{Re}\,\hat{H}_d^0 \end{pmatrix} \mathsf{M}_2 \begin{pmatrix} \text{Re}\,\hat{H}_u^0 \\ \text{Re}\,\hat{H}_d^0 \end{pmatrix} + \begin{pmatrix} \hat{H}_u^+ & \hat{H}_d^- \end{pmatrix}^* \mathsf{M}_3 \begin{pmatrix} \hat{H}_u^- \\ \hat{H}_d^+ \end{pmatrix}, \\
\mathsf{M}_1 &= \frac{m_A^2}{2} \begin{pmatrix} 1+\cos 2\beta & \sin 2\beta \\ \sin 2\beta & 1-\cos 2\beta \end{pmatrix}, \\
\mathsf{M}_2 &= \frac{1}{2} \begin{pmatrix} m_A^2(1+\cos 2\beta) + m_Z^2(1-\cos 2\beta) & (m_A^2+m_Z^2)\sin 2\beta \\ (m_A^2+m_Z^2)\sin 2\beta & m_A^2(1-\cos 2\beta) + m_Z^2(1+\cos 2\beta) \end{pmatrix}, \\
\mathsf{M}_3 &= \frac{m_W^2+m_A^2}{2} \begin{pmatrix} 1+\cos 2\beta & \sin 2\beta \\ \sin 2\beta & 1-\cos 2\beta \end{pmatrix}.
\end{aligned}
$$

The mass matrices are not diagonal (and very messy), so we need to diagonalize them. The physical spectrum contains:

- two neutral CP-even Higgs bosons: h^0 (lighter), H^0 (heavier) (from the real parts),
- one neutral CP-odd Higgs boson: A^0 (from the imaginary parts),
- two charged Higgs bosons: $H^\pm$.

The masses are

$$
\begin{aligned}
m_A^2 &= 2|\mu|^2 + m_u^2 + m_d^2, \\
m_{H^\pm}^2 &= m_A^2 + m_W^2 \\
m_{h^0,H^0}^2 &= \frac{1}{2}\left(m_A^2 + m_Z^2 \pm \sqrt{(m_A+m_Z)^2 - 4m_A^2 m_Z^2 \cos^2(2\beta)} \right).
\end{aligned}
$$

Since $\cos(2\beta) \leq 1$, there is an upper bound on the mass of the lightest Higgs boson at tree level $m_{h^0} \leq m_Z$. For $m_A^2 \gg m_{W,Z}^2$, all Higgs masses scale with m_A (except h^0). When we do a loop correction for the h^0 mass, the upper bound is lifted to $m_{h^0} \lesssim 130\,\text{GeV}$, which comes primarily from the top quark and stop squark loops.



Chapter 2

Spin-2 fields

We have already discussed spin-0, spin-1/2 and spin-1 fields at length, and now we will skip spin-3/2 (applicable to gravitinos, described by Rarita–Schwinger equation). This chapter will be dedicated to spin-2 fields and the general relativity as an effective quantum field theory of weak gravitational fields. This will conclude the analysis of elementary relativistic quantum fields, as fields with higher spin cannot be consistently defined or lack any applicability.

2.1 The massive spin-2 field

Remember that vector fields were built from the $(1/2, 1/2)$ representation of the Lorentz group. This representation is equivalent to the vector representation A_μ (which is 4-dimensional). This representation can be reducible as a representation of the little group SO(3): $\mathbf{4} = \mathbf{1} \oplus \mathbf{3}$. Similarly, we will get the spin-2 representation by multiplying two vector representations:

$$(1/2,1/2) \otimes (1/2,1/2) = \underbrace{(1,0) \oplus (0,1)}_{\text{antisymmetric tensor} F_{\mu\nu}} \oplus \underbrace{(1,1) \oplus (0,0)}_{\text{symmetric tensor} h_{\mu\nu}}.$$

However, this representation is not irreducible as it can be represented by any Lorentz tensor of rank 2 $B_{\mu\nu}$. It can be split into symmetric (10 dimensions) and antisymmetric parts (6 dimensions, a sum of self-dual and anti-self-dual form). Under the little group SO(3), the symmetric tensor decomposes as

$$\mathbf{10} = \mathbf{1} \oplus \mathbf{1} \oplus \mathbf{3} \oplus \mathbf{5}.$$

The parts are h_{00}, h_{ii}, h_{0i} and $h_{ij} - \delta_{ij} h_{kk}/3$. With the spin-1 field, we demanded that only the non-trivial part of the little group representation remains dynamical (the **3** part). We split

$$A_\mu = A^T_\mu + \partial_\mu \pi, \quad \text{with } \partial^\mu A^T_\mu = 0.$$

The most general bilinear terms in the Lagrangian for A_μ are

$$\begin{aligned}
\mathcal{L}_0 &= -\frac{a}{4} F_{\mu\nu} F^{\mu\nu} + \frac{b}{2} (\partial_\mu A^\nu)(\partial^\mu A_\nu) + \frac{c}{2} (\partial_\mu A^\mu)^2 + \frac{m^2}{2} A_\mu A^\mu \\
&= -\frac{a}{4} F^T_{\mu\nu} F^{T,\mu\nu} + \frac{b}{2} (\partial_\mu A^{T,\nu})(\partial^\mu A^T_\nu) + \frac{1}{2} m^2 A^T_\mu A^{T,\mu} + \\
&\quad + \frac{b}{2} (\partial_\mu \partial_\nu \pi)(\partial^\mu \partial^\nu \pi) + \frac{c}{2} (\partial^2 \pi)^2 + \frac{1}{2} m^2 (\partial_\mu \pi)(\partial^\mu \pi).
\end{aligned}$$

Note that the first and the second term of A_μ^T are not independent. The propagator for the longitudinal mode π contains

$$\frac{i}{m^2k^2-(b+c)k^4} = \frac{1}{m^2}\left(\frac{i}{k^2} - \frac{i}{k^2-m^2/(b+c)}\right).$$

The negative sign in the second term indicates a negative norm state (ghost). To avoid this, we set $b=-c$. Together with the dependence of the first two terms, we can set $b=c=0$ and rescale the fields to get $a=1$ (canonical normalization):

$$\mathcal{L}_0 = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}m^2A_\mu A^\mu.$$

If we calculate the conjugate momentum, we see that its time component vanishes identically.

Now, we will repeat the same procedure for the symmetric tensor $h_{\mu\nu}$. The most general bilinear terms are (using the shorthand $h = h^\mu{}_\mu$)

$$\mathcal{L}_0 = \frac{1}{2}ah^{\mu\nu}\Box h_{\mu\nu} - \frac{1}{2}bh^{\mu\nu}\partial_\mu\partial^\alpha h_{\nu\alpha} + \frac{1}{2}ch\Box h + \frac{1}{2}dh\partial_\mu\partial_\nu h^{\mu\nu} - \frac{1}{2}m^2h^{\mu\nu}h_{\mu\nu} - \frac{1}{2}n^2h^2.$$

The terms were already collected together as much as possible using integration by parts. We decompose the field into irreducible representations of the little group:

$$h_{\mu\nu} = h^{\mathrm{T}}_{\mu\nu} + \partial_\mu\pi_\nu + \partial_\nu\pi_\mu = h^{\mathrm{T}}_{\mu\nu} + \partial_\mu\pi^{\mathrm{T}}_\nu + \partial_\nu\pi^{\mathrm{T}}_\mu + 2\partial_\mu\partial_\nu\pi^{\mathrm{L}},$$

with

$$\partial^\mu h^{\mathrm{T}}_{\mu\nu} = 0, \quad \partial^\mu\pi^{\mathrm{T}}_\mu = 0, \quad \pi_\mu = \pi^{\mathrm{T}}_\mu + \partial_\mu\pi^{\mathrm{L}}.$$

After inserting this into the Lagrangian, we want to avoid terms that have more than two derivatives on π^{L} or π^{L} (that would imply ghosts). The problematic terms are:

$$\mathcal{L}_0 \supset -\frac{1}{2}\pi^{\mathrm{L}}\partial^4\pi^{\mathrm{L}}(4m^2+4n^2) + \frac{1}{2}\pi^{\mathrm{L}}\partial^6\pi^{\mathrm{L}}4(a+b+c+d) + \frac{1}{2}\pi^{\mathrm{T},\mu}\partial^4\pi^{\mathrm{T}}_\mu(-2a-b) + \frac{1}{2}h\partial^4\pi^{\mathrm{L}}(8c+4d).$$

In order to avoid ghosts, we set

$$m^2+n^2=0, \quad b=-2a, \quad c=-a, \quad d=2a$$

and $a=-1$ for the canonical normalization. The Lagrangian for the free massive spin-2 field becomes

$$\mathcal{L}_0 = \frac{1}{2}\partial^\lambda h_{\mu\nu}\partial_\lambda h^{\mu\nu} - \partial_\mu h^{\mu\lambda}\partial^\nu h_{\nu\lambda} - \frac{1}{2}(\partial^\lambda h)(\partial_\lambda h) + \partial^\mu h\partial^\nu h_{\mu\nu} - \frac{1}{2}m^2(h^{\mu\nu}h_{\mu\nu} - h^2).$$

This is called the *Fierz–Pauli Lagrangian* (free theory of a massive spin-2 particle). The tensor $h^{\mathrm{T}}_{\mu\nu}$ has six components (not 5), however, the equations of motion give the constraint $h=0$ (requiring the tensor to be transverse and traceless). The Euler–Lagrange equations give

$$\partial^2h_{\mu\nu} - \partial_\mu\partial^\lambda h_{\lambda\nu} - \partial_\nu\partial^\lambda h_{\lambda\mu} - \eta_{\mu\nu}\partial^2h + \eta_{\mu\nu}\partial^\lambda\partial^\omega h_{\lambda\omega} + \partial_\mu\partial_\nu h = -m^2(h_{\mu\nu} - \eta_{\mu\nu}h).$$

When we differentiate with ∂^μ, the left-hand side vanishes, and it follows (for $m\neq 0$) $\partial^\mu h_{\mu\nu} = \partial_\nu h$. We insert this into the Euler–Lagrange equation and get

$$\partial^2h_{\mu\nu} - \partial_\mu\partial_\nu h = -m^2(h_{\mu\nu} - \eta_{\mu\nu}h).$$

The trace of this is zero on the left-hand side, so we get $h=0$. Moreover, we have $\partial^\mu h_{\mu\nu}=0$ and $(\partial^2+m^2)h_{\mu\nu}=0$ for each component. This means that every component of $h_{\mu\nu}$ satisfies the Klein–Gordon equation and that $h_{\mu\nu}$ carries only 5 dynamical degrees of freedom (the **5** representation of the little group). In the nature, however, there are no known massive spin-2 particles. We will thus move to the massless case.

2.2 The massless spin-2 field

There are several problems that we encounter in the limit $m \to 0$. The issues are exactly the same as those in the massless spin-1 case:

- We cannot construct a Lorentz covariant field theory for helicity ± 2 particles from the symmetric tensor representation. We would need to have a representation (A, B) such that $|A - B| = 2$. The simplest choice is $(2, 0) \oplus (0, 2)$, but this does not generate a $1/r$ static force. This is the same reason why we didn't choose $(1, 0) \oplus (0, 1)$ for the photon.

- The polarization sum
$$\sum_{\lambda=0,\pm1,\pm2} \varepsilon^{\mu\nu}(p,\lambda)\varepsilon^{*\rho\sigma}(p,\lambda)$$
contains singular $1/m^2$ terms (from $\lambda = 0, \pm 2$). If we direct the momentum along the z-axis, the $\lambda = \pm 2$ polarizations are of the form
$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & \pm i & 0 \\ 0 & \pm i & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

- We can try to manually remove the unphysical polarizations and construct the free field
$$h_{\mu\nu}(x) = \sum_{\lambda=\pm2} \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3 2E_{\mathbf{p}}} \left(e^{-ipx}\varepsilon_{\mu\nu}(p,\lambda)a(p,\lambda) + \text{h.c.}\right).$$
Under a Lorentz transformation, we get additional terms that do not vanish:
$$\mathcal{U}(\Lambda)h^{\mu\nu}(x)\mathcal{U}^{-1}(\Lambda) = (\Lambda^{-1})^{\mu}{}_{\rho}(\Lambda^{-1})^{\nu}{}_{\sigma}h^{\rho\sigma}(\Lambda x) + \partial^{\mu}\xi^{\nu}(\Lambda x) + \partial^{\nu}\xi^{\mu}(\Lambda x),$$
where ξ^{μ} is some function involving the parameters of the little group transformation. For the non-covariance to be irrelevant for the dynamics, the action must be invariant under the gauge symmetry:
$$h_{\mu\nu}(x) \to h_{\mu\nu}(x) + \partial_{\mu}\xi_{\nu}(x) + \partial_{\nu}\xi_{\mu}(x).$$

- The Fierz–Pauli Lagrangian has a smooth limit $m \to 0$:
$$\mathcal{L}_0 = \frac{1}{2}\partial^{\lambda}h_{\mu\nu}\partial_{\lambda}h^{\mu\nu} - \partial_{\mu}h^{\mu\lambda}\partial^{\nu}h_{\nu\lambda} - \frac{1}{2}\partial^{\lambda}h\partial_{\lambda}h + \partial^{\mu}h\partial^{\nu}h_{\mu\nu}.$$
It is invariant (up to a total derivative) under the gauge transformation. Note that we explicitly constructed the Lagrangian to be ghost-free in the massive case, so we can choose $\pi^{\mu} = \xi^{\mu}$ and the gauge invariance is guaranteed.

Suppose we want to couple $h_{\mu\nu}$ to some "matter current" by adding the term
$$-\int \mathrm{d}^4x\, h^{\mu\nu}T_{\mu\nu}$$
to the action. Then the gauge invariance requires $\partial_{\mu}T^{\mu\nu} = 0$. This gives four Noether currents and corresponding conserved charges, which need to be a 4-vector
$$P^{\nu} = \int \mathrm{d}^3\mathbf{x}\, T^{0\nu}(x).$$

We only know one conserved 4-vector in nature: the energy-momentum 4-vector. This implies that $T^{\mu\nu}$ is the energy-momentum tensor and the gauge symmetry is local space-time translations $x^\mu \to x^\mu + \xi^\mu$. This also makes clear why one cannot have massless fields of spin more than 2 with long-range interactions. They would have to couple to a current associated with a tensorial charge, but there is no symmetry that can give such a charge.

Let's count the degrees of freedom. Naively, the symmetric tensor $h_{\mu\nu}$ has 10 components. The gauge symmetry removes 4 components and the equations of motion (the constraints $\partial^\mu h_{\mu\nu} - \partial_\nu h = 0$) remove another 4 components. We are left with 2 dynamical degrees of freedom, which correspond to the helicity ± 2 states. The remaining dynamical states reside in the transverse-traceless part (s^{T}_{ij}) of the tensor h^{T}_{ij}:

$$h_{ij} = \frac{1}{3}\delta_{ij}h_{kk} + s_{ij} = \frac{1}{3}\delta_{ij}h_{kk} + s^{\mathrm{T}}_{ij} + s^{\mathrm{V}}_{ij} + s^{\mathrm{S}}_{ij}.$$

In the non-relativistic limit, it holds $T^{00} \propto m$, so we conclude that this force has to couple to mass. Another hint towards gravity is that the gauge symmetry is a local version of general coordinate transformations (which is the symmetry of general relativity). With this we can construct GR from the bottom up from $h_{\mu\nu}$. Remember first how the non-Abelian gauge transformations were constructed:

$$A^A_\mu \to A^A_\mu + \partial_\mu \varepsilon^A + f^{ABC}\varepsilon^C A^B_\mu.$$

The other index of the gauge field A^A_μ can be replaced by a space-time index for $h_{\mu\nu}$:

$$h_{\mu\nu} \to h_{\mu\nu} + \partial_\mu \varepsilon_\nu + \partial_\nu \varepsilon_\mu + \text{terms like } \varepsilon \times h.$$

However, the extra terms at the end can make the Lagrangian non-Lorentz invariant. To avoid this, we have to add new terms of higher order in $h_{\mu\nu}$ to the Lagrangian. Repeating this procedure infinitely many times, we get the full Einstein–Hilbert action of general relativity:

$$S_{\mathrm{GR}} = \frac{1}{16\pi G_{\mathrm{N}}} \int \mathrm{d}^4x \, \sqrt{-g}R,$$

where $g = \det(g_{\mu\nu})$ and R is the Ricci scalar.

2.3 GR as an effective field theory of gravitation

We have shown that GR can be constructed from the massless spin-2 field $h_{\mu\nu}$. Now, we will treat GR as an effective quantum field theory of weak gravitational fields.

2.3.1 Classical general relativity

GR is a classical relativistic field theory with the fundamental field being $g_{\mu\nu}(x)$. Why do we choose this field?

- The charge is P^μ, the current is $T^{\mu\nu}$, so the field should be a symmetric tensor like $T^{\mu\nu}$.
- Gravity is a long-range force and always attractive (and couples to mass). Non-relativistically, the Higgs field can couple to mass as $V(r) = -2m_1m_2/v^2 \cdot r^{-1}$. However, the Higgs doesn't couple to photons, to which the gravity does couple (bending of light). So $g_{\mu\nu}$ has to be spin-2.

The Einstein–Hilbert action follows almost uniquely from the nature of the fundamental field, the requirements of covariance and charge conservation and the Newtonian limit, $g_{\mu\nu}$ as a field and second order in derivative.

From the metric, one constructs the affine connection (*Christoffel symbols*)

$$\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2} g^{\lambda\sigma}(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu}).$$

The symmetry of the theory is the group of general coordinate transformations (diffeomorphisms). The covariant derivative for a general tensor is defined with the Christoffel symbols

$$\nabla_\lambda T^{\mu\nu\dots}{}_{\rho\sigma\dots} = \partial_\lambda T^{\mu\nu\dots}{}_{\rho\sigma\dots} + \Gamma^{\mu}{}_{\lambda\alpha} T^{\alpha\nu\dots}{}_{\rho\sigma\dots} + \Gamma^{\nu}{}_{\alpha\lambda} T^{\mu\alpha\dots}{}_{\rho\sigma\dots} + \dots - \Gamma^{\alpha}{}_{\lambda\rho} T^{\mu\nu\dots}{}_{\alpha\sigma\dots} - \dots$$

From the connection, one constructs the *Riemann curvature tensor*

$$R^{\rho}{}_{\sigma\mu\nu} = \partial_\mu \Gamma^{\rho}{}_{\nu\sigma} - \partial_\nu \Gamma^{\rho}{}_{\mu\sigma} + \Gamma^{\rho}{}_{\mu\lambda}\Gamma^{\lambda}{}_{\nu\sigma} - \Gamma^{\rho}{}_{\nu\lambda}\Gamma^{\lambda}{}_{\mu\sigma},$$

and further the *Ricci tensor* and the *Ricci scalar*

$$R_{\mu\nu} = R^{\lambda}{}_{\mu\lambda\nu}, \quad R = g^{\mu\nu} R_{\mu\nu}.$$

The general relativity has a nice geometric interpretation. The line element is $\mathrm{d}s^2 = g_{\mu\nu}(x)\mathrm{d}x^\mu\,\mathrm{d}x^\nu$. The theory is invariant under the general coordinate transformation

$$x^\mu \to x'^\mu(x), \quad \mathrm{d}x'^\mu = \frac{\partial x'^\mu}{\partial x^\nu}(x)\,\mathrm{d}x^\nu.$$

The Jacobian of the transformation is x-dependent. There is also the inverse of the metric $g^{\mu\nu}$, which satisfies $g^{\mu\lambda}g_{\lambda\nu} = \delta^\mu{}_\nu$. Using the appropriate Newtonian limit and the second order in derivatives, we get the *Einstein field equations*:

$$\boxed{R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \Lambda g_{\mu\nu} = -8\pi G_{\mathrm{N}} T_{\mu\nu}.}$$

This is the equation of motion for the *Einstein–Hilbert action*

$$S_{\mathrm{EH}} = \frac{1}{16\pi G_{\mathrm{N}}} \int \mathrm{d}^4x\, \sqrt{-g}(R - 2\Lambda)$$

after varying with respect to $g_{\mu\nu}$ (without matter coupling). The full action is $S = S_{\mathrm{EH}}[g_{\mu\nu}] + S_{\mathrm{M}}[g_{\mu\nu}, \phi, \dots]$, where S_{M} is the matter action. The energy-momentum tensor is *defined* as

$$\delta S_{\mathrm{M}} = -\frac{1}{2} \int \mathrm{d}^4x\, \sqrt{-g} T^{\mu\nu}(x)\, \delta g_{\mu\nu}.$$

The mass dimensions of these objects are $[g_{\mu\nu}] = 0$, $[R] = 2$, $[\Lambda] = 2$, $[G_{\mathrm{N}}] = -2$ and $[T_{\mu\nu}] = 4$. The action is a scalar and is thus invariant under general coordinate transformations. The metric transforms as

$$g_{\mu\nu}(x) \to g'_{\mu\nu}(x') = \frac{\partial x^\alpha}{\partial x'^\mu} \frac{\partial x^\beta}{\partial x'^\nu} g_{\alpha\beta}(x).$$

Infinitesimally ($x'^\mu = x^\mu + \varepsilon^\mu(x)$), this becomes

$$\delta g_{\mu\nu}(x) = g'_{\mu\nu}(x) - g_{\mu\nu}(x) = -\partial_\mu \varepsilon^\rho g_{\rho\nu} - \partial_\nu \varepsilon^\rho g_{\mu\rho} - \varepsilon^\rho \partial_\rho g_{\mu\nu}.$$

If $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ with $|h_{\mu\nu}| \ll 1$, we get $\delta h_{\mu\nu} = -\partial_\mu \varepsilon_\nu - \partial_\nu \varepsilon_\mu$, which is exactly the gauge transformation of the massless spin-2 field.

2.3.2 Quantization of weak gravitational fields

We will quantize fluctuations around the flat Minkowski space-time: $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$. The Minkowski metric is the vacuum expectation value of the metric field

$$\langle\Omega|g_{\mu\nu}|\Omega\rangle = \eta_{\mu\nu}.$$

This also means that $\eta_{\mu\nu}$ has to be a solution of the classical field equations in vacuum, which is the case if we set $\Lambda = 0$ (if $\Lambda \neq 0$ we would need to quantize around de Sitter or anti-de Sitter background, which is quite difficult). The Lagrangian we are left with is

$$\mathcal{L}_{\text{EH}} = \frac{2}{\kappa^2}\sqrt{-g}R, \qquad \kappa = \sqrt{32\pi G_{\text{N}}}.$$

We want to expand this Lagrangian in powers of $h_{\mu\nu}$:

$$-g = -\det(\eta_{\mu\rho}(\delta^\rho{}_\nu + h^\rho{}_\nu)) = -\det(\eta_{\mu\nu})\exp(\text{tr}(\ln(\mathbb{1} + h))) = \exp(h^\alpha{}_\alpha - h^\alpha{}_\beta h^\beta{}_\alpha/2 + \dots).$$

Including also the square root, we get

$$\sqrt{-g} = 1 + \frac{1}{2}h^\alpha{}_\alpha + \frac{1}{8}(h^\alpha{}_\alpha)^2 - \frac{1}{4}h^\alpha{}_\beta h^\beta{}_\alpha + \dots$$

The expansion for R is

$$\begin{aligned} R &= g^{\mu\nu}R_{\mu\nu} = (\eta^{\mu\nu} - h^{\mu\nu})(R^{(1)}_{\mu\nu} + R^{(2)}_{\mu\nu} + \dots) \\ &= (\eta^{\mu\nu} - h^{\mu\nu})(\frac{1}{2}\partial^\rho\partial_\mu h_{\nu\rho} + \frac{1}{2}\partial^\rho\partial_\nu h_{\mu\rho} - \frac{1}{2}\partial^2 h_{\mu\nu} - \frac{1}{2}\partial_\mu\partial_\nu h^\rho{}_\rho + \dots). \end{aligned}$$

The easiest approach is to expand the Lagrangian in powers of h:

$$\mathcal{L}_{\text{EH}} = \mathcal{L}^{(1)} + \mathcal{L}^{(2)} + \dots$$

The first order term is (using $h = h^\alpha{}_\alpha$)

$$\frac{\kappa^2}{2}\mathcal{L}^{(1)} = \partial^\rho\partial^\mu h_{\mu\rho} - \partial^2 h,$$

and it is a total derivative, so it can be dropped. The second order term is exactly the same as the kinetic term of the massless spin-2 field:

$$\mathcal{L}^{(2)} = \frac{1}{\kappa^2}\left(\frac{1}{2}\partial^\lambda h_{\mu\nu}\partial_\lambda h^{\mu\nu} - \partial_\mu h^{\mu\lambda}\partial^\nu h_{\nu\lambda} - \frac{1}{2}\partial^\lambda h\partial_\lambda h + \partial^\mu h\partial^\nu h_{\mu\nu}\right).$$

In order to do quantum field theory, we need to understand the propagators and vertices. For the propagator, we need to write the Lagrangian in a suitable form

$$\mathcal{L}_{\text{EH}} = \frac{1}{\kappa^2}\left(-\frac{1}{2}h_{\mu\nu}\text{D}^{\mu\nu\rho\sigma}h_{\rho\sigma} + \mathcal{O}(h^3)\right) = -\frac{1}{2}\hat{h}_{\mu\nu}\text{D}^{\mu\nu\rho\sigma}\hat{h}_{\rho\sigma} + \kappa\mathcal{O}(\hat{h}^3) + \kappa^2\mathcal{O}(\hat{h}^4),$$

where $\text{D}^{\mu\nu\rho\sigma}$ is some second-order differential operator. We defined the canonically normalized field $\hat{h}_{\mu\nu} = h_{\mu\nu}/\kappa$ with $[\hat{h}] = 1$. As long as $|\kappa\hat{h}| \ll 1$, the expansion is valid. Sadly, the operator $\text{D}^{\mu\nu\rho\sigma}$ is not invertible because of the gauge symmetry. So, let's fix the gauge! We need to make a choice, whether to break covariance or gauge invariance. Most of the time, the covariance is preferred, so we choose the physical gauge. This is the analogue of transverse gauge in QED,

$\nabla \cdot \mathbf{A} = 0$, where the two degrees of freedom are transverse to the direction of motion. The electric potential A^0 vanishes in the vacuum, however in the presence of charges, it obeys the Poisson equation $\nabla^2 A^0 = -4\pi\rho$. With the spin-2 field, the transverse gauge (4 constraints) reads

$$\partial_i h_{i0} = 0, \quad \partial_i s_{ij} = 0$$

and the equations of motion give the additional 4 constraints

$$\partial^\mu h_{\mu\nu} = \partial_\nu h, \quad h = 0, \quad \partial^\mu h_{\mu\nu} = 0, \quad \partial^2 h_{\mu\nu} = 0.$$

The first two constraints give $\dot{h}_{00} =$ and then $\nabla^2 h_{00} = 0$. The vacuum initial conditions give $h_{00} = 0$. The spatial trace also vanishes $h^i{}_i = h - h_{00} = 0$. The third constraint gives for $\nu = i$

$$0 = \dot{h}_{0i} - \partial_j h_{ji} = \dot{h}_{0i} - \partial_j(\delta_{ij} h^k_k + s_{ij}) = \dot{h}_{0i}.$$

With the initial condition and $\nabla^2 h_{0i} = 0$, the vacuum solution is $h_{0i} = 0$. To recap, we have

$$h_{00} = h_{0i} = h_{0\mu} = 0$$

in the absence of sources. If there are sources, these components are non-zero, but non-dynamical, determined from

$$\nabla^2 h_{0\mu} \propto T_{0\mu}.$$

The only dynamical components are in the transverse-traceless part s^{T}_{ij}, which, in the absence of sources ($\partial^2 s^{\mathrm{T}}_{ij} = 0$), admit plane wave solutions with helicity ± 2. When we quantize them, we get operators in the Fock space of gravitons:

$$s^{\mathrm{T}}_{ij}(x) = \sum_{\lambda=\pm 2} \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3} \frac{1}{2E_{\mathbf{k}}} \left(e^{-ikx} \varepsilon_{ij}(p,\lambda) a(p,\lambda) + \text{h.c.} \right).$$

The operators a and $a^\dagger$ destroy and create *gravitons* with helicity ± 2.

Equally, we could use the *de Donder gauge* (harmonic gauge):

$$\partial^\nu h_{\mu\nu} - \frac{1}{2} \partial_\mu h^\nu{}_\nu = 0.$$

Applying the *Faddeev–Popov method* gives the generating functional integral

$$Z[\mathcal{J}_{\mu\nu}] = N \int \mathcal{D}[h_{\mu\nu}]\, G[f_\mu[h]] \det(M) \exp\left(iS_{\mathrm{EH}} + i \int \mathrm{d}^4x\ \mathcal{J}_{\mu\nu} h^{\mu\nu} \right).$$

The gauge-fixing functional is usually taken to be Gaussian

$$G[f_\mu] = \exp\left(+ \frac{i}{\xi} \int \mathrm{d}^4x\ f_\mu(x) f^\mu(x) \right),$$

which is an equivalent of the R_ξ gauge in non-Abelian gauge theories. The Faddeev–Popov determinant $\det(M)$ is

$$M_{\mu\nu}(x,y) = \left. \frac{\delta f_\mu[h^\varepsilon_{\mu\nu}(x)]}{\delta \varepsilon^\nu(y)} \right|_{\varepsilon^\nu = 0} = M_{\mu\nu}[h](x) \delta^{(4)}(x-y).$$

The determinant can be written as a path integral over Grassmann fields η_μ and $\overline{\eta}_\mu$ (ghost fields)

$$\det M = N \int \mathcal{D}[\eta_\mu, \overline{\eta}_\mu] \exp\left(i \int \mathrm{d}^4x \, \overline{\eta}^\mu M_{\mu\nu}[h](x) \eta^\nu \right).$$

In summary the full Lagrangian consists of the graviton self-interactions, the gauge-fixing term, the ghost term and the matter coupling:

$$\mathcal{L} = \mathcal{L}_{\mathrm{EH}} + \mathcal{L}_{\mathrm{gf}} + \mathcal{L}_{\mathrm{gh}} + \mathcal{L}_M.$$

The bilinear terms in the Lagrangian can be taken together to get the graviton propagator

$$\mathcal{L}_{\mathrm{bilin}} = -\frac{1}{2} h_{\mu\nu} M^{\mu\nu,\rho\sigma} \partial^2 h_{\rho\sigma},$$

where

$$M_{[\mu\nu][\rho\sigma]} = \frac{1}{2} \left(\eta_{\mu\rho}\eta_{\nu\sigma} + \eta_{\mu\sigma}\eta_{\nu\rho} - \eta_{\mu\nu}\eta_{\rho\sigma} \right).$$

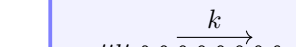
The propagator will be

$$\frac{i}{k^2 + i\varepsilon} N_{[\mu\nu][\rho\sigma]}$$

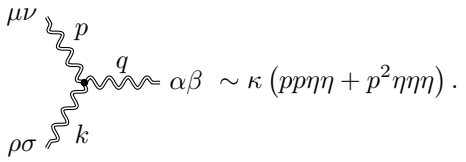
where the tensor $N_{[\mu\nu][\rho\sigma]}$ satisfies

$$M^{[\mu\nu][\alpha\beta]} N_{[\alpha\beta][\rho\sigma]} = \frac{1}{2} (\delta^\mu{}_\rho \delta^\nu{}_\sigma + \delta^\mu{}_\sigma \delta^\nu{}_\rho).$$

The graviton propagator in the de Donder gauge is

$\mu\nu$  $\rho\sigma$ $= \dfrac{i}{k^2 + i\varepsilon} \dfrac{1}{2} \left(\eta_{\mu\rho}\eta_{\nu\sigma} + \eta_{\mu\sigma}\eta_{\nu\rho} - \eta_{\mu\nu}\eta_{\rho\sigma} \right).$

The vertices of the graviton self-interactions go to arbitrary high order in κ (and the number of legs). For n legs, the prefactor is κ^{n-2} and there are always two derivatives acting on the fields (from Riemann tensor). The three point self-interaction vertex is (roughly)



$\sim \kappa \left(pp\eta\eta + p^2\eta\eta\eta \right).$

Similarly, the 4-point vertex is $\kappa^2(pp\eta\eta\eta + p^2\eta\eta\eta\eta)$.

2.3.3 Interaction with matter

The graviton also couples to matter:

$$S_M = \int \mathrm{d}^4x \, \sqrt{-g} \mathcal{L}_M(\psi, \nabla_\lambda \psi).$$

It is easiest to start with a scalar field ϕ (because the covariant derivative is just the partial derivative)

$$\mathcal{L}_\phi = \frac{1}{2}(\partial_\mu\phi\partial^\mu\phi - m^2\phi^2) = \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}m^2\phi^2.$$

The coupling goes via the energy-momentum tensor

$$S_M = S_{\phi,\text{free}} - \kappa\int \mathrm{d}^4x\, h^{\mu\nu}T_{\mu\nu} + \mathcal{O}(h^2),$$

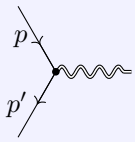
where

$$T_{\mu\nu} = \frac{1}{2}\left(\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}\eta_{\mu\nu}\partial_\rho\phi\partial^\rho\phi + \frac{1}{2}m^2\eta_{\mu\nu}\phi^2\right).$$

This is called the *minimal coupling* of matter to gravity. There could also exist an extra term

$$S_M \supset -\frac{1}{2}\int \mathrm{d}^4x\,\sqrt{-g}\xi R\phi^2$$

where ξ is dimensionless coupling constant. However, we won't examine this interaction in this course. For the minimally coupled scalar vertex (for $\xi = 0$) we get



$$= -\frac{i\kappa}{2}(p_\mu p'_\nu + p_\nu p'_\mu - \eta_{\mu\nu}(p\cdot p' - m^2)).$$

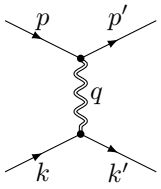
The arrows on the propagators only denote the direction of the momentum and not the fermion nature of the particle. Coupling to the electromagnetic field is

$$S_A = \int \mathrm{d}^4x\,\sqrt{-g}\left(-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}\right) = -\frac{1}{4}\int \mathrm{d}^4x\,\sqrt{-g}g^{\mu\alpha}g^{\nu\beta}F_{\mu\nu}F_{\alpha\beta},$$

where

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

Now, we will consider the scattering of non-relativistic scalar particles (aka planets) through graviton exchange (note that the scalars are denoted by fermion lines in order to signal the direction of momentum flow):



$$= \frac{-i\kappa}{2}(p_\mu p'_\nu + p_\nu p'_\mu - \eta_{\mu\nu}(p\cdot p' - m^2))\frac{i(\eta^{\mu\rho}\eta^{\nu\sigma} + \eta^{\mu\sigma}\eta^{\nu\rho} - \eta^{\mu\nu}\eta^{\rho\sigma})/2}{q^2 + i\varepsilon}\cdot$$
$$\cdot\frac{-i\kappa}{2}(k_\rho k'_\sigma + k_\sigma k'_\rho - \eta_{\rho\sigma}(k\cdot k' - M^2)).$$

In the non-relativistic limit, $p^\mu \to mv^\mu$ and $v^\mu = (1, \mathbf{0})$. Both vertex factors become $2m^2v^\mu v^\nu$ and after the contractions, we get

$$\mathcal{M} = \frac{(-i\kappa)^2}{4}2m^2 2M^2\frac{i}{q^2 + i\varepsilon}\frac{1}{2} \approx \frac{i\kappa^2}{2}\frac{m^2M^2}{\mathbf{q}^2}.$$

We can compare this result to the non-relativistic exchange photons:

$$\mathcal{M} = ie(p+p')^\mu \frac{-i\eta_{\mu\nu}}{q^2+i\varepsilon} ie(k+k')^\nu \approx -ie^2 \frac{4mM}{\mathbf{q}^2}.$$

As we can see, the coupling e is replaced by κm. It is more inspiring to compare the potentials in position space. The Fourier transform of $1/\mathbf{q}^2$ is

$$\int \frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3} e^{i\mathbf{q}\cdot\mathbf{r}} \frac{1}{\mathbf{q}^2} = -\frac{1}{4\pi r}.$$

The factors $2m$ and $2M$ are absorbed as the relativistic normalization of states. The potentials are thus

$$V_{\mathrm{em}}(r) = \frac{e^2}{4\pi r}, \quad V_{\mathrm{gr}}(r) = -\frac{\kappa^2}{32\pi}\frac{Mm}{r} = \frac{GMm}{r}.$$

The electromagnetic potential is repulsive for like charges, while the gravitational potential is always attractive. We should not be surprised by the familiar result, since we have constructed the theory to have the correct non-relativistic limit. Another important point is the mass in the Feynman diagram is inertial mass, and exactly this mass comes into the gravitational potential, which is a manifestation of the equivalence principle.

2.3.4 Renormalization

As we discussed before the Lagrangian is

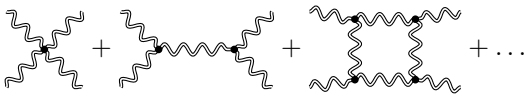
$$\mathcal{L} = \mathcal{L}_{\mathrm{EH}} + L_{\mathrm{GF+ghost}} = -\frac{1}{2} h_{\mu\nu} D^{\mu\nu\rho\sigma} h_{\rho\sigma} + \mathcal{L}_{\mathrm{ghost}} + \kappa\mathcal{O}(\partial^2 h^3) + \kappa^2 \mathcal{O}(\partial^2 h^4).$$

Since $[h] = 1$ and $[\kappa] = -1$ *all interactions are non-renormalizable* (starting at triple graviton self-interaction at dimension 5)! A vertex with n gravitons at scattering energy $E \ll 1/\kappa$ contributes a factor $(\kappa E)^{n-2}$. This is around the Planck scale

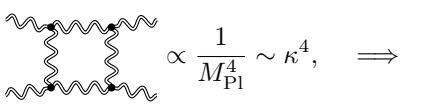
$$\frac{1}{\kappa} = \frac{1}{\sqrt{32\pi G_{\mathrm{N}}}} = \frac{M_{\mathrm{Pl}}}{2} = 1.218 \cdot 10^{18}\,\mathrm{GeV}.$$

For an effective (non-renormalizable) QFT, quantum correction can be consistently computed as an expansion in E/M_{Pl} by adding all higher-dimensional operators (up to the dimension required by the desired accuracy) compatible with the symmetries of the action.

Example 1. Consider the graviton-graviton scattering



The $2 \to 2$ scattering is dimensionless, and the loops are UV-divergent. We need to add a counter-term to cancel the divergence.



This means that we have to add a local operator of the form $\partial\partial\partial\partial hhhh/M_{\mathrm{Pl}}^4$ to the Einstein–Hilbert action as a counter-term. We cannot absorb this into a redefinition of the parameters

in the original Lagrangian (like gravitational constant). The terms that are fourth order in derivatives and h are R^2, $R_{\mu\nu}R^{\mu\nu}$ and $R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}$. However, the last term can be written as a combination of the first two terms (Gauss–Bonnet theorem) and is not independent up to a topological term.

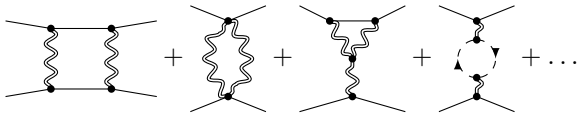
In pure gravity, the effective action is

$$S_{\text{eff}} = \int \mathrm{d}^4x\, \sqrt{-g}\left(-\tilde{\Omega} + \frac{1}{\kappa^2}R + c_1R^2 + c_2R_{\mu\nu}R^{\mu\nu} + \dots\right).$$

As a QFT, EH action has to be extended to the general EFT. The divergences at 1-loop are absorbed into renormalized c_1 and c_2 and so on.

Note. For $S_M = 0$, we have $R = R_{\mu\nu} = 0$ by equations of motion. So the one-loop graviton scattering amplitude is not divergent. But for $S_M \neq 0$, the terms R^2, $R_{\mu\nu}R^{\mu\nu}$ must appear.

We can return to the scattering of two massive scalars via graviton exchange and calculate the quantum corrections to the Newtonian potential.



When the momentum transfer is small, $|\mathbf{q}| \ll m, M$ (and $\mathbf{r} \gg 1/m_1, 1/m_2$) we can expand the scattering amplitude as

$$\mathcal{M} = -i\frac{G_{\text{N}}m_1m_2}{q^2}\left(\tilde{a} + \tilde{b}G_{\text{N}}(m_1+m_2)\sqrt{-q^2} + \tilde{c}G_{\text{N}}q^2\left(\ln\frac{-q^2}{\mu^2} + \tilde{d}\right) + \dots\right).$$

The last constant $\tilde{d}$ is dependent on the renormalization scheme. In the position space, the potential is

$$V(r) = -\frac{G_{\text{N}}m_1m_2}{r}\left(1 + b\frac{G_{\text{N}}(m_1+m_2)}{r} + c\frac{G_{\text{N}}}{r^2} + \dots\right) + d\delta^{(3)}(\mathbf{r}).$$

In 2017, the values of b and c were calculated to be $b = 3$ and $c = 41/(10\pi)$. To restore the units, we have

$$[G_{\text{N}}] = \frac{\mathrm{m}^3}{\mathrm{kg\,s}^2}, \quad [\hbar] = \mathrm{J\,s} = \frac{\mathrm{kg\,m}^2}{\mathrm{s}}, \quad [c] = \frac{\mathrm{m}}{\mathrm{s}}.$$

The potential in physical units is thus

$$V(r) = -\frac{G_{\text{N}}m_1m_2}{r}\left(1 + 3\frac{G_{\text{N}}(m_1+m_2)}{rc^2} + \frac{41}{10\pi}\frac{G_{\text{N}}\hbar}{r^2c^3} + \dots\right)$$

We can see that the first correction is classical (doesn't depend on $\hbar$, sometimes called "classical loop") and the second correction is quantum (depends on $\hbar$). The smallest distances in terrestrial high energy physics are of order 1 TeV, which amounts to $G_{\text{N}}/r^2 \sim 10^{-34}$. Experimentally, we are very far from testing these corrections. However, in the everyday life the gravity still plays the most important role, because the gravitational charges exist in exorbitant amounts (like the Earth mass).

2.4 van Dam–Veltman–Zakharov (vDVZ) discontinuity

For this section, we return to the massive spin-2 field. Remember that massive spin-1 fields are good EFTs for $E \ll \Omega = 4\pi m/g$. What happens with spin-2? For $m = 0$, the cut-off of the theory is M_{Pl}.

2.4.1 Tree level and 1-loop considerations

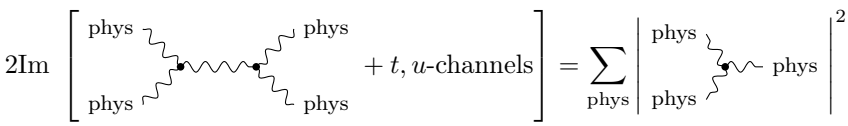
We expect the discontinuity in the limit $m \to 0$ for a field with the spin of more than 1/2 because of missing degrees of freedom. In the spin-1 case, we have the following polarization sum

$$\sum_{\lambda=0,\pm1} \varepsilon^\mu(k,\lambda)\varepsilon^\nu(k,\lambda) = -g^{\mu\nu} + \frac{k^\mu k^\nu}{m^2} \xrightarrow{\text{rest frame}} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

and the massless sum is

$$\sum_{\lambda=\pm1} \varepsilon^\mu(k,\lambda)\varepsilon^\nu(k,\lambda) = -g^{\mu\nu} + \frac{k^\mu \overline{k}^\nu + \overline{k}^\mu k^\nu}{k\cdot\overline{k}} \xrightarrow{\text{rest frame}} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

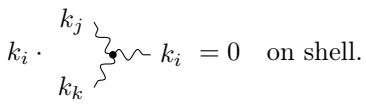
where $k = (|\mathbf{k}|, \mathbf{k})$ and $\overline{k} = (|\mathbf{k}|, -\mathbf{k})$. We have to check if the unitarity of the S-matrix is preserved in the limit $m \to 0$. If we denote $S = \mathbb{1} + iT$, then it has to hold $T^\dagger T = 2\text{Im}\, T$. This equation can be drawn with Feynman diagrams as



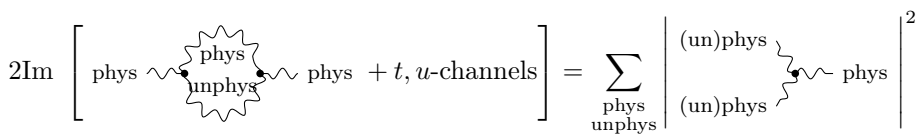
The relation is satisfied for $m \neq 0$ and $m = 0$ with smooth limit $m \to 0$, because

$$\text{Im}\left[\frac{1}{p^2 - m^2 + i\varepsilon}\right] = -i\pi\delta(p^2 - m^2).$$

For the on-shell legs the terms $k^\mu k^\nu/m^2$ drop out

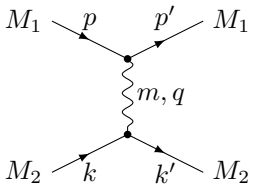


For the 1-loop case, we get into trouble, because the polarizations in the loop can also be non-physical. We fix this by introducing ghosts.



One finds that for $m \neq 0$, the vertex factor between the gauge boson and two ghosts must be a factor $1/\sqrt{2}$ smaller than for $m = 0$ to satisfy the unitarity relation. Hence, the limit $m \to 0$

is discontinuous at 1-loop level. One other thing to consider is non-relativistic scattering of the gauge boson between scalars:



$$\approx (-i)4M_1M_2\frac{e_1e_2}{\mathbf{q}^2+m^2}$$

which gives Yukawa potential in position space. The limit $m^2 \to 0$ is smooth, because the q^μ terms in the propagator/polarization sum do not contribute: from the vertex factor we get $(p+p')^\mu$ and from the propagator q_μ – their product is $q^\mu(p-p')_\mu = p^2-(p')^2 = 0$.

Let's return to the spin-2 case. The polarization tensors for massless graviton are

$$\varepsilon^{(+)}_{\mu\nu} = \frac{1}{\sqrt{2}}\begin{pmatrix}0&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&0\end{pmatrix}, \quad \varepsilon^{(\times)}_{\mu\nu} = \frac{1}{\sqrt{2}}\begin{pmatrix}0&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&0\end{pmatrix}.$$

The polarization sum becomes

$$\sum_{\lambda=+,\times}\varepsilon^{(\lambda)}_{\mu\nu}\varepsilon^{(\lambda)}_{\rho\sigma} = \begin{cases}\frac{1}{2}(\eta_{\mu\rho}\eta_{\nu\sigma}+\eta_{\mu\sigma}\eta_{\nu\rho}-\eta_{\mu\nu}\eta_{\rho\sigma}) & \mu,\nu,\rho,\sigma\in\{1,2\},\\ 0 & \text{otherwise.}\end{cases}$$

For the massive spin-2 field, the polarization tensor is a symmetric traceless matrix with five degrees of freedom

$$\varepsilon^{(1)}_{\mu\nu} = \begin{pmatrix}0&0&0&0\\0&\cdot&\cdot&\cdot\\0&\cdot&\cdot&\cdot\\0&\cdot&\cdot&\cdot\end{pmatrix}$$

The polarization sum is very similar to the massless case, however the limit is not continuous:

$$\sum_{\lambda=1}^{5}\varepsilon^{(\lambda)}_{\mu\nu}\varepsilon^{(\lambda)}_{\rho\sigma} = \begin{cases}\frac{1}{2}(\eta_{\mu\rho}\eta_{\nu\sigma}+\eta_{\mu\sigma}\eta_{\nu\rho})-\frac{1}{3}\eta_{\mu\nu}\eta_{\rho\sigma} & \mu,\nu,\rho,\sigma\in\{1,2,3\},\\ 0 & \text{otherwise.}\end{cases}$$

There is a difference even when $\{\mu,\nu,\rho,\sigma\}\in\{1,2\}$, which arises from the scalar polarization

$$\varepsilon^{(s=0)}_{\mu\nu} = \sqrt{\frac{2}{3}}\begin{pmatrix}0&0&0&0\\0&1/2&0&0\\0&0&1/2&0\\0&0&0&-1\end{pmatrix} \quad\to\quad \frac{1}{\sqrt{6}}\eta_{\mu\nu} \text{ for } \mu,\nu\in\{1,2\}.$$

The propagator for $m=0$ in de Donder gauge $\xi=1$ is simply

$$\frac{i}{k^2+i\varepsilon}\frac{1}{2}(\eta_{\mu\rho}\eta_{\nu\sigma}+\eta_{\mu\sigma}\eta_{\nu\rho}-\eta_{\mu\nu}\eta_{\rho\sigma})$$

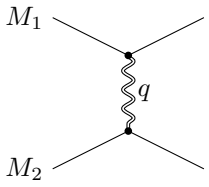
and in the massive case (homework) it is

$$\frac{i}{k^2-m^2+i\varepsilon}\left(\frac{1}{2}(\eta_{\mu\rho}\eta_{\nu\sigma}+\eta_{\mu\sigma}\eta_{\nu\rho})-\frac{1}{3}\eta_{\mu\nu}\eta_{\rho\sigma}\right).$$

Consider the scattering of massive bodies. At linear order, we have the interaction

$$T^{\mu\nu} \quad h_{\mu\nu} \quad = -\kappa.$$

The non-relativistic scattering amplitude is

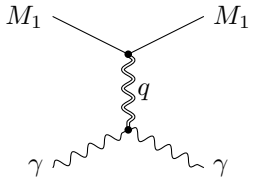


$$= (-i\kappa)^2 T^{\mu\nu}_{M_1} (\text{propagator})_{\mu\nu\rho\sigma} T^{\rho\sigma}_{\gamma}$$

$$\approx (-i\kappa)^2 T^{00}_{M_1} (\text{propagator})_{0000} T^{00}_{M_2}$$

$$= \begin{cases} -\frac{1}{2}\kappa^2 T^{00}_{M_1} T^{00}_{M_2} \frac{i}{q^2+i\varepsilon} & m = 0, \\ -\frac{2}{3}\kappa^2 T^{00}_{M_1} T^{00}_{M_2} \frac{i}{q^2-m^2+i\varepsilon} & m \neq 0. \end{cases}$$

Even at tree-level there is a discontinuity in the limit $m \to 0$. This is called the *vDVZ discontinuity*. However, this difference is not observable if the only measurements of G_{N} come from non-relativistic scattering (planetary motion). But, if we also include relativistic sources (like light bending around the Sun), then the difference becomes observable. The non-relativistic energy-momentum tensor for the photon still contains all terms, however, due to the symmetric properties of the propagator, only T^{00} contributes in the non-relativistic limit when coupled to massive bodies:



$$\approx (-i\kappa)^2 T^{00}_{M} (\text{propagator})_{00\rho\sigma} T^{\rho\sigma}_{\gamma}$$

$$= (-i\kappa)^2 T^{00}_{M} (\text{propagator})_{0000} T^{00}_{\gamma}$$

$$= -\kappa^2 T^{00}_{M} T^{00}_{\gamma} \cdot \begin{cases} \frac{i}{q^2-i\varepsilon} & m = 0, \\ \frac{i}{q^2-m^2+i\varepsilon} & m \neq 0. \end{cases}$$

There is *the same* factor in front of massless and massive case. If G_{N} was determined from, say, planetary motion, then we would predict that the light bending angle was

$$\lim_{m\to 0} \theta(m) = \frac{3}{4}\theta(m = 0).$$

Thus, any infinitesimal mass of the graviton would be detected by the bending of the light.

2.4.2 Significance of the vDVZ discontinuity

Remember the *Stuckelberg trick* for the massive, non-Abelian SU(2) vector field (first problem of the homework):

$$\mathcal{L} = -\frac{1}{2}\,\mathrm{tr}(F^{\mu\nu}F_{\mu\nu}) + m^2\,\mathrm{tr}(A^{\mu}A_{\mu}).$$

Perform a gauge transformation

$$A \to \mathcal{U}A\mathcal{U}^{\dagger} + \frac{i}{g}\mathcal{U}\partial_{\mu}\mathcal{U}^{\dagger}, \quad \mathcal{U} = e^{i\varepsilon^A T^A}, \quad T^A = \sigma^A/2.$$

The Lagrangian becomes

$$\mathcal{L}' = \mathcal{L} + \frac{m^2}{g^2}\,\mathrm{tr}((\partial_\mu \mathcal{U}^\dagger)(\partial^\mu \mathcal{U})) + \frac{2m^2}{g}\,\mathrm{tr}(A^\mu (i\partial_\mu \mathcal{U}^\dagger)\mathcal{U}).$$

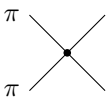
We make this formally gauge invariant by introducing $\mathcal{U} = e^{ig\pi^A T^A/m}$ with $\pi^A = m\varepsilon^A/g$ as a new scalar field by assigning to π^A a suitable transformation. The transformation is complicated, which at linear order acts as shift symmetry

$$\pi^A \to \pi^A + \varepsilon^A + \mathcal{O}(\varepsilon^2, \varepsilon\pi, \pi^2).$$

Because of the shift symmetry, we call π the Goldstone field. When we expand the Lagrangian to the fourth order in self-interactions of π (see first homework), we get

$$\mathcal{L}_{\text{Goldstone}} = \frac{1}{2}\partial_\mu \pi^A \partial^\mu \pi^A + \frac{g^2}{m^2}\pi^2 (\partial_\mu \pi)^2.$$

The last term dominates at high energies $E \gg m$, since every derivative brings a factor of E. The term $A^\mu(\partial_\mu \pi^A)\pi^A$ is thus sub-leading. This is the content of the *Goldstone boson equivalence theorem*: at high energies, the scattering amplitudes of longitudinally polarized massive vector bosons are equivalent to the scattering amplitudes of the corresponding Goldstone bosons. At tree level, the 4-point self interaction is



$$\sim \left(\frac{gE}{m}\right)^2.$$

The tree-level unitarity is broken at energies $E > 4\pi m/g$. The strong coupling scale $\Lambda > m$ is multiplied by a factor of $4\pi/g$, which is the sole reason that we actually can approach the theory from the EFT point of view. However, for $E > \Lambda$ we need the UV completion, an embedding of EFT into a more fundamental theory.

Let's go back to gravitons. We start with Fierz–Pauli Lagrangian and perform the gauge transformation $x^\mu \to x'^\mu = x^\mu + \varepsilon^\mu(x)$. From the tensor transformation of $g_{\mu\nu}$ we can derive

$$h_{\mu\nu}(x) \to h'_{\mu\nu}(x) = h_{\mu\nu}(x) + \partial_\mu \varepsilon_\nu + \partial_\nu \varepsilon_\mu + \partial_\mu \varepsilon^\alpha \partial_\nu \varepsilon_\alpha + \mathcal{O}(\varepsilon^3, \varepsilon h).$$

We interpret the ε_μ as new fields π^μ, which are pseudo-Goldstone bosons of a spontaneously broken diffeomorphism invariance. Hence, we substitute

$$h'_{\mu\nu}(x) = h_{\mu\nu}(x) + \partial_\mu \pi_\nu + \partial_\nu \pi_\mu + \partial_\mu \pi^\alpha \partial_\nu \pi_\alpha$$

into the Fierz–Pauli Lagrangian. It is also useful to decompose π_μ into transverse and longitudinal parts

$$\pi_\mu = \pi^{\mathrm{T}}_\mu + \partial_\mu \pi^{\mathrm{L}} = A_\mu + \partial_\mu \pi^{\mathrm{L}},$$

and we only keep the longitudinal part (we are only interested in $s = 0$ polarization, which is represented by π^{L}, also π^{L} has one derivative more and hence is enhanced more at higher energies). The new Lagrangian is the same as the original one with a few extra terms. After integrating these extra terms by parts to get rid of the derivative on the field π^{L}, we obtain

$$\mathcal{L}' = \mathcal{L}_{\mathrm{FP}} - \frac{m^2}{2} 4\pi^{\mathrm{L}} (\partial_\mu \partial_\nu h^{\mu\nu} - \Box h).$$

We eliminate the $\pi^{\rm L}\partial^2 h$ bilinear mixing terms (to diagonalize the kinetic term) by redefinitions of h:

$$h_{\mu\nu} \to h_{\mu\nu} - a\eta_{\mu\nu} m^2 \pi^{\rm L}, \quad h \to h - 4am^2\pi^{\rm L}.$$

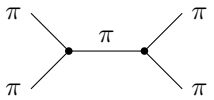
The constant a is fixed by requiring no mixing terms. Then

$$\mathcal{L} = L_{\rm FP} + h\pi^{\rm L} + \mathcal{L}_{\rm Goldstone}.$$

The term $h\pi^{\rm L}$ is sub-leading in relation to the Goldstone self-interactions at high-energy limit. The Goldstone part is

$$\begin{aligned}\mathcal{L}_{\rm Goldstone} &= \frac{M_{\rm Pl}^2}{2}\left(m^4\partial_\mu\pi^{\rm L}\partial^\mu\pi^{\rm L} + bm^2(\partial_\mu\partial_\nu\pi^{\rm L})(\partial^\mu\partial^\alpha\pi^{\rm L})(\partial^\nu\partial_\alpha\pi^{\rm L}) + \dots\right)\\ &= \frac{1}{2}(\partial_\mu\pi)^2 + \frac{1}{m^4 M_{\rm Pl}}(\partial^2\pi)^3 + \dots\end{aligned}$$

We have omitted the numerical factors and b is some constant. In the second line, we used the canonical normalization $\pi \to \pi/(m^2 M_{\rm Pl})$. These particles represent the $s = 0$ polarization of the graviton. The $2 \to 2$ scattering gives



$$\sim \left(\frac{E^6}{m^4 M_{\rm Pl}}\right)^2 \frac{1}{E^2}.$$

This gives the strong coupling scale at

$$\Lambda = \sqrt[5]{m^4 M_{\rm Pl}}.$$

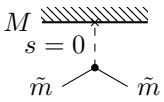
The validity of massive gravity is restricted to energies much smaller than for $m = 0$, namely $\Lambda \ll M_{\rm Pl}$. The vDVZ discontinuity appears after the redefinition of the field $h + am^2\pi$

$$T^{\mu\nu}h_{\mu\nu} \to T^{\mu\nu}h_{\mu\nu} - am^2\eta_{\mu\nu}T^{\mu\nu}\pi \xrightarrow{\text{canonical normalization}} T^{\mu\nu}h_{\mu\nu} - \frac{a}{M_{\rm Pl}}T^\mu{}_\mu\,\pi.$$

The $s = 0$ polarization (π degree of freedom) does not decouple in the limit $m \to 0$.

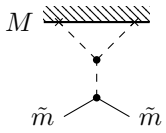
The last point of discussion in this chapter will be the *Vainshtein mechanism*. Consider a heavy body of mass M generating a classical field $\phi_{\rm cl}$ around itself. We want to estimate when the non-linear terms in the Goldstone Lagrangian become important. In the following diagrams the massive object will be denoted by a cross, and since it is so heavy, it can produce arbitrarily many gravitons without being affected itself.

At linear order (excluding the coupling to the test particle, forgetting $G_{\rm N}^{1/2}\tilde{m}$), the classical potential is



$$\phi_{\rm cl} \sim \frac{G_{\rm N}^{1/2}M}{r} \sim \frac{M}{M_{\rm Pl}}\frac{1}{r}.$$

The Goldstone 3-point interaction term gives the correction to the classical field as



$$\sim \frac{1}{m^4 M_{\rm Pl}} = \frac{1}{\Lambda^5}$$

and from this we can estimate the correction to the classical field as

$$\phi_{\rm cl}^{\rm non\text{-}lin} \sim \left(\frac{M}{M_{\rm Pl}}\right)^2 \frac{1}{\Lambda^5}\frac{1}{r^6}.$$

The exponent of r was determined by dimensional analysis and the two factors of $M/M_{\rm Pl}$ came from two vertices on the object M. The non-linear interaction dominates when

$$r \lesssim \frac{1}{\Lambda}\left(\frac{M}{M_{\rm Pl}}\right)^{1/5} = \left(\frac{G_{\rm N}M}{m^4}\right)^{1/5} \equiv r_{\rm V}.$$

This is called the *Vainshtein radius*. It can take macroscopic values if the mass of the graviton m is small enough. Within $r_{\rm V}$ the potential is almost the same as if the graviton was massless. Computation of the potential of heavy bodies breaks down at a scale

$$\frac{1}{r_{\rm V}} \sim m_{\rm grav}(G_{\rm N}M m_{\rm grav})^{-1/5} \ll m_{\rm grav}.$$

This implies no perturbative theory of massive gravitons at the scale of the graviton mass itself.

Chapter 3

Quantum field theory at finite temperature

We will now study QFT at finite temperature T and also at finite chemical potential μ. The uses of this treatment are the following:

- QCD at $T \gg \Lambda_{\mathrm{QCD}} \approx 200\,\mathrm{GeV} \sim 10^{12}\,\mathrm{K}$ when the quarks and gluons are not confined into hadrons,
- cosmology and EW interaction ($T \sim 150\,\mathrm{GeV}$).

We will assume that the system is in thermal equilibrium with the *grand canonical ensemble*, which means that the temperature T and chemical potential μ are fixed. The density matrix/operator of the system is

$$\rho = \frac{1}{Z} e^{-\beta(H-\mu N)}$$

and the key quantity we are interested is the *partition function*

$$Z = \mathrm{tr}\left(e^{-\beta(H-\mu N)}\right), \quad \beta = \frac{1}{T}.$$

We will use the natural units $\hbar = c = k_{\mathrm{B}} = 1$ and $\beta = 1/(k_{\mathrm{B}}T)$. The free energy of the theory is

$$F = -\frac{1}{\beta} \ln Z$$

in the canonical ensemble ($\mu = 0$) and we call it the *grand potential* ϕ in the grand canonical ensemble. The thermodynamic variables can be derived from ϕ:

$$\begin{aligned}
\text{pressure} \quad & p = -\left(\frac{\partial F}{\partial V}\right)_{T,\mu}, \\
\text{entropy} \quad & S = -\left(\frac{\partial F}{\partial T}\right)_{V,\mu}, \\
\text{energy} \quad & E = -\left(\frac{\partial \ln Z}{\partial \beta}\right)_{V,\mu} + \mu N.
\end{aligned}$$

For the canonical ensemble, we replace ϕ with F, and we don't keep track of μ. The energy density is

$$\frac{E}{V} = g_s \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \frac{\omega(\mathbf{p})}{e^{\beta(\omega(\mathbf{p})-\mu)} \mp 1}, \tag{$*$}$$

where $\omega(\mathbf{p}) = \sqrt{m^2 + \mathbf{p}^2}$. Positive sign is for fermions and negative for bosons.

3.1 Non-interacting gas

In elementary statistical physics, the partition function of the ideal Bose or Fermi gas is calculated by taking the trace in the Fock space basis, which are eigenstates of the number operator N and the Hamiltonian H. Consider the simplest case of non-interacting scalar field

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - \frac{1}{2}m^2\phi^2.$$

Our goal will be to formulate the path integral representation of $\ln Z$ and then to generalize it to correlation functions $\langle\mathcal{O}\rangle = \mathrm{tr}(\rho\mathcal{O})$, where $\mathcal{O} = T(\phi(x_1), \dots, \phi(x_n))$ is an operator made of fields. We would like to make the generalization

$$\langle\Omega|\dots|\Omega\rangle \xrightarrow{T\neq 0} \mathrm{tr}(\rho\dots).$$

The other way is evident, since

$$\mathrm{tr}(\rho\mathcal{O}) = \sum_n e^{-\beta E_n} \langle n|\mathcal{O}|n\rangle \xrightarrow{T=0} \langle\Omega|\mathcal{O}|\Omega\rangle .$$

When the temperature goes to zero, only the ground state contributes to the trace.

3.1.1 Path representation of the partition function

We will use the coordinate (field) eigenstate basis. The action of the Heisenberg field operator on the eigenstate is

$$\hat{\phi}(t_i, \mathbf{x}) |\phi_i, t_i\rangle = \phi_i(\mathbf{x}) |\phi_i, t_i\rangle , \quad \forall\mathbf{x},$$

where $\phi_i(\mathbf{x})$ is the classical field configuration. The path integral representation of the time evolution operator is

$$\begin{aligned}\langle\phi_\mathrm{f}, t_\mathrm{f}|\phi_\mathrm{i}, t_\mathrm{i}\rangle &= \int\limits_{\phi(t_\mathrm{i},\mathbf{x})=\phi_\mathrm{i}(\mathbf{x})}^{\phi(t_\mathrm{f},\mathbf{x})=\phi_\mathrm{f}(\mathbf{x})} \mathcal{D}[\phi(x), \pi(x)] \exp\left(i \int_{t_\mathrm{i}}^{t_\mathrm{f}} \mathrm{d}t \int \mathrm{d}^3\mathbf{x} \left(\pi(x)\dot{\phi}(x) - \mathcal{H}(\phi(x), \pi(x))\right)\right) \\ &= \langle\phi_\mathrm{f}|e^{-iH(t_\mathrm{f}-t_\mathrm{i})}|\phi_\mathrm{i}\rangle , \end{aligned} \tag{$**$}$$

where ϕ_f and ϕ_i are eigenstates of $\phi(t = 0, \mathbf{x})$. We want to compare this to

$$Z = \mathrm{tr}(e^{-\beta(H-\mu N)}) = \int \mathcal{D}[\phi(\mathbf{x})] \langle\phi(\mathbf{x})|e^{-\beta(H-\mu N)}|\phi(\mathbf{x})\rangle .$$

In order to do this, we make the following substitutions in $(**)$:

- Put $\phi_\mathrm{f} = \phi_\mathrm{i} = \phi(\mathbf{x})$ and replace $\phi(\mathbf{x})$.

- Set $t_\mathrm{i} = 0$ and $t_\mathrm{f} = -i\beta$ (this is called the imaginary time formalism).
- Put $H \to H - \mu N$ if $N = N(\phi, \pi)$ and if it is a conserved quantity.

The result is

$$Z = \int\limits_{\substack{\phi(\beta,\mathbf{x})=\phi(0,\mathbf{x}) \\ \text{periodic}}} \mathcal{D}[\phi(x), \pi(x)] \exp\left(\int_0^\beta \mathrm{d}\tau \int \mathrm{d}^3\mathbf{x} \left(i\pi(x) \frac{\partial \phi(x)}{\partial \tau} - \mathcal{H}(\phi, \pi) + \mu N(\phi, \pi) \right) \right).$$

3.1.2 Free, real scalar field

For the free, real scalar field there are no conserved quantities, so $\mu = 0$. First, we will do the path integral over π. This sets

$$\pi = \frac{\partial \mathcal{L}}{\partial(\partial_0 \phi)} = \frac{\partial \phi}{\partial t} = i \frac{\partial \phi}{\partial \tau}$$

and leads to

$$Z = \int\limits_{\phi \text{ periodic on } [0,\beta]} \mathcal{D}[\phi(x)] e^{-S_\mathrm{E}[\phi]},$$

where

$$S_\mathrm{E}[\phi] = \frac{1}{2} \int_0^\beta \mathrm{d}\tau \int \mathrm{d}^3\mathbf{x} \left(\left(\frac{\partial \phi}{\partial \tau} \right)^2 + (\nabla \phi)^2 + m^2 \phi^2 \right).$$

It is the *Euclidean action*, because the Laplacian operator is positive definite. This is an exceptional case, because we can actually calculate Z exactly:

$$Z = \mathcal{N} \frac{1}{\sqrt{\det D}}, \quad D = -\beta^2 \left(\frac{\partial^2}{\partial \tau^2} + \nabla^2 - m^2 \right).$$

The prefactor $\mathcal{N}$ is related to the measure and is not relevant for the following discussion, where we will compute $\ln Z$ (the minus sign in the determinant comes from the integration by parts). We will temporarily use finite volume V and the following discretization of momenta:

$$\sum_{\mathbf{p}} \to V \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}.$$

The Fourier-ansatz for the eigenfunctions of D is

$$\phi(\tau, \mathbf{x}) = \sqrt{\frac{\beta}{V}} \sum_{n=-\infty}^{\infty} \sum_{\mathbf{p}} e^{i(\omega_n \tau + \mathbf{p}\cdot\mathbf{x})} \phi_n(\mathbf{p})$$

with $\omega_n = 2\pi n/\beta = 2\pi n T$. The choice of ω_n comes from the periodic boundary conditions in the imaginary time direction. They are called the (bosonic) *Matsubara frequencies*. The eigenvalues of D are $\beta^2(\omega_n^2 + \omega(\mathbf{p})^2)$. Hence,

$$Z = N \prod_{n=-\infty}^{\infty} \prod_{\mathbf{p}} \frac{1}{\beta \sqrt{\omega_n^2 + \omega(\mathbf{p})^2}} = N \prod_{\mathbf{p}} \frac{1}{\beta \omega(\mathbf{p})} \prod_{n=1}^{\infty} \frac{1}{\beta^2(\omega_n^2 + \omega(\mathbf{p})^2)}.$$

We could directly compute the infinite sum, but instead we will compute logarithm of Z:

$$\begin{aligned}
\ln Z &= -\frac{1}{2}\sum_{\mathbf{p}}\sum_{n=-\infty}^{\infty}\ln(\beta^2(\omega_n^2+\omega(\mathbf{p})^2))+C\\
&= \frac{1}{2}\sum_{\mathbf{p}}\sum_{n=-\infty}^{\infty}\ln((2\pi n)^2+\beta^2\omega(\mathbf{p})^2)+C\\
&= -\frac{1}{2}\sum_{\mathbf{p}}\sum_{n=-\infty}^{\infty}\int_1^{(\beta\omega(\mathbf{p}))^2}\frac{\mathrm{d}z}{z+(2\pi n)^2}+C(\beta\text{-independent})\\
&= -\frac{1}{2}\sum_{\mathbf{p}}\int_1^{(\beta\omega(\mathbf{p}))^2}\mathrm{d}z\sum_n\frac{1}{z+(2\pi n)^2}+C\\
&= -\sum_{\mathbf{p}}\int_1^{\beta\omega(\mathbf{p})}\mathrm{d}t\,\left(\frac{1}{2}+\frac{1}{e^t-1}\right)\\
&= V\int\frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\left(-\frac{1}{2}\beta\omega(\mathbf{p})-\ln(1-e^{-\beta\omega(\mathbf{p})})\right).
\end{aligned}$$

We used the following identity:

$$\sum_{n=-\infty}^{\infty}\frac{1}{z^2+(2\pi n)^2}=\frac{1}{2}\coth\left(\frac{z}{2}\right)=\frac{1}{2z}\left(1+\frac{2}{e^z-1}\right).$$

When we take the derivative of $\ln Z$ with respect to β, we get the familiar formula for the energy density of free boson gas:

$$\frac{E}{V}=-\frac{1}{V}\frac{\partial\ln Z}{\partial\beta}=\int\frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\left(\frac{1}{2}\omega(\mathbf{p})+\frac{\omega(\mathbf{p})}{e^{\beta\omega(\mathbf{p})}-1}\right).$$

The first term is the zero-point energy of the bosonic field. Expanding the equation for relativistic particles, we get

$$\frac{F}{V}=-\frac{\ln Z}{V\beta}=\int\frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\left(\frac{1}{2}\sqrt{m^2+\mathbf{p}^2}+T\ln\left(1-\exp\left(-\beta\sqrt{m^2+\mathbf{p}^2}\right)\right)\right).$$

The first term is quartically divergent and equals $m^4/(16\pi^2\varepsilon)$ in the dimensional regularization with $d=4-2\varepsilon$. This divergence is not observable (in the absence of gravity) and we can simply set it to zero. In the second term, we can recognize a dimensionless quantity $x=|\mathbf{p}|\beta$:

$$\frac{F}{V}=\frac{T^4}{2\pi^2}\int_0^{\infty}\mathrm{d}x\,x^2\ln\left(1-e^{-\sqrt{x^2+m^2/T^2}}\right).$$

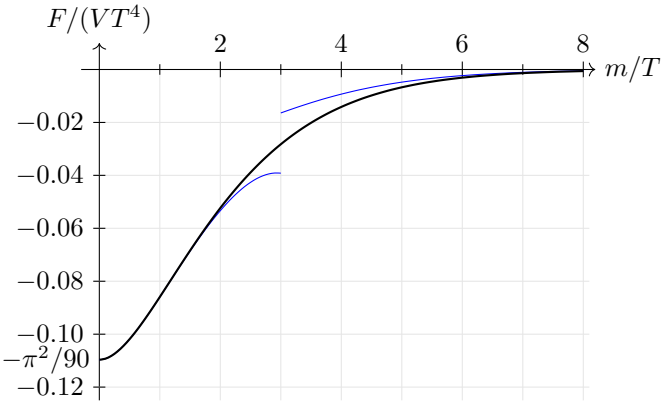
The plot of this function is given below. In the limit $T\to 0$, the free energy goes to zero exponentially fast

$$\frac{E}{V}=-T^4\left(\frac{m}{2\pi T}\right)^{3/2}e^{-m/T}\left(1+\mathcal{O}(T/m,e^{m/T})\right).$$

The approximation for the high-temperature limit $T\gg m$ is

$$\frac{F}{V}=-\frac{\pi^2T^4}{90}+\frac{m^2T^2}{24}-\frac{m^3T}{12\pi}-\frac{m^4}{32\pi^2}\left(\ln\left(\frac{me^{\gamma_{\mathrm{E}}}}{4\pi T}\right)-\frac{3}{4}\right)+\mathcal{O}(m^6/T^6).$$

These non-analytic terms (IR effects) are related exclusively to the zero Matsubara frequency mode $n=0$. Both limits are visible in the plot below (coloured in blue).



3.1.3 Free charged scalar and Bose–Einstein condensation

The Lagrangian for the free, complex scalar field is

$$\mathcal{L} = \partial_\mu \phi^* \partial^\mu \phi - m^2 \phi \phi^*.$$

This Lagrangian has a conserved U(1) current $\phi \to e^{i\alpha}\phi$

$$j^\mu = i(\phi^* \partial_\mu \phi - \phi \partial_\mu \phi^*) = -i(\phi_1 \partial_\mu \phi_2 - \phi_2 \partial_\mu \phi_1)$$

where $\phi = (\phi_1 + i\phi_2)/\sqrt{2}$. In this case, we have to include the chemical potential μ in the partition function. However, the chemical potential comes with the charge operator

$$Q = \int \mathrm{d}^3\mathbf{x}\, j^0(\mathbf{x}) = \int \mathrm{d}^3\mathbf{x}\, (\phi_2 \pi_1 - \phi_1 \pi_2).$$

The partition function is

$$\begin{aligned} Z &= \mathrm{tr}\left(e^{-\beta(H-\mu Q)}\right) \\ &= \int\limits_{\substack{\phi_1,\phi_2 \\ \text{periodic}}} \mathcal{D}[\phi_1, \phi_2, \pi_1, \pi_2] \exp\left(\int_0^\beta \mathrm{d}\tau \int \mathrm{d}^3\mathbf{x} \left[i\pi_1 \frac{\partial \phi_1}{\partial \tau} + i\pi_2 \frac{\partial \phi_2}{\partial \tau} - \mathcal{H}(\phi_i, \pi_i) + \mu(\phi_2 \pi_1 - \phi_1 \pi_2) \right] \right). \end{aligned}$$

Next, we collect the quadratic terms in π_i and integrate them out. The result is

$$\begin{aligned} Z = \int\limits_{\substack{\phi_1,\phi_2 \\ \text{periodic}}} \mathcal{D}[\phi_1, \phi_2] \exp\Bigg(-\frac{1}{2} \int_0^\beta \mathrm{d}\tau \int \mathrm{d}^3\mathbf{x} \Bigg[\left(\frac{\partial \phi_1}{\partial \tau} - i\mu\phi_2 \right)^2 + \left(\frac{\partial \phi_2}{\partial \tau} + i\mu\phi_1 \right)^2 \\ + \nabla^2\phi_1 + \nabla^2\phi_2 + m^2(\phi_1^2 + \phi_2^2) \Bigg] \Bigg). \end{aligned}$$

The same as before, we define the Euclidean action

$$\begin{aligned} S_\mathrm{E} &= \int_0^\beta \mathrm{d}\tau \int \mathrm{d}^3\mathbf{x} \left(\partial_\tau \phi^* \partial_\tau \phi + \nabla\phi^* \cdot \nabla\phi + m^2 \phi^* \phi + \mu j^0 - \mu^2 \phi^* \phi \right) \\ &= \int_0^\beta \mathrm{d}\tau \int \mathrm{d}^3\mathbf{x} \left((\partial_\tau + \mu)\phi^* (\partial_\tau - \mu)\phi + \nabla\phi^* \cdot \nabla\phi + m^2 \phi^* \phi \right). \end{aligned}$$

We can see that the effective potential effectively shifts the Matsubara frequencies:

$$\omega_n \to \omega_n + i\mu.$$

The Fourier expansion of the eigenfunctions is

$$\phi = ve^{i\theta} + \sqrt{\frac{\beta}{V}} \sum_{n=-\infty}^{\infty} \sum_{\mathbf{p}} e^{i(\omega_n \tau + \mathbf{p}\cdot\mathbf{x})} \phi_n(\mathbf{p}),$$

where we have allowed for a possible condensate v in the zero mode. Plugging this into the action gives

$$S_{\mathrm{E}} = \beta V v^2(m^2 - \mu^2) + \frac{1}{2} \sum_n \sum_{\mathbf{p}} (\phi_{1,-n}(-\mathbf{p}), \phi_{2,-n}(-\mathbf{p})) D \begin{pmatrix} \phi_{1,n}(\mathbf{p}) \\ \phi_{2,n}(\mathbf{p}) \end{pmatrix},$$

where

$$D = \beta^2 \begin{pmatrix} \omega_n^2 + \omega(\mathbf{p})^2 - \mu^2 & -2\mu\omega_n \\ 2\mu\omega_n & \omega_n^2 + \omega(\mathbf{p})^2 - \mu^2 \end{pmatrix}.$$

The logarithm of the partition function is

$$\begin{aligned} \ln Z &= -\beta V v^2(m^2 - \mu^2) - \frac{1}{2} \sum_{\pm} \sum_{\mathbf{p}} \sum_n \ln\left(\beta^2(\omega_n^2 + (\omega(\mathbf{p}) \pm \mu)^2)\right) \\ &= -\beta V v^2(m^2 - \mu^2) - V \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \left[\beta\omega(\mathbf{p}) + \ln(1 - e^{-\beta(\omega(\mathbf{p})-\mu)}) + \ln(1 - e^{-\beta(\omega(\mathbf{p})+\mu)})\right]. \end{aligned}$$

Evidently, there are two particles with the same energies but with opposite chemical potentials.

The ground state is the extremum of F with respect to all parameters, including v:

$$0 = \frac{\partial \ln Z}{\partial v} = -2\beta V(m^2 - \mu^2)v.$$

This means that for $m^2 \neq \mu^2$ it must be $v = 0$. In the case $m^2 = \mu^2$, v is still undetermined. On the other hand, we know $|\mu| \leq m$ (otherwise F would be singular). But what happens in the case $|\mu| = m$? The charge density is

$$\rho = j^0 = \frac{Q}{V} = -\frac{1}{V}\frac{\partial}{\partial\mu}\left(\frac{1}{\beta}\ln Z\right)_{\beta,V} = 2mv^2 + \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\left(\frac{1}{e^{\beta(\omega(\mathbf{p})-m)} - 1} - \frac{1}{e^{\beta(\omega(\mathbf{p})+m)} - 1}\right).$$

The first term is the ground state and the second term is the contribution of excited states ($\rho_{\text{excitations}}$). As we lower the temperature, $|\mu|$ grows so that it keeps the second term constant and $v = 0$. When β increases so much that $|\mu|$ reaches m (at β_{crit}), the $\rho_{\text{excitations}}$ decreases and the vacuum expectation value v starts to grow

$$v^2 = \frac{\rho - \rho_{\text{excitations}}}{2m}.$$

The temperature β_{crit} is determined from the second term of the previous equation

$$\rho = \frac{1}{2\pi^2}\int_0^\infty \mathrm{d}p\, p^2 \frac{\sinh(\beta m)}{\cosh(\beta\sqrt{m^2 + p^2}) - \cosh(\beta m)}\bigg|_{\beta=\beta_{\text{crit}}}.$$

This is not analytically invertible, but we can look at some limits. For $\beta m \gg 1$ we get

$$\rho = \left(\frac{m}{2\pi\beta_{\text{crit}}}\right)^{3/2} \zeta(3/2) \quad \Longrightarrow \quad T_{\text{crit}} = \frac{2\pi}{m}\left(\frac{\rho}{\zeta(3/2)}\right)^{2/3}.$$

For $\beta m \ll 1$ it holds

$$T_{\mathrm{cr}} = \left(\frac{3\rho}{m}\right)^{1/2}.$$

From this it seems that the critical temperature can be arbitrarily high, however, for very small mass it is very difficult to produce net charge, as both particles and anti-particles are excited.

3.1.4 Kubo–Martin–Schwinger relation

The two-point Green's function at finite temperature for a scalar field is defined as (*Kubo–Martin–Schwinger relation*)

$$G(\tau, 0; \mathbf{x}, \mathbf{y}) = \frac{1}{Z} \operatorname{tr}\left(e^{-\beta H} T\{\phi(\tau, \mathbf{x}), \phi(0, \mathbf{y})\}\right) = G(\tau, \beta; \mathbf{x}, \mathbf{y}).$$

The two-point function is β-periodic.

3.1.5 Fermions

Dirac's fermions are described by the Lagrangian

$$\mathcal{L} = \overline{\Psi}(i\not{\partial} - m)\Psi$$

and the corresponding Hamiltonian is

$$\mathcal{H} = \Pi\dot{\Psi} - \mathcal{L} = \overline{\Psi}(-i\gamma^i\nabla^i + m)\Psi.$$

The conserved current $j^\mu = \overline{\Psi}\gamma^\mu\Psi$ gives the conserved charge Q. The partition function is

$$Z = \int\limits_{\Psi(\beta,\mathbf{x})=-\Psi(0,\mathbf{x})} \mathcal{D}[\Psi, \overline{\Psi}] \exp\left(\int_0^\beta \mathrm{d}\tau \int \mathrm{d}^3\mathbf{x} \left(i\Pi\frac{\partial\Psi}{\partial\tau} - \mathcal{H} + \mu j^0\right)\right).$$

The minus sign in the boundary condition comes from the anticommuting nature of fermions (anti-periodicity), which comes from the KMS relation for the fermions. The Euclidean action becomes

$$S_{\mathrm{E}} = \int_0^\beta \mathrm{d}\tau \int \mathrm{d}^3\mathbf{x}\, \overline{\Psi}\left(\gamma^0\frac{\partial}{\partial\tau} - i\boldsymbol{\gamma}\cdot\nabla + m - \mu j^0\right)\Psi.$$

The factor of i in the spatial part is bothering us, so we redefine the gamma matrices as

$$\tilde{\gamma}_0 = \gamma_0, \quad \tilde{\gamma}_i = i\gamma_i.$$

With this, the slash notation becomes $\not{\partial} = \tilde{\gamma}_0\partial_\tau + \tilde{\gamma}_i\nabla^i$ and $\{\tilde{\gamma}_\mu, \tilde{\gamma}_\nu\} = 2\delta_{\mu\nu}$. The Fourier mode expansion is

$$\Psi_\alpha(\tau, \mathbf{x}) = \frac{1}{\sqrt{V}} \sum_{n=-\infty}^{\infty} \sum_{\mathbf{p}} e^{i(\omega(\mathbf{p})\tau + \mathbf{p}\cdot\mathbf{x})} \Psi_{\alpha;n}(\mathbf{p}).$$

There is no factor of $\sqrt{\beta}$ since the dimension of the field is already $[\Psi] = 3/2$. The Matsubara frequencies for fermions are

$$\omega_n = (2n+1)\pi T.$$

As we can see from anti-periodicity, there is no zero mode, which gave rise to nonanalyticities in the bosonic case. A more compact form of the Euclidean potential is

$$S_{\mathrm{E}} = \sum_n \sum_{\mathbf{p}} \overline{\Psi}_{\alpha;n}(\mathbf{p}) \mathrm{D}_{\alpha\beta} \Psi_{\beta;n}(\mathbf{p}),$$

where

$$D = \beta\gamma^0((i\omega_n - \mu) + \gamma^0 \boldsymbol{\gamma}\cdot\mathbf{p} + \gamma^0 m) = \beta(i\not{p} + m).$$

The partition function will again depend on the determinant of this operator, but this time the determinant will be in the numerator (due to the Grassmann nature of the variables) and it will not be under the square root as we have two fields to integrate: $Z = \mathcal{N} \det D$. The logarithm of the partition function is

$$\ln Z = \ln \prod_{\mathbf{p}} \prod_n \det(D_{n,\mathbf{p}}) + \text{const.}$$

Note that the product is the same if one lets $n \to -n$, $\mathbf{p} \to -\mathbf{p}$ and $\mu \to -\mu$. Then $\mathrm{D} \to \beta(-i\not{p} + m)$. Therefore,

$$\begin{aligned}
\ln Z &= \ln \prod_n \prod_{\mathbf{p}} \sqrt{\det \beta^2 (i\not{p} + m)(-i\not{p} + m)} \\
&\ln \prod_n \prod_{\mathbf{p}} \sqrt{\det \beta^2 (p^2 + m^2) \mathbb{1}_{4\times 4}} \\
&= \frac{1}{2} \cdot 4 \cdot \sum_n \sum_{\mathbf{p}} \ln(\beta^2(p^2 + m^2)) \\
&= 2 \sum_n \sum_{\mathbf{p}} \sum_{\pm} \ln(\beta^2(\omega_n^2 + (\omega(\mathbf{p}) \pm \mu)^2)) \\
&= \sum_{\mathbf{p}} \sum_n \sum_{\pm} \int_1^{\beta^2(\omega(\mathbf{p})\pm\mu)^2} \frac{\mathrm{d}z}{z + (2n+1)^2\pi^2} \\
&= 2V \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \left[2 \cdot \frac{1}{2}\beta\omega(\mathbf{p}) + \ln(1 + e^{-\beta(\omega(\mathbf{p})-\mu)}) + \ln(1 + e^{-\beta(\omega(\mathbf{p})+\mu)}) \right].
\end{aligned}$$

The factor 2 in front indicates two spin states of the fermion, and again we have two particles with opposite chemical potentials. Notice the difference of $-$ sign in the logarithms compared to the bosonic case. The difference is better evident if we look at the energy density

$$\frac{E}{V} = -\frac{\partial \ln Z}{\partial \beta} + \mu Q = 2 \sum_{\pm} \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \left(-\frac{1}{2}\omega(\mathbf{p}) + \frac{\omega(\mathbf{p})}{e^{\beta(\omega(\mathbf{p})\pm\mu)} + 1} \right).$$

Note the negative zero-point energy.

3.2 Interactions

3.2.1 Propagators, Feynman rules and renormalization

With the imaginary time prescription, we moved from $\mathbb{R}^4$ (Minkowski space) to $S^1 \times \mathbb{R}^3$ (4-dimensional Euclidean cylinder). We will split the Euclidean action into the free and interaction parts $S_{\mathrm{E}} = S_0 + S_{\mathrm{int}}$. We want to construct a perturbative expansion in S_{int}. First, we will

study the zero-point correlation functions (the vacuum diagrams). For $T = 0$, we defined another function $W[\mathcal{J}]$ with $Z[\mathcal{J}] = e^{iW[\mathcal{J}]}$ that generated only connected diagrams. At $T \neq 0$ we are interested in T-dependence of $\ln Z$. The contributions only come from connected vacuum diagrams. Consider the interaction term $\lambda\phi^4/4$. The Euclidean action is

$$S_{\mathrm{E}} = \int_0^\beta \mathrm{d}\tau \int \mathrm{d}^3\mathbf{x} \left(\frac{1}{2}(\partial_\tau \phi)^2 + \frac{1}{2}((\nabla\phi)^2 + m^2\phi^2) + \frac{\lambda}{4}\phi^4 \right) = -\frac{1}{2}\int_x \int_y \phi(x) D(x,y) \phi(y) + S_{\mathrm{int}}$$

where

$$\int_x = \int_0^\beta \mathrm{d}\tau \int \mathrm{d}^3\mathbf{x}, \quad D(x,y) = \delta^{(3)}(x-y)(-\partial_\tau^2 - \nabla^2 + m^2).$$

The Fourier transformation of the two point function is

$$\Delta(x,y) = \sum\!\!\!\!\!\!\int e^{ip(x-y)} \frac{1}{p^2 + m^2}$$

where

$$\sum\!\!\!\!\!\!\int = \frac{1}{\beta} \sum_n \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}.$$

Writing the propagator out explicitly, we get

$$\Delta(\tau_1, \tau_2; \mathbf{x}, \mathbf{y}) = \frac{1}{\beta} \sum_n \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} e^{i(\omega_n(\tau_1 - \tau_2) + \mathbf{p}\cdot(\mathbf{x}-\mathbf{y}))} \Delta(\omega_n, \mathbf{p})$$

where we introduced the Matsubara propagator

$$\Delta(\omega_n, \mathbf{p}) = \frac{1}{\omega_n^2 + \mathbf{p}^2 + m^2}.$$

As we can see, there are no poles in the Euclidean space (except $\mathbf{p} = 0$ and $m = 0$). The Feynman rules are similar to the zero-temperature case, with the following differences:

- external momenta for "particles" $(\omega_{n,i}, \mathbf{p}_i)$,
- internal lines with momentum $(\omega_{n,i}, \mathbf{k}_i)$:

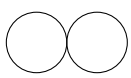
$$\frac{1}{\beta} \sum_{n_i} \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3} \Delta(\omega_{n,i}, \mathbf{k}_i),$$

- momentum conservation at vertices:

$$\beta \cdot \delta \left(\sum_k \omega_{n,k} \right) \cdot \delta^{(3)} \left(\sum_k \mathbf{p}_k \right), \quad \text{all ingoing},$$

- vertex $-\lambda/4$,
- for connected diagrams, the number of loops is $L = I - V + 1$,
- overall momentum conservation $\beta(2\pi)^3\delta^{(3)}(0) \to \beta V$ (the volume of space-time).

An example for ϕ^4 theory is

$$\ln Z = \ln Z_{\text{free}} + 3 \cdot$$  $$+ \left[36 \cdot \right.$$  $$\left. + 12 \cdot \right.$$  $$\left. \right] + \mathcal{O}(\lambda^3).$$

The renormalization at finite temperature works the same way as at zero temperature. The counter-terms are not temperature dependent. Intuitively, this should hold, because the counter-terms arise from the UV divergences, where the momentum is greater than any other scale, including the temperature. Another way to see this is to notice that the temperature-dependent part of the space-time measure is only at the upper bound of time. The divergences arise when all four components are small and that region is not affected by temperature. In dimensional regularization, we choose

$$\frac{1}{V}\sum_{\mathbf{p}} \to \tilde{\mu}^{2\varepsilon} \int \frac{\mathrm{d}^d \mathbf{p}}{(2\pi)^d}, \quad d = 3 - 2\varepsilon.$$

The renormalization conditions are the same as at $T = 0$:

$$\phi_0 = \sqrt{Z}\phi, \quad m_0^2 = m^2 + \delta m^2, \quad \lambda_0 = Z_\lambda \lambda \tilde{\mu}^{2\varepsilon}.$$

We are working in $\overline{\text{MS}}$-scheme

$$\tilde{\mu} = \mu \left(\frac{e^{\gamma_{\text{E}}}}{4\pi} \right)^{1/2}.$$

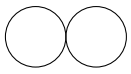
The expanded Lagrangian is

$$\mathcal{L}_{\text{E}} = \frac{1}{2}(\partial_\mu \phi \partial^\mu \phi + m^2 \phi^2) + \frac{\lambda}{4}\phi^4 + \frac{1}{2}(Z-1)\partial_\mu \phi \partial^\mu \phi + \frac{1}{2}((Z-1)m^2 + \delta m^2 Z)\phi^2 + (Z_\lambda Z^2 - 1)\frac{\lambda}{4}\tilde{\mu}^{2\varepsilon}\phi^4.$$

In ϕ^4 theory, the first corrections to Z arise at $\mathcal{O}(\lambda^2)$.

3.2.2 $\mathcal{O}(\lambda)$ correction to $\ln(Z)$ and the two-point function

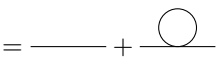
The $\mathcal{O}(\lambda)$ parts of $\ln Z$ are

$$3 \cdot$$  $$- \frac{1}{2}\delta m^2 \cdot$$  $$= 3\left(-\frac{\lambda}{4}\right)\beta V I^2 - \delta m^2 \beta V I,$$

where

$$I = \frac{1}{\beta}\sum_n \int \frac{\mathrm{d}^d \mathbf{p}}{(2\pi)^d} \Delta(\omega_n, \mathbf{p})$$

Note that the field renormalization is not needed at $\mathcal{O}(\lambda)$, as it only starts contributing at λ^2. The two-point function is

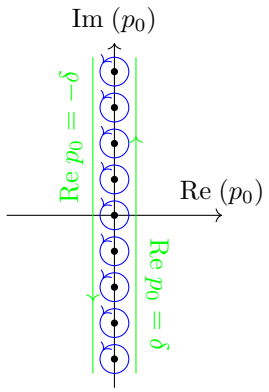
$$G(\omega_n, \mathbf{p}) =$$  $$= \Delta(\omega_n, \mathbf{p}) - \frac{\lambda}{4}\Delta(\omega_n, \mathbf{p})^2 \cdot 12 I - 2 \cdot \frac{1}{2}\delta m^2 \Delta_n(\omega_n, \mathbf{p})^2.$$

The next goal is to calculate I:

$$I = T \int \frac{\mathrm{d}^d \mathbf{p}}{(2\pi)^d} \sum_n \frac{1}{\omega_n^2 + \omega(\mathbf{p})^2}.$$

What happens to be generally useful is a solution to a more general problem, which turns out to have a nice solution

$$T\sum_{n=-\infty}^{\infty} f(p_0) = \frac{T}{2\pi i}\oint_C \mathrm{d}p_0\, f(p_0)\frac{\beta}{2}\coth\left(\frac{\beta p_0}{2}\right)$$
$$= \frac{1}{2\pi i}\int_{-i\infty+\delta}^{i\infty+\delta} \mathrm{d}p^0\ \left(f(p^0)+f(-p^0)\right)\left[\frac{1}{2}+\frac{1}{e^{\beta p_0}-1}\right].$$



The function $\coth(\beta/2)$ is not unexpected, since we have already seen it in a similar context before. The contour C encircles all poles of $\coth(\beta p_0/2)$, which are at $p_0 = i\omega_n$. This contour can be deformed into two parallel lines on either side of the imaginary axis. Of course, we have to be careful that the function is decreasing fast enough for big absolute values of p_0. For more details, see homework.

In our calculation, we take $f(p_0) = 1/(\omega(\mathbf{p})^2 - p_0^2)$, and we split $I = I_{T=0} + I_{T\neq 0}$. For the first part (called the vacuum part), we substitute $p_0 = -ip_4$ and get a result in $(d+1)$-dimensional Euclidean space

$$I_{T=0} = \int \frac{\mathrm{d}^{d+1}p}{(2\pi)^{d+1}}\frac{1}{p^2+m^2}$$

which is exactly the Euclidean equivalent of the self-energy. For $T \neq 0$ we close the contour in the right half-plane and pick up the residue at $p_0 = \omega(\mathbf{p})$. Here, the "four-momenta" are on-shell, so $p_0 = \omega(\mathbf{p}) = \sqrt{m^2+\mathbf{p}^2}$:

$$I_{T\neq 0} = \int \frac{\mathrm{d}^{d}\mathbf{p}}{(2\pi)^{d}}\frac{1}{\omega(\mathbf{p})}\frac{1}{e^{\beta\omega(\mathbf{p})}-1}$$

and the full expression for I is

$$I = \int \frac{\mathrm{d}^{d+1}p}{(2\pi)^{d+1}}\left(\frac{i}{p^2-m^2+i\varepsilon}+2\pi\frac{\delta(p^2-m^2)}{e^{\beta|p_0|}-1}\right),$$

where $p = (p_0, \mathbf{p})$ in the Minkowski space-time. The vacuum is quadratically divergent and determines the mass counter-term

$$\delta m^2 = \frac{\lambda m^2}{16\pi^2\varepsilon} + \text{finite}.$$

Here, choose the finite part such that $\Pi(p)|_{p^2=m^2} = 0$ (the on-shell renormalization condition). The two point function can be to the first order expanded as

$$G = \Delta - \Delta\Pi\Delta,$$

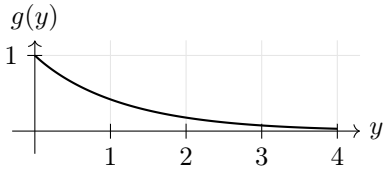
where $\Pi(p) = 3\lambda I + \delta m^2$. Together with the renormalization condition, we get $\delta m^2 = -3\lambda I_{T=0}$ and

$$\Pi(p) = 3\lambda\int\frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{\omega(\mathbf{p})}\frac{1}{e^{\beta\omega(\mathbf{p})}-1}.$$

To evaluate the integral, we define two dimensionless quantities $y = m/T$ and $x = |\mathbf{p}|/m$:

$$\Pi(p) = \frac{\lambda T^2}{4}\underbrace{\frac{6y^2}{\pi^2}\int_0^\infty \mathrm{d}x\,\frac{x^2}{\sqrt{1+x^2}}\frac{1}{e^{y\sqrt{1+x^2}}-1}}_{g(y)}.$$

We can plot the function $g(y)$ and look at its asymptotic behaviour.



For $T \to 0$ we get $\Pi \to 0$ and for $T \gg m$ we get $\Pi \propto \lambda T^2$. The two point function can be resummed to give the *thermal mass*

$$G(\omega_n, \mathbf{p}) = \frac{1}{\omega_n^2 + \mathbf{p}^2 + m^2 + \Pi(\omega_n, \mathbf{p})}.$$

The thermal mass is obtained when we evaluate the self-energy at zero mass ($m \to 0$):

$$m_{\mathrm{T}} = \sqrt{\lambda}\frac{T}{2}.$$

Let's return to the partition function

$$\begin{aligned}
\frac{1}{\beta V} \ln Z|_{\mathcal{O}(\lambda)} &= 3 \cdot \quad - \frac{1}{2}\delta m^2 \cdot \\
&= -\frac{3\lambda}{4} I^2 - \frac{1}{2}\delta m^2 I \\
&= -\frac{3}{4}\lambda \left((I_{T=0} + I_{T\neq 0})^2 - 2 I_{T=0}(I_{T=0} + I_{T\neq 0}) \right) \\
&= -\frac{3}{4}\lambda \left(I_{T\neq 0}^2 - I_{T=0}^2 \right).
\end{aligned}$$

For example, the pressure would be calculated as

$$\begin{aligned}
p = -\left.\frac{\partial F}{\partial V}\right|_{T,N} &= -\frac{F}{V} = \frac{1}{\beta V} \ln Z \\
&= -\left.\frac{F}{V}\right|_{T\text{-dep}}^{\mathcal{O}(\lambda^0)} - \frac{3}{4}\lambda \left(\int \frac{\mathrm{d}^4\mathbf{p}}{(2\pi)^3} \frac{1}{\omega(\mathbf{p})} \frac{1}{e^{\beta\omega(\mathbf{p})} - 1} \right)^2 + \dots \\
&\overset{m=0}{=} \frac{\pi^2 T^4}{90} - \frac{\lambda T^4}{192} + \dots
\end{aligned}$$

The pressure decreases with the interaction term.

3.3 Infrared divergences and resummation

Let's focus on $\mathcal{O}(\lambda^2)$ contributions to $\ln Z$ in ϕ^4 theory. There are two diagrams:

$$\ln Z|_{\mathcal{O}(\lambda^2)} = \left[36 \cdot \quad + 12 \cdot \right].$$

The Euclidean propagator in Euclidean space is

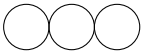
$$\frac{1}{\omega_n^2 + \mathbf{p}^2 + m^2}.$$

This can diverge only for $m = 0$ and the zero Matsubara mode $n = 0$. Let's look at the second diagram:



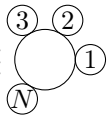
$$\sim \int \mathrm{d}^3\mathbf{p}_1\, \mathrm{d}^3\mathbf{p}_2\, \mathrm{d}^3\mathbf{p}_3\, \frac{1}{\mathbf{p}_1^2\mathbf{p}_2^2\mathbf{p}_3^2(\mathbf{p_1}+\mathbf{p}_2+\mathbf{p}_3)^2}.$$

This integral is not IR divergent. However, the first diagram is



$$\sim \int \frac{\mathrm{d}^3\mathbf{p}_1}{\mathbf{p}_1^2} \int \frac{\mathrm{d}^3\mathbf{p}_2}{\mathbf{p}_2^2} \int \frac{\mathrm{d}^3\mathbf{p}_3}{\mathbf{p}_3^4}.$$

This does diverge in the IR limit, and, moreover, any *ring diagram* will diverge:



$$\propto [-\Pi_{\text{1-loop}}]^N \frac{\mathrm{d}^3\mathbf{p}}{(\mathbf{p}^2)^N}.$$

The full expression for all ring diagrams is

$$\begin{aligned}
&\beta V \frac{1}{\beta} \sum_n \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \sum_N \frac{1}{N!} \frac{(N-1)!}{2} \left[-\Pi_{\text{1-loop}}(\omega_n, \mathbf{p})\Delta(\omega_n, \mathbf{p})\right]^N \\
&= -\frac{V}{2} \sum_n \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \left[\ln\left(1 + \Pi(\omega_n, \mathbf{p})\Delta(\omega_n, \mathbf{p})\right) - \Pi(\omega_n, \mathbf{p})\Delta(\omega_n, \mathbf{p})\right].
\end{aligned}$$

The preceding βV comes from the momentum conservation and the combinatoric factor comes from four parts: $1/N!$ from N-th order in perturbation theory, $N!$ from arranging N loops, $1/N$ for the loops being cyclically symmetric and $1/2$ for the circle being unordered. We have used the Taylor expansion of the logarithm

$$\ln(x+1) = -\sum_{n=1}^{\infty} \frac{(-x)^n}{n}.$$

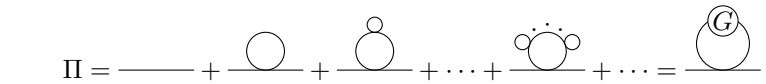
In the limit $n = 0$ and $\mathbf{p}, m \to 0$, we have

$$\Pi_{\text{1-loop}}\Delta(\omega_0, \mathbf{p} = 0) = \frac{\lambda T^2}{4} \cdot \frac{1}{m^2}.$$

Hence, we need the resummation when $m^2 < \lambda T^2/4$. The leading contribution in this case is $(m = 0)$

$$\frac{V}{2} \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \left[\ln\left(1 + \frac{\lambda T^2}{4} \frac{1}{\omega_0^2 + \mathbf{p}^2}\right) - \frac{\lambda T^4}{4} \frac{1}{\omega_0^2 + \mathbf{p}^2}\right] = \beta V \frac{\lambda^{3/2} T^4}{96\pi}$$

which is IR finite. However, the result is non-analytic in λ and also non-perturbative. The origin of the non-analyticity is the zero mode of the bosonic field, and it comes from the power counting and the relation $\mathbf{p} \sim \sqrt{\lambda} T$. The self-energy also includes ring diagrams:

$\Pi =$ 

$$= 3\lambda \sum_n \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \frac{1}{\omega_n^2 + \mathbf{p}^2 + m^2 + \Pi(\omega_n, \mathbf{p})}.$$

We have used that

$$G(\omega_n, \mathbf{p}) = \frac{1}{\omega_n^2 + \mathbf{p}^2 + m^2 + \Pi(\omega_n, \mathbf{p})}.$$

If $n \neq 0$, then we can expand in λ and there is no IR divergence, but for $n = 0$ (and $m = 0$) the cut-off regularized integral gives

$$\int_\Lambda \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \frac{1}{\mathbf{p}^2 + \lambda T^2/4} = \Lambda - \frac{\sqrt{\lambda}}{4}\pi T + \mathcal{O}(\Lambda^{-1}).$$

The T- dependent part of Π at $\lambda^{3/2}$ is

$$\Pi|^{T\text{-dep}}_{\mathcal{O}(\lambda^{3/2})} = -\frac{3\lambda^{3/2}T^2}{8\pi}, \quad \Longrightarrow \quad \Pi = \frac{\lambda}{4}T^2\left(1 - \frac{3\sqrt{\lambda}}{2\pi}\right).$$

3.3.1 Effective field theory for the zero mode

At high temperatures, we can construct an effective field theory (EFT) for the zero mode of the bosonic field. The radius of the Euclidean time circle is small, we effectively lose one dimension. Assume $|\mathbf{p}| \ll T$, and we split the fields for $n = 0$ and $n \neq 0$ (which we integrate out):

$$\phi(\tau, x) = \sqrt{\frac{\beta}{V}} \sum_n \sum_{\mathbf{p}} e^{i(\omega_n \tau + \mathbf{p}\cdot\mathbf{x})} \phi_n(\mathbf{p}).$$

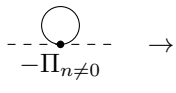
The zero mode contributes to EFT the field φ

$$\varphi(\mathbf{x}) = \sqrt{\frac{\beta}{V}} \sum_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x}} \varphi(\mathbf{p}).$$

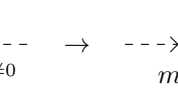
We will use the dashed propagator for $n = 0$ mode and the solid one for $n \neq 0$. Assuming $\mathbb{Z}_2$ symmetry of the ϕ^4 theory, the effective action reads

$$S_{\text{eff}} = \beta \int \mathrm{d}^d\mathbf{x} \left(\frac{1}{2}(\nabla\varphi)^2 + \frac{1}{2}m_{\mathrm{T}}^2\varphi^2 + \frac{K_4}{4}\varphi^4 + \frac{K_6}{6}\varphi^6 + \dots\right)$$

The assumption for EFT is that the mass is smaller than $m^2 \leq \lambda T^2$ and that all scales are much smaller than T (but can be around $\sim \sqrt{\lambda}T$). To calculate the EFT coefficients, we do the matching at $m = 0$:

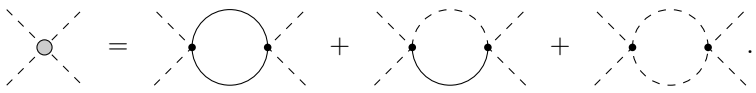


$\rightarrow$



$$, \quad \Longrightarrow \quad m_{\mathrm{T}}^2 = \frac{\lambda}{4}T^2.$$

For K_4, for example, we have



The six-point vertex K_6 arises as



3.4 Fermions, sums and loops

The Euclidean action for the free Dirac fermion is

$$S_{\mathrm{E}}^{\mathrm{free}} = \beta \sum_n \sum_{\mathbf{p}} \overline{\Psi}(i\not{p} + m)\Psi,$$

where

$$\not{p} = p_0\tilde{\gamma}_0 + \mathbf{p}\cdot\tilde{\boldsymbol{\gamma}}, \quad p_0 = \omega_n + i\mu$$

and the propagator is expressed as

$$S(\omega_n, \mathbf{p}) = (i\not{p} + m)^{-1} = \frac{-i\not{p} + m}{p^2 + m^2}.$$

The integral representation for the thermal sums ($\mu = 0$) is

$$\Sigma_{\mathrm{f}} f(T) = T \sum_{n=-\infty}^{\infty} f(\omega_n = (2\pi + 1)\pi T).$$

If we compare it to the bosonic sum

$$\Sigma_{\mathrm{b}} f(T) = T \sum_{n=-\infty}^{\infty} f(\omega_n = 2\pi n T)$$

we find

$$\Sigma_{\mathrm{f}} f(T) = T \sum_n (f(n\pi T) - f(2\pi n T)) = 2\Sigma_b(f(T/2)) - \Sigma_{\mathrm{b}}(T).$$

From here, we can continue with the bosonic case

$$\Sigma_{\mathrm{b}} f(T) = \frac{1}{2\pi i} \int_{-i\infty+\delta}^{i\infty+\delta} \mathrm{d}p_0 \, (f(p_0) + f(-p_0)) \left[\frac{1}{2} + \frac{1}{e^{\beta p_0} - 1}\right],$$

whence it follows

$$\Sigma_{\mathrm{f}} f(T) = \frac{1}{2\pi} \int_{-\infty-i\delta}^{\infty-i\delta} \mathrm{d}k_0 \, \left[\frac{1}{2} - \frac{1}{e^{i\beta k_0} + 1}\right] (f(k_0) + f(-k_0)).$$

If $\mu \neq 0$, we substitute $k_0 \to k_0 + i\mu$. We always have to be careful with the minus signs from the loops of fermions. There is also no IR divergences, because there are no zero Matsubara modes.

3.5 QCD at finite temperature

3.5.1 Path integral representation

The starting point is the QCD Lagrangian

$$\mathcal{L}_{\mathrm{QCD}} = -\frac{1}{4} G^A_{\mu\nu} G^{A,\mu\nu} + \sum_{\mathrm{f}} \overline{\Psi}_{\mathrm{f}}(i\not{D} - m_{\mathrm{f}})\Psi_{\mathrm{f}} + \mathcal{L}_{\text{gauge-fix}} + \mathcal{L}_{\text{ghost}}.$$

The construction of the path integral at finite temperature follows the same steps as for $T = 0$. First we identify the canonical variables, $\pi^A_\mu = G^A_{\mu 0}$. We choose $\pi^A_0 = 0$, so that A^A_0 is not a

canonical variable. To fix the gauge we could use $A_0^A = 0$ together with Gauss constraint, but we rather work with $A_3^A = 0$, which only leaves A_i^A for $i = 1, 2$. Then we write the Hamiltonian path integral, where we require the periodic boundary condition for A_i^A:

$$Z = \int\limits_{A_i^A(\beta,\mathbf{x})=A_i^A(0,\mathbf{x})} \mathcal{D}[A_i^A, \Pi_i^A] \exp\left(\int_0^\beta \mathrm{d}\tau \int \mathrm{d}^3\mathbf{x}\, \left((-i)\Pi^{A,i}(\mathbf{x})A^{A,i}(\mathbf{x}) - \mathcal{H}(A,\Pi)\right)\right),$$

where

$$G_{0i}^A \equiv \partial_\tau A_i^A - \partial_i A_0^A + g f^{ABC} A_0^B A_i^C$$

and G_{ij}^A is defined as usual. Then we reintroduce A_0^A that was eliminated by the gauge fixing by adding it as an extra integration. After performing the Gaussian integral over Π_i^A, going to covariant gauge $f^A(x) = \partial^\mu A_\mu^A(x)$ and representing the Faddeev-Popov determinant as a path integral over ghost fields, we get the partition function for QCD at finite temperature:

$$Z = \mathcal{N} \int\limits_{\text{periodic}} \mathcal{D}[A_0^A, A_i^A] \int\limits_{\text{periodic}} \mathcal{D}[\overline{c}^A, c^A] \int\limits_{\text{antiperiodic}} \mathcal{D}[\Psi, \overline{\Psi}] \exp\left(-\int_0^\beta \mathrm{d}\tau \int \mathrm{d}^3\mathbf{x}\, \mathcal{L}_\mathrm{E}\right)$$

where

$$\mathcal{L}_\mathrm{E} = \frac{1}{4} G_{\mu\nu}^A G_{\mu\nu}^A + \frac{1}{2\xi} f^A f^A + \overline{c}^A \frac{\delta f^A}{\delta \varepsilon^B} c^B + \overline{\Psi}(\gamma_\mu \mathrm{D}_\mu + m)\Psi.$$

Remark.

(1) Gauge fields and ghosts are periodic on $(0, \beta)$. The field A_3^A (in the axial gauge) is periodic from the definition of the path integral $Z = \operatorname{tr} e^{\beta H}$ (it needed to be periodic even before choosing a gauge). Ghosts have to be periodic because they are introduced as a representation of the Faddeev-Popov determinant, which is a functional determinant of a bosonic (periodic) operator.

(2) When we shift from Minkowski to Euclidean space, we perform the substitution

$$\mathcal{L}_\mathrm{E} = -\mathcal{L}_\mathrm{M}(t \to -i\tau, A_0^A \to iA_0^A, \mathrm{D}_0 = \partial_t - igA_0^A \to i\mathrm{D}_\tau = i\partial_\tau + gA_0^A).$$

(3) For the interaction part of the Lagrangian, the substitution is

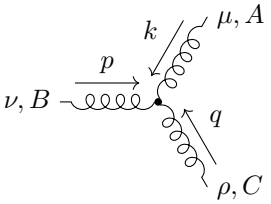
$$-\mathcal{L}_\mathrm{M}|_{t\to -i\tau} = -\overline{\Psi}(i\gamma^0\partial_t - i\boldsymbol{\gamma}\cdot\nabla - m)\Psi|_{\partial_t \to i\partial_\tau, \gamma_i = -i\boldsymbol{\gamma}}.$$

(4) The index contractions are now done with $\delta_{\mu\nu}$, where $\mu = 0$ means τ.

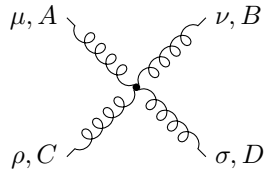
The Feynman rules are as follows. We use the convention $p_\mu = (\omega_n, \mathbf{p})$ and the projectors $P_{\mu\nu}^\mathrm{T} = \delta_{\mu\nu} - p_\mu p_\nu/p^2$, $P_{\mu\nu}^\mathrm{L} = p_\mu p_\nu/p^2$. The projectors satisfy $P^\mathrm{T}P^\mathrm{L} = 0$, $(P^\mathrm{T})^2 = P^\mathrm{T}$ and $(P^\mathrm{L})^2 = P^\mathrm{L}$.

$$\begin{aligned}
\nu, B \;\text{⌒⌒⌒⌒⌒⌒}\; \mu, A \quad &= \frac{\delta_{\mu\nu} - (1-\xi)p^\mu p^\nu/p^2}{p^2}\delta^{AB} = \frac{\delta^{AB}}{p^2}(P_{\mu\nu}^\mathrm{T} + \xi P_{\mu\nu}^\mathrm{L}),\\
i \longrightarrow j \quad &= \frac{\delta^{ij}}{\not{p} + m} = \delta^{ij}\frac{-i\not{p} + m}{p^2 + m^2},\\
B \dashrightarrow A \quad &= \frac{\delta^{AB}}{p^2}.
\end{aligned}$$

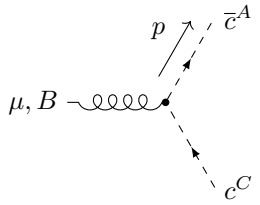
The vertex rules follow by the substitution $i \to -1$, $g_{\mu\nu} \to -\delta_{\mu\nu}$, $-G^2/4 \to G^2/4$.



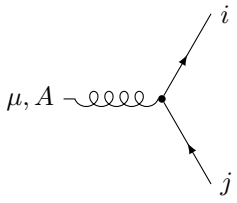
$$= igf^{ABC}\left(\delta^{\mu\nu}(k-p)^\rho + \delta^{\nu\rho}(p-q)^\mu + \delta^{\rho\mu}(q-k)^\nu\right),$$



$$= \begin{aligned} &-g^2[f^{ABE}f^{CDE}(\delta^{\mu\rho}\delta^{\nu\sigma} - \delta^{\mu\sigma}\delta^{\nu\rho}) + f^{ACE}f^{BDE}(\delta^{\mu\nu}\delta^{\rho\sigma} - \delta^{\mu\sigma}\delta^{\nu\rho}) \\ &+ f^{ADE}f^{BCE}(\delta^{\mu\nu}\delta^{\rho\sigma} - \delta^{\mu\rho}\delta^{\nu\sigma})] \end{aligned},$$



$$= igf^{ABC}p^\mu,$$



$$= ig\gamma^\mu T^A_{ij}.$$

The free energy density for free single scalar/fermion field at one-loop order is

$$\frac{F}{V} = -\frac{\ln Z}{\beta V} = \bigcirc = \mp\frac{1}{2}\sum_n \sum_{\mathbf{p}} \ln\left(\beta^2\left(\omega_n^2 + \mathbf{p}^2 + m^2\right)\right)\frac{-1}{\beta V} \overset{m^2=0}{=} -\frac{\pi^2 T^4}{90}\begin{cases}1, & \text{boson;}\\ 7/8, & \text{fermion.}\end{cases}.$$

With QCD, we have to consider more types of loops:

$$\frac{F}{V} = -\frac{\ln Z}{\beta V} = \bigcirc + \bigcirc + \bigcirc.$$

The eigenvalues of the determinant will always be the same for one loop $\beta^2(\omega_n^2 + \mathbf{p}^2 + m^2)$, so the only task is to find the proportionality factors:

- internal indices: gluons and ghosts $\delta^{AA} = N^2 - 1$, quarks $N_\mathrm{f}\delta^{aa} = N_\mathrm{f}N$,
- Lorentz indices: quarks $\mathrm{Tr}(\mathbb{1}_{4\times4}) = 4$, ghosts -2 (minus is for the Fermionic nature, factor 2 from the two ghost fields) and gluons $M_{\mu\nu} = P^\mathrm{T}_{\mu\nu} + P^\mathrm{L}_{\mu\nu} = (d-1) + 1 = d$.

The longitudinal part of the gluon propagator contributes $1/\xi$ to the determinant, but it only contributes a constant to the free energy, which we can ignore. In the end, we have

$$\frac{F}{V} = -\frac{\pi^2 T^4}{90}\left((N^2-1)(3+1-2) + \frac{7}{8}4N_\mathrm{f}N\right).$$

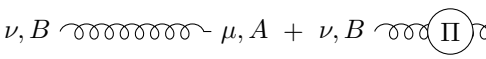
For QED, we have $N = 1$, and $N^2 - 1 = 1$. Using the above computation, we can calculate the properties of the quark-gluon plasma, however only for $T \gg \Lambda_\mathrm{QCD}$ to avoid quark confinement.

3.5.2 Gluon self-energy

The problems that arise in finite-temperature QCD self-energy calculation are the following

- we have additional scale T,
- spatial indices are distinguished (no Lorentz covariance),
- gluons and ghosts are bosons, which carry the zero Matsubara mode and cause IR divergences,
- spin.

The full 2-point function is

$$G^{AB}_{\mu\nu}(\omega_n,\mathbf{p}) =$$ 

Since spatial part is distinguished from the temporal one, we have to split the transversal projection operator:

$$P^{\mathrm{T}} = P^{\mathrm{TT}} + P^{\mathrm{TL}}, \quad P^{\mathrm{TT}}_{\mu\nu} = \begin{cases} 0 & \mu \text{ or } \nu = 0; \\ \delta_{ij} - \frac{p_i p_j}{\mathbf{p}^2} & \mu,\nu = i,j = 1,2,3, \end{cases}, \quad P^{\mathrm{TL}}_{\mu\nu} = P^{\mathrm{T}}_{\mu\nu} - P^{\mathrm{TT}}_{\mu\nu}.$$

The following relations also hold

$$(P^{\mathrm{TT}})^2 = P^{\mathrm{TT}}, \quad P^{\mathrm{TT}}\mathcal{P}^{\mathrm{T}} = P^{\mathrm{TT}}, \quad P^{\mathrm{TL}}P^{\mathrm{TT}} = 0.$$

We have

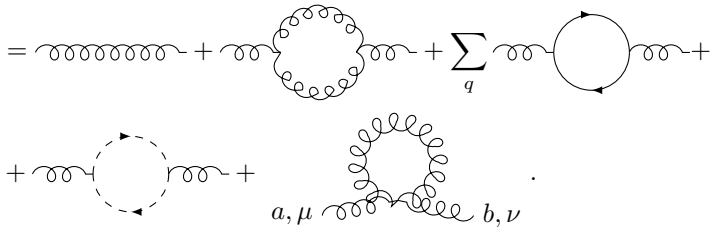
ν, B 

$$\begin{aligned} &= (-1)\frac{1}{(p^2)^2} p^2 \delta^{AB} \left(P^{\mathrm{T}}_{\mu\mu'} + \xi P^{\mathrm{L}}_{\mu\mu'}\right)\left(P^{\mathrm{TT}}_{\mu'\nu'}\Pi_{\mathrm{TT}} + P^{\mathrm{TL}}_{\mu'\nu'}\Pi_{\mathrm{TL}} + P^{\mathrm{L}}_{\mu'\nu'}\Pi_{\mathrm{L}}\right)\left(P^{\mathrm{T}}_{\nu'\nu} + \xi P^{\mathrm{L}}_{\nu'\nu}\right) \\ &= (-1)\frac{\delta^{AB}}{p^2}\left(P^{\mathrm{TT}}_{\mu\nu}\Pi_{\mathrm{TT}} + P^{\mathrm{TL}}_{\mu\nu}\Pi_{\mathrm{TL}} + \xi^2 P^{\mathrm{L}}_{\mu\nu}\Pi_{\mathrm{L}}\right). \end{aligned}$$

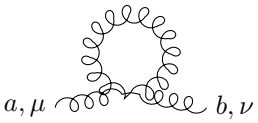
The result for the full propagator is

$$G^{AB}_{\mu\nu}(\omega_n,\mathbf{p}) = \frac{\delta^{AB}}{p^2}\left[\frac{P^{\mathrm{TT}}_{\mu\nu}}{1-\Pi_{\mathrm{TT}}} + \frac{P^{\mathrm{TL}}_{\mu\nu}}{1-\Pi_{\mathrm{TL}}} + \xi\frac{P^{\mathrm{L}}_{\mu\nu}}{1-\xi\Pi_{\mathrm{L}}}\right].$$

All we have to do now is to calculate the one-loop self-energy $\Pi^{AB}_{\mu\nu}(\omega_n,\mathbf{p})$:

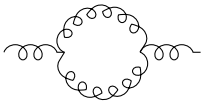
$$\Pi^{AB}_{\mu\nu}(\omega_n,\mathbf{p}) =$$ 

The internal momentum will be denoted by k and the external momentum is $p_\mu = (\omega_n, \mathbf{p})$. We are interested in calculating the thermal mass, so we take the limit $\omega_n = 0$, $\xi = 1$, $\mathbf{p} \to 0$ ($\mathbf{p} \ll \mathbf{k}, T$) and $m = 0$ (or equivalently $T \gg m$). Let's start with the gluon snail diagram:



$$= -\frac{1}{2}g^2 V_4 \frac{1}{\beta}\sum_m \int \frac{\mathrm{d}^d\mathbf{k}}{(2\pi)^d}\Delta(\omega_m,\mathbf{k}) = -g^2 dN\delta^{AB}\delta_{\mu\nu}I.$$

The integral I is defined the same as in the previous section, $1/2$ is the symmetry factor and V_4 is some constant that has Lorentz and gauge indices. Next is the gluon bubble:



$$= -\frac{1}{2}g^2 N_\mathrm{c}\delta^{AB}\frac{1}{\beta}\sum_m\int\frac{\mathrm{d}^d\mathbf{k}}{(2\pi)^d}\frac{V_{\mu\alpha\beta}(p,-k,k-p)V_{\nu\alpha\beta}(-p,k,p-k)}{k^2(p-k)^2},$$

where $V_{\alpha\beta\gamma}$ are the Lorentz factors from the three gluon vertex. The numerator can be simplified to

$$-\delta_{\mu\nu}(4p^2+(p-k)^2+k^2)+(d+4)p_\mu p_\nu+(2-4d)k_\mu k_\nu\to(-1)(2k^2\delta_{\mu\nu}+(4d-2)k_\mu k_\nu).$$

In the second step we assumed $\mathbf{p}\to 0$ and $n=0$. For convenience, we define the following expression:

$$I_{\mu\nu}=\frac{1}{\beta}\sum_m\int\frac{\mathrm{d}^d\mathbf{k}}{(2\pi)^d}\frac{k_\mu k_\nu}{k^4}.$$

Both I and $I_{\mu\nu}$ are scaleless for $T=0$ but are finite since UV and IR divergent terms vanish at $T=0$. Their values are

$$I=\begin{cases}T^2/12, & \text{bosons;}\\ -T^2/24, & \text{fermions.}\end{cases}.$$

For the other integral, it is more convenient to split it according to 3-dimensional rotational invariance:

$$I_{\mu\nu}=\delta_{\mu0}\delta_{\nu0}I_0+\delta_{\mu i}\delta_{\nu i}I_\mathrm{s}.$$

There is no summation over i. We have

$$I_{\{0,\mathrm{s}\}}=\frac{1}{\beta}\sum_m\int\frac{\mathrm{d}^dk}{(2\pi)^d}\frac{\{k_0^2,k_i^2\}}{k^4}.$$

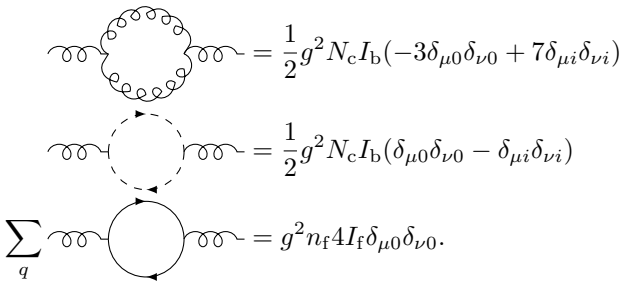
We can notice a special relation in I, that it is almost proportional to its derivative

$$I=\frac{T}{2}\frac{\mathrm{d}}{\mathrm{d}T}I=\frac{T}{2}\frac{\mathrm{d}}{\mathrm{d}T}\left(T\sum_n\int\frac{\mathrm{d}^dk}{(2\pi)^d}\frac{1}{\omega_n^2+\mathbf{k}^2}\right)=-\frac{I}{2}-T\sum_n\int\frac{\mathrm{d}^dk}{(2\pi)^d}\frac{\omega_n^2}{(\omega_n^2+\mathbf{k}^2)^2}=\frac{I}{2}-\frac{I_0}{2}.$$

So, without much calculation, we determined that $I_0=-I/2$. Similarly, we can find I_s:

$$I_\mathrm{s}=\frac{1}{d}\sum_n\int\frac{\mathrm{d}^dk}{(2\pi)^d}\frac{\mathbf{k}^2}{(\omega_n^2+\mathbf{k}^2)^2}=\frac{1}{d}(I-I_0)=\frac{3}{2d}I.$$

This is a version of *Passarino–Veltman reduction*. Putting everything together, the three possible bubbles are:



$$= \frac{1}{2}g^2N_\mathrm{c}I_\mathrm{b}(-3\delta_{\mu0}\delta_{\nu0}+7\delta_{\mu i}\delta_{\nu i})$$

$$= \frac{1}{2}g^2N_\mathrm{c}I_\mathrm{b}(\delta_{\mu0}\delta_{\nu0}-\delta_{\mu i}\delta_{\nu i})$$

$$\sum_q = g^2n_\mathrm{f}4I_\mathrm{f}\delta_{\mu0}\delta_{\nu0}.$$

Bringing all that together into the self-energy, we get

$$\Pi_{\mu\nu}(p)|_{\mathbf{p}\to 0} = m_{\mathrm{T}}^2 \delta_{\mu 0}\delta_{\nu 0} = \frac{1}{6} g^2 (2N_{\mathrm{c}} + n_{\mathrm{f}}) \delta_{\mu 0}\delta_{\nu 0} T^2$$

and the two-point function is

$$G_{\mu\nu} = \frac{\delta^{AB}\delta_{\mu\nu}}{\omega_n^2 + \mathbf{p}^2 + m_{\mathrm{T}}^2 \delta_{\mu 0}\delta_{\nu 0}}.$$

Remark.

(1) $\delta_{\mu i}\delta_{\nu i}$ terms dropped out of the sum as a consequence of transversality (there are no higher loop contributions to Π_{L}, the same as at $T = 0$).

(2) The m_{T}^2 is called the *Debye mass* (electrical thermal mass comes from 0-component).

(3) The magnetic fields are not screened at 1-loop order, i.e. there is no magnetic mass.

(4) The QED limit can be obtained by the substitution $g^2 \to e^2$, $N_{\mathrm{c}} \to 0$, $n_{\mathrm{f}}/2 \to 1$.

There are IR divergent terms, which we need to treat with resummation. The consequence is non-analytic contributions to the perturbative series in g (such as g^3 and $g^4 \ln g$).

3.5.3 Electrostatic and magnetic QCD

We want to construct an effective theory where all Matsubara modes $n \neq 0$ which have mass $\mathcal{O}(\pi T)$ are "integrated out". There is an EFT valued at scales $\mu \ll T$, where only the bosonic zero mode is left (and fermions are integrated out completely). It will still hold $\partial_0 A_\mu = 0$. The gluon field zero mode transforms as

$$A_0 \to \mathcal{U} A_0 \mathcal{U}^{-1}, \quad A_i \to \mathcal{U} A_i \mathcal{U}^{-1} + \frac{i}{g} \mathcal{U} \partial_i \mathcal{U}^{-1}.$$

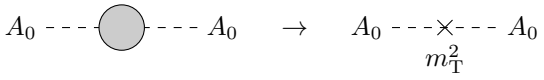
The spatial part corresponds to a gauge field in three dimensions, while A_0 transforms as an adjoint scalar (zero mode is independent of time, thus no inhomogeneous term in the transformation). This means that EFT should be a 3-dimensional Yang–Mills theory with an adjoint scalar A_0. The field strength tensor is $G_{0i}^A = -\mathrm{D}_i^{AB} A_0^B$, so the kinetic term of the Lagrangian becomes

$$\frac{1}{4} G_{ij}^A G_{ij}^A = \frac{1}{4} G_{ij}^A G_{ij}^A + \frac{1}{2} (\mathrm{D}_i^{AB} A_0^B)(\mathrm{D}_i^{AC} A_0^C).$$

Acting with a generator on the temporal part of $G_{\mu\nu}$ gives $T^A \mathrm{D}_i^{AB} A_0^B = [\mathbf{D}_i, \mathbf{A}_0]$. The possible terms in the effective Lagrangian with dimension 4 or less are

$$S_{\mathrm{eff}} = \beta \int \mathrm{d}^3\mathbf{x} \left[\frac{1}{2} \operatorname{tr}(\mathbf{G}_{ij}\mathbf{G}_{ij}) + \operatorname{tr}([\mathrm{D}_i, \mathbf{A}_0][\mathrm{D}_i, \mathbf{A}_0]) + m_{\mathrm{T}}^2 \operatorname{tr}(A_0^2) + \lambda_{41} [\operatorname{tr}(\mathbf{A}_0^2)]^2 + \lambda_{42} \operatorname{tr}(\mathbf{A}_0^4) \right].$$

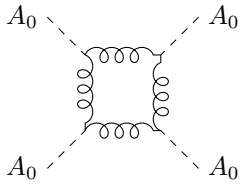
The matching of m_{T}^2 term with the 2-point function (with A_0 legs) in the IR limit gives the electric/Debye thermal mass.



The 2-point function with fields A_i changes the coefficient of $\mathrm{tr}(G_{ij}G_{ij})$ and is absorbed into the definition of the effective gauge coupling g_{eff}:

$$gA_i \to g_{\mathrm{eff}} A_i, \quad g_{\mathrm{eff}} = g(1 - \#g^2 + \dots).$$

The effective description for λ_{41} and λ_{42} come from gluon loops with four A_0 legs (the A_0 four point function that comes from the gauge coupling is zero).



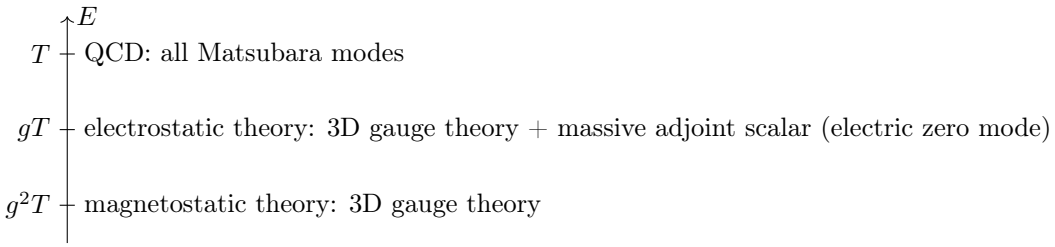
If the field A_0 is massive, its mass is proportional to gT and typical momenta are also proportional to gT (so-called soft scaling). A power counting argument gives

$$T\{A_0, A_0\} \sim T \int \mathrm{d}^3\mathbf{p} \frac{1}{\mathbf{p}^2 + m_{\mathrm{T}}^2} \sim T(gT)^3 \cdot \frac{1}{(gT)^2} \sim gT^2.$$

Hence, $A_0 \sim \sqrt{g}T$ and a similar calculation shows $A_i \sim \sqrt{g}T$ for momenta of size $p \sim gT$. The covariant derivative is made of two parts, the regular derivative scales as gT and the $-ig_{\mathrm{eff}}A_i$ scales as $g^{3/2}T$. This means that gauge boson is perturbative in $\sqrt{g}$. Furthermore,

$$\beta \int \mathrm{d}^3\mathbf{x}\, \lambda_4 A_0^4 \sim \frac{1}{T}\frac{1}{(gT)^3} g^4 (\sqrt{g}T)^4 \sim g^3.$$

In general, any term in S_{eff} is assigned a power g^a. What happens beyond the scale gT? The scalar field A_0 must be integrated out and only A_i (magnetic) fields remain.

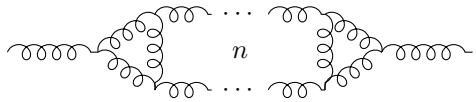


The effective action in the magnetic regime is

$$S_{\mathrm{eff}}^{\mathrm{mag}} = \beta \int \mathrm{d}^3\mathbf{x} \left[\frac{1}{2}\,\mathrm{tr}(\mathbf{G}_{ij}\mathbf{G}_{ij}) + \frac{c}{T^2}\,\mathrm{tr}([\mathrm{D}_i, \mathbf{G}_{ik}][\mathrm{D}_j, \mathbf{G}_{jk}]) + \text{6-dim terms} \right].$$

This is actually a 3-dimensional Yang-Mills theory, which is super-renormalizable and has a dimensional coupling $\hat{g} = g^2T$ (the scale is similar as the magnetic scale). We can rescale the fields and coupling as $A_i = \sqrt{T}\hat{A}_i$ and $g_{\mathrm{eff}} = \hat{g}/\sqrt{T}$. With this rescaling, the field strength tensor is $\hat{G} \sim \partial\hat{A} - \partial\hat{A} + \hat{g}[\hat{A}, \hat{A}]$. But since $\hat{g}$ has the magnetic scale, this theory is non-perturbative. Even if $T \gg \Lambda_{\mathrm{QCD}}$ ($g(T) \ll 1$), the contributions from the zero magnetic mode are non-renormalizable. One can match perturbatively to S_{mag}, but then the magnetic contributions must be obtained non-perturbatively, e.g., by putting $S_{\mathrm{eff}}^{\mathrm{mag}}$ on a 3-dimensional Euclidean lattice. This is the so-called *Linde problem* (1980).

Consider the self-energy of the magnetic mode and assume it has some thermal mass M.

I_n :  $\sim g^{2n}\left(T\int \mathrm{d}^3\mathbf{p}\right)^n \frac{(\mathbf{p}^2)^n}{(\mathbf{p}^2+M^2)^{3n-1}}$

$$\sim \int \frac{g^{2n}T^n p^{5n}}{(p^2+M^2)^{3n-1}}.$$

- Suppose $M^2 = 0$. The loop diagrams become more and more IR divergent when we add loops $I_n \sim g^{2n}T^n/p^{n-2}$.
- Suppose $M^2 \sim (gT)^2$ (this is not true, but we can still try) and $p \sim M$. Then $I_n \sim g^n(gT)^2$, which means that every loop gives a suppression by g, so we can do a perturbative expansion.
- Suppose $M^2 \sim (g^2T)^2$. We have to use $p \sim g^2T$ to regulate the denominator and this gives $I_n = g^4T^2$, which is independent of n. This is a small contribution, but not perturbative – loops are not suppressed. We get a similar behaviour for $\ln Z$ and p (where we consider only vacuum diagrams):

$$p = cT^4(1 + \#g^2 + \#g^3 + \cdots + g^6 \cdot (\text{non-perturbative} + \dots)).$$

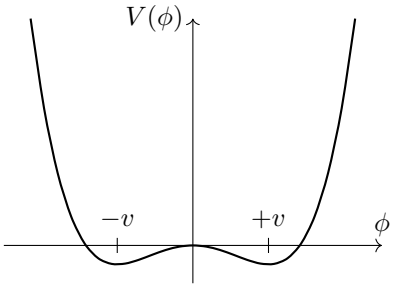
3.6 Phase transitions and symmetry restoration at finite temperature

There are two relevant cases.

- Strong interaction: at $T \approx 150\,\mathrm{MeV}$ quarks and gluons are de-confined and chiral symmetry is restored.
- Weak interaction: at $T \approx 160\,\mathrm{GeV}$ the Higgs field expectation value vanishes and $\mathrm{SU}(2)_\mathrm{L}$ symmetry is restored ($m_{W,Z} \to 0$, $m_q \to 0$, there is still a thermal mass).

We will discuss only $\lambda\phi^4$ interaction. The Euclidean Lagrangian is

$$\mathcal{L}_\mathrm{E} = \frac{1}{2}(\partial_\mu\phi)(\partial_\mu\phi) + V(\phi), \quad V(\phi) = -\frac{1}{2}m^2\phi^2 + \frac{\lambda}{4}\phi^4.$$



The vacuum expectation value of the field is $\langle\phi\rangle = \pm m/\sqrt{\lambda}$. The phase transition can happen if $Z(\mu, T, V)$ is non-analytic in T, μ in the infinite volume limit. If the energy density

$$\frac{E}{V} = \frac{T^2}{V}\frac{\partial}{\partial T}\ln Z|_{\text{fixed } \mu/T}$$

is discontinuous at some T_c, there is a first-order phase transition. If E/V is continuous but $C_V = \partial(E/V)/\partial T$ is discontinuous, there is a second-order phase transition. Phase transitions are associated by an order parameter that is discontinuous (first order) or continuous with a discontinuous derivative (second order) at T_c. In field theory, the order parameter is usually the vacuum expectation value of some operator that breaks the symmetry. If the order parameter is zero at $T > T_c$ and non-zero at $T < T_c$, we say that the symmetry is restored at high temperature.

Recall from QFT2 that the ground state (vacuum) is determined by the absolute minimum of the effective potential (assuming the Lorentz symmetry is not broken)

$$\Gamma[\phi] = W[J] - \int \mathrm{d}^4x\ \phi J \equiv \int \mathrm{d}^4x\ \left(-V_{\mathrm{eff}}(x) + \frac{1}{2}Z[\phi]\partial_\mu\phi\partial^\mu\phi\right)$$

where

$$V_{\mathrm{eff}} = -\frac{1}{VT}\Gamma[\phi = \text{classical field}].$$

In the path integral formalism, we can expand the field around the classical field $\phi \to \varphi + \phi$ and do the Gaussian integral over φ:

$$\begin{aligned}
e^{i\Gamma[\phi]} &= e^{iW[J]}e^{-i\int \mathrm{d}^4x\ \phi J} \\
&= \int \mathcal{D}[\varphi] e^{iS[\varphi]+i\int \mathrm{d}^4x\ J(\varphi-\phi)} \\
&= \int \mathcal{D}[\varphi] \exp\left(iS[\phi] + i\int \mathrm{d}^4x\ \left.\frac{\delta S}{\delta\varphi(x)}\right|_\phi \varphi(x) + \right. \\
&\qquad \left. + i\frac{1}{2}\int \mathrm{d}^4x\ \mathrm{d}^4y\ \left.\frac{\delta^2 S}{\delta\varphi(x)\delta\varphi(y)}\right|_\phi \varphi(x)\varphi(y) + i\int \mathrm{d}^4x\ J\varphi \right) \\
&= e^{iS[\phi]} \int \mathcal{D}[\varphi] \exp\left(\frac{i}{2}\int \mathrm{d}^4x\ \int \mathrm{d}^4y\ \left.\frac{\delta^2 S}{\delta\varphi(x)\delta\varphi(y)}\right|_\phi \varphi(x)\varphi(y) \right) \\
&\equiv \exp\left(-\frac{i}{VT}V_{\mathrm{eff}}[\phi]\right).
\end{aligned}$$

For finite temperature, we replace $VT \to V\beta$, $iS \to -S_{\mathrm{E}}$ and

$$\int \mathrm{d}^4x\ = \int_0^\beta \mathrm{d}\tau \int \mathrm{d}^3\mathbf{x}.$$

The effective potential at one-loop order is

$$e^{-\beta V V_{\mathrm{eff}}(\phi)} = e^{-\beta V V(\phi)} \int \mathcal{D}[\varphi] \exp\left(-\frac{1}{2}\int_0^\beta \mathrm{d}\tau \int \mathrm{d}^3\mathbf{x}\, \varphi(-\partial_\tau^2 - \nabla^2 + m_{\mathrm{eff}}^2(\phi))\varphi \right).$$

The effective mass depends on the classical field as $m_{\mathrm{eff}}^2 = -m^2 + 3\lambda\phi^2$ and

$$\left.\frac{\delta S_{\mathrm{E}}}{\delta\varphi\delta\varphi}\right|_\phi = -\frac{1}{2}(m^2 - 3\lambda\phi^2).$$

The corrections to the effective potential can be expressed perturbatively, so to one loop order, we have

$$V_{\mathrm{eff}}(\phi) = V(\phi) + V_{T=0}^{\text{1-loop}}(\phi) + V_{T\neq 0}^{\text{1-loop}}(\phi).$$

The one loop correction is exactly the one we had at the very beginning of this chapter

$$\frac{F}{V} = -\frac{\pi^2 T^4}{90} + \frac{m_{\text{eff}}^2 T^2}{24} - \frac{m_{\text{eff}}^3 T}{12\pi} + \mathcal{O}(m^4 \ln m).$$

- If we keep only the first term in the correction $m_{\text{eff}}^2 T^2/24$, we get

$$V_{\text{eff}}(\phi, T) = \phi_{T\text{-indep}} + \frac{1}{2}\left(-m^2 + \frac{1}{4}\lambda T^2\right)\phi^2 + \frac{\lambda T^4}{4}.$$

  The vacuum expectation value is

$$\phi_0 = \begin{cases} \frac{1}{2}\sqrt{4m^2/\lambda - T^2}, & T < T_{\text{c}} = 2m/\sqrt{\lambda}; \\ 0, & T \geq T_{\text{c}}. \end{cases}$$

  What is the order of this transition? We have

$$\frac{\mathrm{d}}{\mathrm{d}T} \ln Z|_{\phi_{\min}} = \frac{\mathrm{d}}{\mathrm{d}T} V_{\text{eff}}(\phi_0(T), T) = \frac{\partial V_{\text{eff}}}{\partial \phi_0}\frac{\partial \phi_0}{\partial T} + \left.\frac{\partial V_{\text{eff}}}{\partial T}\right|_{\phi_{\min}}.$$

  V_{eff} is a smooth function of T and ϕ (in this approximation), hence the derivative is not discontinuous, and we have second-order transition. If we add the cubic term $-m_{\text{eff}}^3 T/(12\pi)$, the situation changes. There can now be two local minima for $\lambda > 2\pi^2/3$. However, this is the regime of λ where perturbation theory is not valid, so we cannot trust this result. The size of perturbative correction at three-loop order is

$$\lambda T \int \frac{\mathrm{d}^3\mathbf{p}}{\mathbf{p}^4} \sim \frac{\lambda T}{m_{\text{eff}}} \sim \frac{\lambda T}{\sqrt{\lambda}|\phi|} = \frac{\sqrt{\lambda} T}{|\phi|}.$$

  The perturbative theory fails for $|\phi| < \sqrt{\lambda} T$, and so we cannot determine the order of the phase transition perturbatively.

3.7 Real-time formalism

So far, we only discussed static properties of the thermal system, such as the partition function and the effective potential. However, we are also interested in dynamical properties, such as transport coefficients and particle production rates. The object of interest is the ensemble average of some operator $\mathcal{O}$ as a function of time:

$$\langle \Omega | \mathcal{O}(t) | \Omega \rangle = \operatorname{tr}(\rho \mathcal{O}(t)).$$

3.7.1 Linear response

We split the Hamiltonian into two parts, $H = H_{\text{bath}} + H_{\text{ext}}(t)$, where H_{bath} is time independent and $H_{\text{ext}}(t)$ is a small time-dependent perturbation

$$H_{\text{ext}}(t) = \int \mathrm{d}^3\mathbf{x}\, \mathcal{J}(t, \mathbf{x}) Q(t, \mathbf{x}).$$

The external source $\mathcal{J}$ couples to some operator Q built from the fields of H_{bath}. The Heisenberg equation for the operator $\mathcal{O}$ in thermal bath is

$$\frac{\partial \mathcal{O}}{\partial t} = i[H(t), \mathcal{O}].$$

Let $|\psi\rangle$ be the energy eigenstate of H_{bath} with eigenvalue E. The time evolution operator is

$$\partial_t \langle\psi|\mathcal{O}(t)|\psi\rangle = i\langle\psi|[H_{\text{ext}}(t),\mathcal{O}]|\psi\rangle$$

and the variation of it becomes

$$\delta\langle\psi|\mathcal{O}(t)|\psi\rangle = \langle\psi|\mathcal{O}(t)|\psi\rangle - \langle\psi|\mathcal{O}(t_0)|\psi\rangle = i\int_{t_0}^{t} \mathrm{d}t'\ \langle\psi|[H_{\text{ext}}(t'),\mathcal{O}(t)]|\psi\rangle\,.$$

The operator $\mathcal{O}$ only depends on t in this expression (not t'), since the difference is already of higher order. Since this holds for any eigenstate, we can take the ensemble average with the density matrix $|\psi\rangle \to \mathrm{tr}(\rho)$:

$$\begin{aligned}\delta\langle\mathcal{O}(t,\mathbf{x})\rangle &= i\int_{t_0}^{t}\mathrm{d}t'\ \mathrm{tr}\left(\rho[H_{\text{ext}}(t'),\mathcal{O}(t,\mathbf{x})]\right)\\ &= (-i)\int_{t_0}^{\infty}\mathrm{d}t'\int\mathrm{d}^3\mathbf{x}'\ \underbrace{\mathrm{tr}\left(\rho[\mathcal{O}(t,\mathbf{x}),Q(t',\mathbf{x}')]\theta(t-t')\right)}_{\text{retarded 2-point function}}\cdot\mathcal{J}(t',\mathbf{x}).\end{aligned}$$

The retarded 2-point function is denoted by $-i\Pi^{\mathrm{R}}_{\mathcal{O}Q}(t,\mathbf{x},t',\mathbf{x}')$. If $\mathcal{O}=Q=\phi$ is the elementary scalar field, the $\Pi^{\mathrm{R}}_{\phi\phi}=\Pi^{\mathrm{R}}$ is the exact retarded ϕ propagator (because H_{bath} includes interactions). When H_{bath} is in equilibrium, Π^{R} depends only on the time differences $t-t'$ and $\mathbf{x}-\mathbf{x}'$. Setting $t_0\to-\infty$ and Fourier transforming with

$$f(t,\mathbf{x}) = \int\frac{\mathrm{d}\omega\,\mathrm{d}^3\mathbf{k}}{(2\pi)^4}e^{-i\omega t+i\mathbf{k}\cdot\mathbf{x}}\hat{f}(\omega,\mathbf{k})$$

gives

$$\delta\langle\tilde{\mathcal{O}}(\omega,\mathbf{k})\rangle = \tilde{\mathcal{J}}(\omega,\mathbf{k})\Pi^{\mathrm{R}}_{\mathcal{O}Q}(\omega,\mathbf{k}).$$

The first term is the change in the ensemble average, the second is the external source and the last is the retarded two-point function. This is a very important relation in linear response theory.

3.7.2 Green's functions and spectral function

We will use the notation $\langle\ldots\rangle = \mathrm{tr}(\rho\ldots)$ and since we have real time, the four-vectors are Minkowski $k^\mu=(\omega,\mathbf{k})$, $x^\mu=(t,\mathbf{x})$. The two point correlator (the retarded Green's function) for the complex scalar field ϕ is expressed as

$$\Pi^{\mathrm{R}}(k) = i\int\mathrm{d}^4x\ e^{ikx}\langle[\phi(x),\phi^\dagger(0)]\theta(t)\rangle\,,$$

and we can also define the advanced Green's function

$$\Pi^{\mathrm{A}}(k) = i\int\mathrm{d}^4x\ e^{ikx}\langle-[\phi(x),\phi^\dagger(0)]\theta(-t)\rangle\,.$$

The definition can be generalized to any type of field, we just have to keep in mind that commutator is replaced by an anti-commutator for fermions. We can also define Green's functions

without time ordering:

$$\Pi^{>}(k) = \int \mathrm{d}^4x\, e^{ikx} \langle \phi(x)\phi^\dagger(0)\rangle\,,$$
$$\Pi^{<}(k) = \int \mathrm{d}^4x\, e^{ikx} \langle \phi^\dagger(0)\phi(x)\rangle\,,$$
$$\Pi^{\mathrm{T}}(k) = \int \mathrm{d}^4x\, e^{ikx} \langle T\{\phi(x)\phi^\dagger(0)\}\rangle\,,$$
$$\Pi^{\overline{\mathrm{T}}}(k) = \int \mathrm{d}^4x\, e^{ikx} \langle \overline{T}\{\phi(x)\phi^\dagger(0)\}\rangle\,.$$

Different Green's functions are related to each other:

$$\Pi^{\mathrm{T}} + \Pi^{\overline{\mathrm{T}}} = \Pi^{>} + \Pi^{<}, \quad -i\Pi^{\mathrm{R}} = \Pi^{\mathrm{T}} - \Pi^{<}, \quad -i\Pi^{\mathrm{A}} = \Pi^{\overline{\mathrm{T}}} - \Pi^{>}.$$

Moreover, we define the *spectral function* as

$$\rho(k) = \int \mathrm{d}^4x\, e^{ikx} \langle [\phi(x), \phi^\dagger(0)]/2\rangle = \frac{1}{2}(\Pi^{>} - \Pi^{<}) = -\frac{i}{2}(\Pi^{\mathrm{R}} - \Pi^{\mathrm{A}})$$

and the *statistical function*

$$\Pi_h(k) = \int \mathrm{d}^4x\, e^{ikx} \langle \{\phi(x), \phi^\dagger(0)\}/2\rangle = \frac{1}{2}(\Pi^{>} + \Pi^{<}).$$

We want to see how these functions are related to Euclidean (imaginary time) two-point function

$$\Pi_{\mathrm{E}}(\omega_n, \mathbf{k}) = \int_0^\beta \mathrm{d}\tau \int \mathrm{d}^3\mathbf{x}\, e^{i\omega\tau - i\mathbf{k}\cdot\mathbf{x}} \langle \phi(\tau,\mathbf{x})\phi^\dagger(0)\rangle\,.$$

In thermal equilibrium (at $\mu = 0$), it holds

$$\begin{aligned}\langle \phi(t - i\beta, \mathbf{x})\phi^\dagger(0,\mathbf{0})\rangle &= \frac{1}{Z}\,\mathrm{tr}\left(e^{-\beta H}(e^{iH(-i\beta)}\phi(t,\mathbf{x})e^{-iH(-i\beta)})\phi^\dagger(0,\mathbf{0})\right)\\ &= \frac{1}{Z}\,\mathrm{tr}\left(\phi^\dagger(0,\mathbf{0})\phi(t,\mathbf{x})e^{-\beta H}\right)\\ &= \langle \phi^\dagger(0,\mathbf{0})\phi(t,\mathbf{x})\rangle\,.\end{aligned}$$

In the last step we used the cyclicity of the trace. We want to show that there is only a simple two-point function and all other Green's functions are in thermal equilibrium completely determined from one. There is the KMS (Kubo–Martin–Schwinger) relation between $\Pi^{>}(k)$ and $\Pi^{<}(k)$:

$$\begin{aligned}\Pi^{>}(k) &= \frac{1}{Z}\int \mathrm{d}^4x\, e^{ikx}\,\mathrm{tr}\left(e^{-\beta H}e^{iHt}\phi(0,\mathbf{x})e^{-iHt}\phi^\dagger(0,\mathbf{0})\right)\\ &= \frac{1}{Z}\sum_{m,n}\int \mathrm{d}^4x\, e^{ikx}e^{(-\beta+it)E_m}e^{-itE_n}\,\langle m|\phi(0,\mathbf{x})|n\rangle\,\langle n|\phi^\dagger(0,\mathbf{0})|m\rangle\\ &= \frac{1}{Z}\sum_{m,n}\int \mathrm{d}^3\mathbf{x}\, e^{-i\mathbf{k}\cdot\mathbf{x}}e^{-\beta E_m}2\pi\delta(k^0 + E_m - E_n)\,\langle m|\phi(0,\mathbf{x})|n\rangle\,\langle n|\phi^\dagger(0,\mathbf{0})|m\rangle\,.\end{aligned}$$

Similarly, we can show that

$$\Pi^{<}(k) = \frac{1}{Z}\sum_{m,n}\int \mathrm{d}^3\mathbf{x}\, e^{-i\mathbf{k}\cdot\mathbf{x}}e^{-\beta E_n}2\pi\delta(k^0 + E_m - E_n)\,\langle m|\phi(0,\mathbf{x})|n\rangle\,\langle n|\phi(0,\mathbf{0})|m\rangle\,.$$

Hence, in thermal equilibrium, it holds

$$\boxed{\Pi^{>}(k) = e^{-\beta k_0}\Pi^{<}(k).}$$

The spectral function is expressed in terms of $\Pi^{>}$ and $\Pi^{<}$, so we can now express it in terms of either

$$\rho(k) = \frac{1}{2}(e^{\beta k_0} - 1)\Pi^{>}.$$

But actually, we choose to express everything in terms of spectral function

$$\Pi^{>}(k) = 2n_{\mathrm{B}}(k_0)\rho(k), \quad \Pi^{<}(k) = 2(1 + n_{\mathrm{B}}(k_0))\rho(k), \quad \Pi_h(k) = (1 + 2n_{\mathrm{B}}(k_0))\rho(k)$$

To express the retarded and advanced functions, we use the following prescription for Heaviside function

$$\theta(t) = i\int_{-\infty}^{\infty}\frac{\mathrm{d}\omega}{2\pi}\frac{e^{-i\omega t}}{\omega + i\varepsilon}.$$

The retarded Green's function becomes

$$\Pi^{\mathrm{R}}(k) = i\int \mathrm{d}^4x\, e^{ikx}2\theta(t)\int\frac{\mathrm{d}^4p}{(2\pi)^4}e^{-ipx}\rho(p) = \int_{-\infty}^{\infty}\frac{\mathrm{d}p_0}{2\pi}\frac{2\rho(p_0,\mathbf{k})}{p_0 - k_0 - i\varepsilon}.$$

The calculation for the advanced function is analogous, we just exchange $\varepsilon \to -\varepsilon$. Note that this is not the Feynman prescription. In the latter case, the poles are shifted in different directions from the real axis, but in the advanced/retarded prescription, both poles are displaced in the same direction (above or below the real axis). From $\mathrm{Im}\,(1/(x \pm i\varepsilon)) = \pm\pi\delta(x)$, we can extract the discontinuity in Π:

$$\mathrm{Im}\,\Pi^{\mathrm{R/A}}(k) = \pm\rho(k).$$

Finally, we can relate Π^{T} (and $\Pi^{\overline{\mathrm{T}}}$) to the spectral function using $-i\Pi^{\mathrm{R}} = \Pi^{\mathrm{T}} - \Pi^{<}$. Hence, all Green's functions follow from $\rho(k)$. We still have the Euclidean two-point function to relate. But it is only $\Pi^{>}$ evaluated at $t = -i\beta$. Hence, in the momentum space it holds

$$\begin{aligned}\Pi^{\mathrm{E}}(\omega_n,\mathbf{k}) &= \int_0^{\beta}\mathrm{d}\tau\int\frac{\mathrm{d}p_0}{2\pi}e^{(i\omega_n - p_0)\tau}\Pi^{>}(p_0,\mathbf{k})\\ &= \int\frac{\mathrm{d}p^0}{2\pi}\left[\frac{e^{i(\omega_n - p^0)\tau}}{i\omega_n - p^0}\right]_0^{\beta}2(1 + n_{\mathrm{B}}(k^0))\rho(k)\\ &= \int\frac{\mathrm{d}k_0}{2\pi}\frac{2\rho(k_0,\mathbf{k})}{k^0 - i\omega_n}.\end{aligned}$$

In the first step, we replaced $\mathrm{d}^3\mathbf{x} \to (2\pi)^3\delta^{(3)}(\mathbf{p} - \mathbf{k})$ and $\mathrm{d}t \to 2\pi\delta(k^0 - p^0 - \omega)$. In the last step, we renamed $p^0 \to k^0$. We have seen before that spectral function is simply the discontinuity of the retarded/advanced Green's function. Hence, we can write

$$\rho(k) = \frac{1}{2i}\left(\Pi^{\mathrm{E}}(-i(k^0 + i\varepsilon),\mathbf{k}) - \Pi^{\mathrm{E}}(-i(k^0 - i\varepsilon),\mathbf{k})\right)$$

from which it follows that the retarded Green's function is merely the analytic continuation of the Euclidean one:

$$\boxed{\Pi^{\mathrm{R}}(k) = \Pi^{\mathrm{E}}(-i(k^0 + i\varepsilon),\mathbf{k}).}$$

3.7.3 Free theory

From the known Euclidean free two-point function

$$\Delta^{\mathrm{E}}(\omega_n, \mathbf{k}) = \frac{1}{\omega_n^2 + E_{\mathbf{k}}^2}$$

where $E_{\mathbf{k}} = \sqrt{\mathbf{k}^2 + m^2}$ we obtain the retarded propagator

$$\Delta^{\mathrm{R}}(k) = \frac{-1}{(k^0 + i\varepsilon)^2 - E_{\mathbf{k}}^2} = \frac{-1}{k^2 - m^2 + i\varepsilon\,\mathrm{sgn}(k^0)}$$

and

$$\rho(k) = \mathrm{Im}\ \Delta^{\mathrm{R}}(k) = \pi\delta(k^2 - m^2)(\theta(k^0) - \theta(-k^0)) = \frac{\pi}{2E_{\mathbf{k}}}(\delta(k^0 - E_{\mathbf{k}}) - \delta(k^0 + E_{\mathbf{k}})).$$

There is a relation that allows one to express the Bose–Einstein distribution at some energy and at negative energy $n_{\mathrm{B}}(k^0) + n_{\mathrm{B}}(-k^0) + 1 = 0$. With this, we can express the lesser and greater Green's function

$$\Delta^{<}(k) = 2n_{\mathrm{B}}(k^0)\rho(k) = 2\pi\delta(k^2 - m^2)(\theta(k^0)n_{\mathrm{B}}(k^0) + \theta(-k^0)(1 + n_{\mathrm{B}}(|k^0|))).$$

For the time ordered function, we have

$$\begin{aligned}\Pi^{\mathrm{T}} &= -i\Pi^{\mathrm{R}} + \Pi^{<} \\ &= \frac{i}{k^2 - m^2 - i\varepsilon\,\mathrm{sign}(k^0)} + 2\pi\theta(-k^0)\delta(k^2 - m^2) + 2\pi\delta(k^2 - m^2)n_{\mathrm{B}}(|k^0|) \\ &= \frac{i}{k^2 - m^2 + i\varepsilon} + 2\pi\delta(k^2 - m^2)n_{\mathrm{B}}(|k^0|)\end{aligned}$$

and

$$\Delta^{\mathrm{T}}(k) = \frac{i}{k^2 - m^2 + i\varepsilon} + 2\pi\delta(k^2 - m^2)n_{\mathrm{B}}(|k^0|), \quad \Delta^{\overline{\mathrm{T}}}(k) = \frac{-i}{k^2 - m^2 - i\varepsilon} + 2\pi\delta(k^2 - m^2)n_{\mathrm{B}}(|k^0|).$$

Both propagators separate nicely to vacuum part ($T = 0$) and thermal part.

- All thermal parts are on-shell, because of the delta function $\delta(k^2 - m^2)$.
- For fermions, we replace $n_{\mathrm{B}} \to n_{\mathrm{F}}$ and $1 + n_{\mathrm{B}}(|k^0|) \to -1 + n_{\mathrm{F}}(|k^0|)$. The numerators in the propagators become $1 \to \not{k} + m$.

3.7.4 Schwinger–Keldysh (closed-time-path) formalism

We will develop a path integral representation for $\langle \mathcal{O}(t) \rangle$ and $\langle \phi(x)\phi^\dagger(0) \rangle$ and its cousins. We will also explore the diagrammatic expansion at weak coupling for the path integral.

The assumption is that the initial state is specified at t_0 by

$$\rho = \sum\!\!\!\!\!\!\int_n \rho_n \left|n\right\rangle \left\langle n\right| .$$

The ensemble average of some (Heisenberg) operator $\mathcal{O}$ at time t is

$$\langle \mathcal{O}(t) \rangle = \sum\!\!\!\!\!\!\int_n \rho_n \left\langle n|\mathcal{O}(t)|n\right\rangle = \sum\!\!\!\!\!\!\int_n \sum\!\!\!\!\!\!\int_\alpha \rho_n \left\langle n|\alpha;t\right\rangle \left\langle \alpha;t|\mathcal{O}|n\right\rangle .$$

The integrand can be represented as a path integral

$$\begin{aligned}\langle\alpha;t|\mathcal{O}(t)|n\rangle &= \int \mathrm{d}\phi_1(t)\,\mathrm{d}\phi_1(t_0)\,\langle\alpha;t|\phi_1;t\rangle\,\langle\phi_1;t|\mathcal{O}(t)|\phi_1;t_0\rangle\,\langle\phi_1;t_0|n\rangle \\ &= \int \mathcal{D}[\phi_1(t),\pi_1(t)]\exp\left(i\int_{t_0}^{t}\mathrm{d}t'\int \mathrm{d}^3\mathbf{x}\,(\pi\dot{\phi}-\mathcal{H})\right)\mathcal{O}(\phi_1(t))\,\langle\alpha;t|\phi_1;t\rangle\,\langle\phi_1;t_0|n\rangle \\ &= \operatorname{index}\mathcal{D}[\phi_1(t)]\,\exp\left(i\int_{t_0}^{t}\mathrm{d}t'\int \mathrm{d}^3\mathbf{x}\,\mathcal{L}(\phi_1)\right)\mathcal{O}(\phi_1(t))\,\langle\alpha;t|\phi_1;t\rangle\,\langle\phi_1;t_0|n\rangle\,.\end{aligned}$$

Note that there are no boundary conditions. A similar result arises for $\langle n|\alpha;t\rangle = (\langle\alpha;t|n\rangle)^*$, for which we use field ϕ_2. Joining them together, we have

$$\begin{aligned}\langle\mathcal{O}(t)\rangle = \sum\!\!\!\!\!\!\int_n \rho_n \int_{\phi_1(t,\mathbf{x})=\phi_2(t,\mathbf{x})} \mathcal{D}[\phi_1(x),\phi_2(x)]\,\exp\left(i\int_{t_0}^{t}\mathrm{d}t'\,\mathrm{d}^3\mathbf{x}\,(\mathcal{L}(\phi_1)-\mathcal{L}(\phi_2))\right)\cdot \\ \cdot\,\langle n|\phi_2;t_0\rangle\,\langle\phi_1;t_0|n\rangle\,\mathcal{O}[\phi_1(t,\mathbf{x})] \\ = \int_{\phi_1(t,\mathbf{x})=\phi_2(t,\mathbf{x})} \mathcal{D}[\phi_1(x),\phi_2(x)]\,\exp\left(i\int_{t_0}^{t}\mathrm{d}t'\int \mathrm{d}^3\mathbf{x}\,(\mathcal{L}(\phi_1)-\mathcal{L}(\phi_2))\right)\langle\phi_1;t_0|\rho|\phi_2;t_0\rangle\,\mathcal{O}[\phi_1(t,\mathbf{x})].\end{aligned}$$

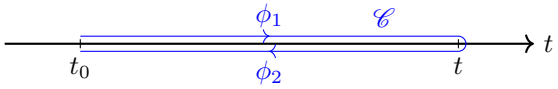
The boundary condition comes from

$$\sum\!\!\!\!\!\!\int_\alpha \langle\phi_2;t|\alpha;t\rangle\,\langle\alpha;t|\phi_1;t\rangle = \delta(\phi_2(t)-\phi_1(t)).$$

To construct Feynman rules for the perturbative expansion, define the generating functional with two sources J_1 and J_2 coupled to fields ϕ_1 and ϕ_2 respectively (it will also depend on the initial density matrix ρ):

$$\begin{aligned}Z[J_1,J_2,\rho] &= \int_{\phi_1(t,\mathbf{x})=\phi_2(t,\mathbf{x})} \mathcal{D}[\phi_1(x),\phi_2(x)]\,\exp\left(i\int_{t_0}^{t}\mathrm{d}t'\int \mathrm{d}^3\mathbf{x}\,(\mathcal{L}(\phi_1)-\mathcal{L}(\phi_2)+J_1\phi_1-J_2\phi_2)\right)\langle\phi_1;t_0|\rho|\phi_1;t_0\rangle \\ &= \exp\left(i\int_{t_0}^{t}\mathrm{d}t'\int \mathrm{d}^3\mathbf{x}\left[\mathcal{L}_{\mathrm{int}}\left(\frac{1}{i}\frac{\delta}{\delta J_1(x)}\right)-\mathcal{L}_{\mathrm{int}}\left(\frac{1}{i}\frac{\delta}{\delta J_2(x)}\right)\right]\right)Z_0[J_1,J_2,\rho] \\ &= \int \mathcal{D}[\phi_a]\,\exp\left[i\int_{\mathscr{C}}\mathrm{d}t'\int \mathrm{d}^3\mathbf{x}\,(\mathcal{L}(\phi_a)+J_a\phi_a)\right]\langle\phi_1;t_0|\rho|\phi_2;t_0\rangle\,,\end{aligned}$$

where $a\in\{1,2\}$ and $\mathscr{C}$ is a closed time path that goes from t_0 to t and back to t_0, where on the upper branch $\phi(t,\mathbf{x})=\phi_1(t,\mathbf{x})$ and on the lower branch $\phi(t,\mathbf{x})=\phi_2(t,\mathbf{x})$.



Since the time path goes backwards for ϕ_2 which naturally gives minus sign at $\mathcal{L}(\phi_2)$. The two-point functions are obtained by functional differentiation:

$$\Pi_{ab}(t_1,\mathbf{x}_1;t_2;\mathbf{x}_2) = \frac{\delta}{i\hat{a}\delta J_a(x_1)}\frac{\delta}{i\hat{b}J_b(x_2)}Z[J_1,J_2,\rho]|_{J_1=J_2=0}.$$

The letters with a hat are $+1$ for $a, b = 1$ and -1 for $a, b = 2$:

$$\hat{a} = \begin{cases} +1, & a = 1; \\ -1, & a = 2; \end{cases}, \quad \hat{b} = \begin{cases} +1, & b = 1; \\ -1, & b = 2. \end{cases}$$

Some authors will denote two different types of fields with ± 1 and with that notation, there is no need for inventing these new functions. The different choices of a and b correspond to

$$\begin{aligned} &\Pi_{11} \to \phi_1(x_1)\phi_1(x_2) \to \Pi^{\mathrm{T}}, \text{ T-ordered 2-point function} \\ &\Pi_{22} \to \phi_2(x_1)\phi_2(x_2) \to \Pi^{\overline{\mathrm{T}}}, \text{ } \overline{\mathrm{T}}\text{-ordered 2-point function} \\ &\Pi_{12} \to \phi_1(x_1)\phi_2(x_2) \to \Pi^{<}, \text{ no time ordering} \\ &\Pi_{21} \to \phi_2(x_1)\phi_1(x_2) \to \Pi^{>}, \text{ no time ordering} \end{aligned}$$

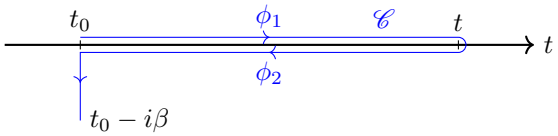
We thus define a generalized time ordering

$$T_{\mathscr{C}}(\phi(x)\phi^\dagger(y)) = \theta(x^0, y^0)\phi(x)\phi^\dagger(y) + \theta(y^0, x^0)\phi^\dagger(y)\phi(x),$$

where

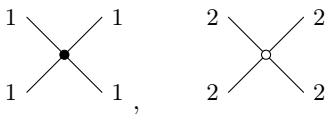
$$\theta(y^0, x^0) = \begin{cases} 1, & x^0 > y^0 \text{ and } x^0 \text{ and } y^0 \text{ are on the upper contour;} \\ 1, & x^0 < y^0 \text{ and } x^0 \text{ and } y^0 \text{ are on the lower contour;} \\ 0, & x^0 \text{ on the upper, } y^0 \text{ on the lower contour;} \\ 1, & x^0 \text{ on the lower, } y^0 \text{ on the upper contour.} \end{cases}$$

So far, we assumed a general form for ρ. Now, we will assume $\rho = e^{-\beta H}/Z$. Effectively, this adds a line from t_0 to $t_0 - i\beta$ to the previous path.



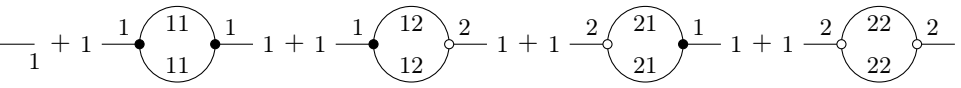
However, when we send $t_0 \to -\infty$, the interactions are switched off (by $i\varepsilon$ prescription). Hence, the only effect of ρ is that in the Feynman diagrams we use the free propagators in thermal equilibrium. The Feynman rules are as follows

- Vertices: there are two types of vertices, type 1 (from $\mathcal{L}_1(\phi_1)$) which we denote by a filled dot $+iV$ and type 2 (from $-\mathcal{L}_2(\phi_2)$) which we denote by an open dot $-iV$.



- Propagators: see the previous section.
- Diagrams: write all possible contractions $\overbrace{\phi_1\phi_1}$, $\overbrace{\phi_1\phi_2}$, $\overbrace{\phi_2\phi_1}$, $\overbrace{\phi_2\phi_2}$.
- Symmetry factors: as usual.

Example 2. Consider $\mathcal{L}_{\text{int}} = -g\phi^3/3!$ with 1-loop corrections to the ϕ propagator:

$\Pi_{11} =$  .

In all cases, the 1-loop correction can be written as (omitting symmetry factors and vertex factors)

$$\Pi_{ab}^{1-\text{loop}} = \sum_{a',b'} \Delta_{aa'} \Delta_{a'b'}^2 \Delta_{b'b}$$

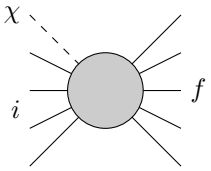
3.8 Particle production/decay in a medium

Suppose we have a thermal bath with two scalar particle types ϕ, ψ. There is also another scalar χ that interacts weakly with ϕ, ψ. The interaction Lagrangian is $\mathcal{L}_{\text{int}} = -g\phi\psi\chi$. First, we will discuss the production of χ. The amplitude for $i \to f + \chi$ at first order in g is

$$S_{fi} = -g \int \mathrm{d}^4x \, e^{iqx} \, \langle f|\mathcal{O}(x)|i\rangle \,,$$

where $\mathcal{O}(x) = (\phi\psi)(x)$ and $|i\rangle, |f\rangle$ only contain ϕ and ψ particles. The particle production per unit time and volume is

$$\mathrm{d}\Gamma = \frac{1}{VT} \sum_{i,f} \frac{1}{Z} e^{\beta E_i} |S_{fi}|^2 \frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3} \frac{1}{2E_{\mathbf{q}}}.$$



Let's do some calculations. The first thing to calculate is

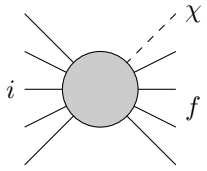
$$\begin{aligned}
\sum_f |S_{fi}|^2 &= g^2 \sum_f \int \mathrm{d}^4x \, \mathrm{d}^4y \, e^{iq(x-y)} \, \langle i|\mathcal{O}^\dagger(y)|f\rangle \, \langle f|\mathcal{O}(x)|i\rangle \\
&= g^2 \sum_f (2\pi)^4 \delta^{(4)}(q + p_i - p_f) \int \mathrm{d}^4x \, e^{iqx} \, \langle i|\mathcal{O}^\dagger(0)|f\rangle \, \langle f|\mathcal{O}(x)|i\rangle \\
&= VTg^2 \int \mathrm{d}^4x \, e^{iqx} \, \langle i|\mathcal{O}^\dagger(0)\mathcal{O}(x)|i\rangle \,.
\end{aligned}$$

In the second step, we used the momentum basis and $\mathcal{O}^\dagger(y) = e^{i(p_i - p_f)y}\mathcal{O}^\dagger(0)$. Now we plug this back into the production rate equation

$$\begin{aligned}
\frac{\mathrm{d}\Gamma}{\mathrm{d}^3\mathbf{q}} &= \frac{g^2}{(2\pi)^3 2E_{\mathbf{q}}} \sum_i \frac{1}{Z} e^{-\beta E_i} \int \mathrm{d}^4x \, e^{iqx} \, \langle i|\mathcal{O}^\dagger(0)\mathcal{O}(x)|i\rangle \\
&= \frac{g^2}{(2\pi)^3 2E_{\mathbf{q}}} \int \mathrm{d}^4x \, e^{iqx} \, \mathrm{tr}(\rho_{\text{thermal}} \mathcal{O}^\dagger(0)\mathcal{O}(x)) \\
&= \frac{g^2}{(2\pi)^3} \frac{1}{2E_{\mathbf{q}}} \Pi^<(q^0 = E_{\mathbf{q}}, \mathbf{q}).
\end{aligned}$$

The decay in medium can be calculated similarly (and it is one of the tasks on the homework). The final result is

$$\frac{\mathrm{d}\Gamma_{\text{decay}}}{\mathrm{d}^3\mathbf{q}} = \frac{g^2}{(2\pi)^3}\frac{1}{2E_{\mathbf{q}}}\Pi^{>}(q^0 = E_{\mathbf{q}}, \mathbf{q}).$$



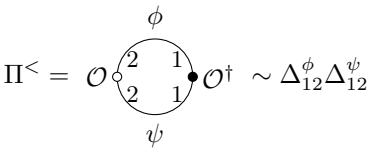
Remark.

(1) The lesser and greater Green's functions are related to the spectral function:

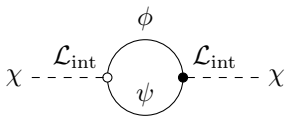
$$\Pi^{<}(E_{\mathbf{q}}, \mathbf{q}) \propto n_{\mathrm{B}}(E_{\mathbf{q}})\rho, \quad \Pi^{>}(E_{\mathbf{q}}) = (1 + n_{\mathrm{B}}(E_{\mathbf{q}}))\rho.$$

This means that at $T \to 0$ the production rate vanishes, while the decay rate reduces to the vacuum decay rate.

(2)



Here, all fields are on-shell because of the thermal parts of the propagators. Similarly, we can think about this as

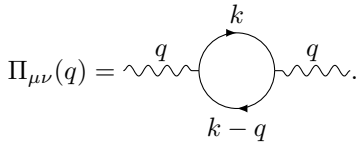


3.9 The hard-thermal loop (HTL) approximation

We will treat the IR divergences in real time (i.e. Minkowski). The propagator $1/(p^2 - m^2 + i\varepsilon)$ has soft and collinear divergences when the propagator goes on-shell. This contrast the Euclidean case where the only divergence occurred at $m = 0, p^\mu = 0$. The calculation of the propagator can be done by analytic continuation from the Euclidean result (for which we need to know the dependence on all ω_m):

$$\Pi_{\mathrm{R}} = \Pi_{\mathrm{E}}(\omega_n \to -i(k + i\varepsilon), \mathbf{k}).$$

IR behaviour is not from a single zero mode as in Euclidean space, but a continuum of small $k^0, \mathbf{k}$. The divergence happens approximately, when the energy and momentum have approximately equal amplitude, $k^0/|\mathbf{k}| \sim \mathcal{O}(1)$. The HTL approximation is valid when external momenta are soft $k^0, |\mathbf{k}| \sim gT \ll T$ and loop momenta are hard $p^0, |\mathbf{p}| \sim T$. For example, the photon HTL self-energy is



This can be calculated by setting the fermion mass to zero. As mentioned before, the retarded Green's function can be obtained by the analytic continuation of the Euclidean one. The latter is expressed with the sum of Matsubara frequencies:

$$T\sum_n \frac{1}{(k_n^2+a^2)((k_n-q_m)^2+b^2)},$$

which can be converted into a sum of terms with Fermi-Dirac distributions $n_{\rm F}(a, b\pm\mu)$ (see homework). In the HTL approximation, the results depend on $q_m, \mathbf{q} \ll k_n, \mathbf{k}$. We carefully expand in q, including the Fermi–Dirac distributions of the form $n_{\rm F}(E_{\mathbf{k}-\mathbf{q}}\pm\mu)$. Define $\mathbf{v}=\mathbf{k}/k$ and $k=|\mathbf{k}|$ and write

$$\int \frac{{\rm d}^3\mathbf{k}}{(2\pi)^3}\frac{1}{2\pi^2} = \int {\rm d}k\, k^2 {\rm d}\Omega_{\mathbf{v}},$$

where $\int {\rm d}\Omega_{\mathbf{v}} = 1$. We can then integrate over ${\rm d}k$ including $n_{\rm F}(E_{\mathbf{k}})$ and $n'_{\rm F}(E_{\mathbf{k}})$ and analytically continue $q_m \to -iq_0$, which gives

$$\Pi_{ij}(q) = m_{\rm T}^2 \int {\rm d}\Omega_{\mathbf{v}} \frac{q^0 v_i v_j}{q^0 - \mathbf{v}\cdot\mathbf{q} + i\varepsilon}.$$

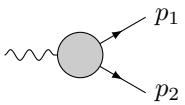
The indices run from 1 to 3, and $m_{\rm T}$ is the thermal mass

$$m_{\rm T}^2 = \frac{g^2}{6}\left((2N+n_{\rm f})T^2 + \frac{3\mu^2}{\pi^2}\right).$$

The fermion self energy becomes

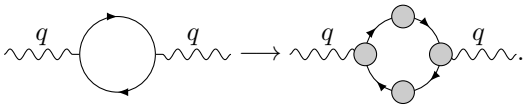
$$\Sigma_{\alpha\beta}(p) = -m_{\rm F}^2 \int {\rm d}\Omega_v \frac{\not{v}}{v\cdot p},$$

where we take $v=(1,\mathbf{v})$. This integral modifies the dispersion relation of particles in the plasma (different polarizations modify the dispersion differently). These different excitations can be identified as quasi-particles, called plasmons. For $q^0 \to 0$ one recovers the Debye screening. The denominator $1/(\mathbf{v}\cdot\mathbf{q}+i\varepsilon)$ can be zero and the $i\varepsilon$ prescription gives an imaginary part, so $\omega(\mathbf{k})$ becomes complex which leads to damping, called *Landau damping*. Three-point functions also contain HTL pieces



$$= -m_{\rm F}^2 \int {\rm d}\Omega_v \frac{v^\mu \not{v}}{(v\cdot p_1)(v\cdot p_2)}, \quad \text{at 1-loop.}$$

This leads to HTL-resummed perturbation theory, where we replace every n-point function by the resummed one:



With this, we can write an HTL effective Lagrangian, for example for QCD at one loop order:

$$\mathcal{L}_{\rm QCD}^{\rm HTL} = -\frac{1}{2}{\rm tr}(G_{\mu\nu}^2) + \frac{m_{\rm T}^2}{2}\int {\rm d}\Omega_v\ {\rm tr}\left(\left[\frac{v^\lambda G_{\lambda\mu}}{v\cdot {\rm D}}\right]\left[\frac{v^\rho G_\rho{}^\mu}{v\cdot {\rm D}}\right]\right) + \overline{\Psi}(i\not{\rm D}-m)\Psi - m_{\rm F}^2\int {\rm d}\Omega_v\, \overline{\Psi}\frac{\not{v}}{iv{\rm D}}\Psi.$$

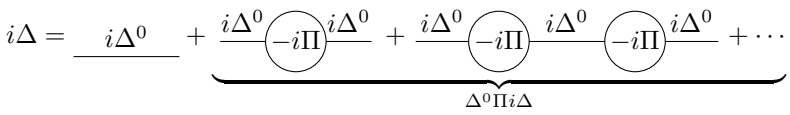
This Lagrangian gives any A^n and $A^{n-2}\overline{\Psi}\Psi$ interactions for $n\geq 2$ from expanding $1/(iv\partial + gA)$ in powers of A. However, for QED (Abelian case), there is only A^2 self interaction from HTL (and all $A^n\overline{\Psi}\Psi$ for $n\geq 0$).

3.10 Out-of-equilibrium kinetic theory from QFT

We start the discussion from the *Boltzmann equation* for the phase space distribution $f(\mathbf{r}, \mathbf{v}, t)$ with a collision term from the reaction $12 \to 1'2'$:

$$\frac{\partial f}{\partial t} + \mathbf{p} \cdot \frac{\partial f}{\partial \mathbf{r}} + \frac{\mathbf{F}}{m} \cdot \frac{\partial f}{\partial \mathbf{v}} = \frac{\partial f}{\partial t}\bigg|_{\text{collisions}} = \int \mathrm{d}^3\mathbf{v}_2 \, \mathrm{d}^3\mathbf{v}_{1'} \, \mathrm{d}^3\mathbf{v}_{2'} (f_{1'} f_{2'} - f f_2) W(\mathbf{v}, \mathbf{v}_2, \mathbf{v}_{1'}, \mathbf{v}_{2'}).$$

The term $f_{1'} f_{2'}$ is called the *gain term* and $f f_2$ the *loss term* (they come from $12 \to 1'2'$ and the reverse reactions). We will not assume thermal equilibrium, so f is not the Bose–Einstein or Fermi–Dirac distribution. We will also change the notation, so that the two point function will be $i\Delta_{ab}$ (instead of Π_{ab}) and the propagator is $i\Delta^0_{ab}$ (instead of Δ). The symbol Π will be reserved for the self-energy. The full propagator can be written in terms of the following Feynman diagrams:



Using $i\Delta^0(-i\Pi)i\Delta = i\Delta(-i\Pi)i\Delta^0$, this can be summed into the *Schwinger–Dyson equation*

$$\Delta_{ab}(x, y) = \Delta^0_{ab}(x, y) - \int_{\mathscr{C}} \mathrm{d}^4 z \int_{\mathscr{C}} \mathrm{d}^4 z' \sum_{a', b' = 1, 2} \Delta^0_{aa'}(x, z) \Pi_{a'b'}(z, z') \Delta_{bb'}(z' y).$$

This is not a perturbative equation, it is exact. $\Pi_{a'b'}$ depends on Δ_{ab} and on the (exact) vertex functions. For the scalar field, we apply $\partial^2_{(x)} - m^2$:

$$\begin{aligned} (-\partial^2_{(x)} - m^2)\Delta(x, y) &= \delta^{(4)}_{\mathscr{C}}(x, y) - \int_{\mathscr{C}} \mathrm{d}^4 z \, \Pi(x, z) \Delta(z, y), \\ (-\partial^2_{(y)} - m^2)\Delta(x, y) &= \delta^{(4)}_{\mathscr{C}}(x, y) - \int_{\mathscr{C}} \mathrm{d}^4 z \, \Delta(x, z) \Pi(z, y). \end{aligned} \tag{$**$}$$

where the contour delta function is defined as

$$\delta^{(4)}_{\mathscr{C}}(x, y) = \begin{cases} \delta^{(4)}(x - y), & x^0, y^0 \text{ on upper contour;} \\ -\delta^{(4)}(x - y), & x^0, y^0 \text{ on lower contour;} \\ 0, & \text{otherwise.} \end{cases}$$

We dropped the CTP (closed time path) indexes for brevity. Now, we write out the CTP sum and take $ab = 12, 21$ to get the collision term:

$$\begin{aligned} \mathcal{C} &\equiv (-\partial^2_{(x)} - m^2 + \Pi_\delta(x))\Delta^{\lessgtr}(x, y) - \int_{\mathscr{C}} \mathrm{d}^4 z \, \left(\Pi^{\mathrm{h}}(x, z) \Delta^{\lessgtr}(z, y) + \Pi^{\lessgtr}(x, z) \Delta^{\mathrm{h}}(z, y) \right) \\ &= \frac{1}{2} \int \mathrm{d}^4 z \left(\Pi^{>}(x, z) \Delta^{<}(z, y) - \Pi^{<}(x, z) \Delta^{>}(z, y) \right). \end{aligned}$$

The term $\mathcal{C}$ is called the *collision term* and the equation is called the *Kadanoff–Baym equation*. The term Π_δ is defined via

$$\Pi(x, y) = \Pi_\delta(x) \delta^{(4)}_{\mathscr{C}}(x, y) + \Pi^{>}(x, y) \theta_{\mathscr{C}}(x^0, y^0) + \Pi^{<}(x, y) \theta_{\mathscr{C}}(y^0, x^0).$$

Next, we go to the *Wigner space* by defining the centre-of-mass coordinate $X = (x+y)/2$ and the relative coordinate $u = x - y$:

$$f(X,p) = \int \mathrm{d}^4 u\, e^{ipu} f(X - u/2, X + u/2).$$

We will assume that the system is homogeneous on scales much larger than the scale particle of particle interactions (mean free path). This allows us to do a *gradient expansion* in ∂_X, because it is small relative to p^μ. Keeping only the leading order in gradients, we have

$$\begin{aligned}\int \mathrm{d}^4 z\, f(x,z)g(z,y) &\to f(X,p)g(X,p),\\ h(x)g(x,y) &\to h(X)g(X,p),\\ h(y)g(x,y) &\to h(X)g(X,p),\\ \partial_{(x)}h(x)g(x,y) &\to (\partial_X h(X))g(X,p) + ih(X)pg(X,p),\end{aligned}$$

The collision becomes

$$(p^2 - m^2 + \Pi_\delta(x) - ip^\mu\partial_\mu)\Delta^{\lessgtr}(X,p) - \left[\Pi^{\mathrm{h}}\Delta^{\lessgtr} + \Pi^{\lessgtr}\Delta^{\mathrm{h}}\right](X,p) = \mathcal{C} = \frac{1}{2}(\Pi^{>}\Delta^{<} - \Pi^{>}\Delta^{<})(X,p).$$

Take the second equation of Eq. (**) and calculate the sum and the difference. The difference gives the *transport equation*

$$p^\mu\partial_\mu\Delta^{\lessgtr}(X,p) = i\mathcal{C} = \frac{i}{2}(\Pi^{>}\Delta^{<} - \Pi^{<}\Delta^{>})(X,p)$$

and the sum gives the *constraint equation*

$$(p^2 - M^2(X))\Delta^{\lessgtr}(X,p) = \left[\Pi^{\mathrm{h}}\Delta^{\lessgtr} + \Pi^{\lessgtr}\Delta^{\mathrm{h}}\right](X,p),$$

where $M(X)^2 = m^2 - \Pi^2(X)$. In the non-equilibrium case, we are very much hopeless to solve these equations exactly. Thus, we will make the *quasi-particle approximation*, where we assume that the spectral function is sharply peaked around the mass shell:

$$\rho(X,p) = \pi\delta(p^2 - M^2(X))(\theta(p^0) - \theta(-p^0)).$$

The phase-space distribution $f(X,p)$ function is defined as

$$\theta(p^0)i\Delta^{<}(X,p) = 2\theta(p^0)\rho(X,p)f(X,p).$$

Using the relations between the different Green's functions, we have

$$\begin{aligned}i\Delta^{<}(X,p) &= 2\pi\delta(p^2 - M^2(X))(\theta(p^0)f(X,p) + \theta(-p^0)(1 + f(X,-p))),\\ i\Delta^{>}(X,p) &= 2\pi\delta(p^2 - M^2(X))(\theta(p^0)(1 + f(X,p)) + \theta(-p^0)f(X,-p)).\end{aligned}$$

In the thermal equilibrium case, $f(X,p) \to n_{\mathrm{B}}(|p^0|)$. In general, X depends on X and $\mathbf{p}$ and $p^0 = E_{\mathbf{p}(X)}$, where $E_{\mathbf{p}(X)} = \sqrt{m^2 + \mathbf{p}^2(X)}$. Now, we can plug these into the transport equation and integrate over p^0, using the following:

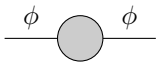
$$iE_{\mathbf{p}(X)}\int \frac{\mathrm{d}p^0}{2\pi} 2\theta(p^0).$$

This gives the *Boltzmann equation*

$$p^\mu\partial_\mu f(X,p)|_{p^0 = E_{\mathbf{p}(X)}} = E_{\mathbf{p}(x)}\int \frac{\mathrm{d}p^0}{2\pi}\theta(p^0)[\Pi^{>}\Delta^{<} - \Pi^{<}\Delta^{>}](X,p).$$

The right-hand side is the QFT expansion of the collision term.

- To calculate it, we need the self-energies $\Pi^{\lessgtr}$ of the scalar ϕ.



  We assume that all other particles are in thermal equilibrium. If the interactions in the bath are weak and the bath particles are in the thermal equilibrium, then we can calculate (at 1-loop) in terms of the equilibrium propagators of the bath particles.

- If the plasma is homogeneous (spatially), there is no $\mathbf{X}$ dependence and the left-hand side of the Boltzmann equation reduces to $E_{\mathbf{p}(X)}\partial_t f(t,\mathbf{p})$.

Chapter 4

Localized, topological field configurations and instantons

When we treated the Higgs field, we assumed that the value of the field is close to the vacuum expectation value, and we only quantized the excitations $\phi(x) = v + \hat{\phi}(x)$. If the minimum of the potential $V(\phi)$ is at $V(\phi_0) = 0$, we can without loss of generality assume $S[\phi_0(x)] = 0$.

In this chapter, we will study the solutions of the classical field equations with finite action $S[\phi_{\text{cl}}] < \infty$ or energy (for static solutions). Such solutions have $S = c/\lambda$, where λ is the coupling constant and c is some number. These solutions are non-perturbative, because they cannot be expanded in powers of λ ($e^{-c/\lambda}$). Why should we care about such solutions? Some effects cancel in perturbation theory to all loop orders and only arise when non-perturbative effects are included. Such solutions are often *topological*. This means that we cannot deform them to $\phi_0(x) = c$ by continuous paths in field space without passing through infinite action configurations. They have conserved quantum numbers which do not aries from symmetries (Noether currents). There are two types of such solutions:

- Static (time-independent) solutions with finite energy: particle-like objects *solitons, (kinks, vortices, monopoles).*
- Solutions to the Euclidean equations of motion with finite action: *instantons, bounces.* These solutions are not particle-like and are relevant to tunnelling in QFT.

4.1 Solitons in one (space) dimension (the kink)

Consider a simple scalar field in 1+1 dimensions with the Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - V(\phi),$$

where

$$V(\phi) = \frac{\lambda}{4}(\phi^2 - v^2)^2 = \text{const} - \frac{1}{2}m^2\phi^2 + \frac{\lambda}{4}\phi^4.$$

The constants are related by $m^2 = \lambda v^2$, so $m_\varphi^2 = \sqrt{2}m^2$ is the mass of the excitations of the φ around the vacuum value $\phi = v + \varphi$. The dimension of the field is $[\phi] = 0$ and the coupling

constant has dimension $[\lambda] = 2$. This means that the theory is super-renormalizable, but this is not important for us now.

4.1.1 The kink solution

Consider only the static solution $\phi(t, x) = \phi(x)$. The Euler–Lagrange equation is

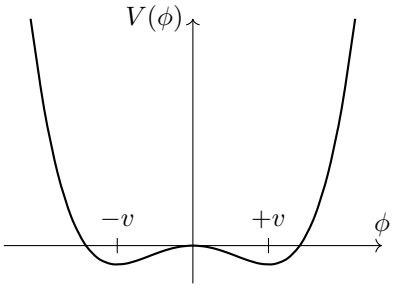
$$\partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} = \frac{\partial \mathcal{L}}{\partial \phi}, \quad \boldsymbol{\partial} \cdot \frac{\partial U}{\partial (\boldsymbol{\partial}\phi)} = \frac{\partial U}{\partial \phi}$$

where

$$U = \frac{1}{2}\left(\frac{\mathrm{d}\phi}{\mathrm{d}x}\right)^2 + V(\phi(x)).$$

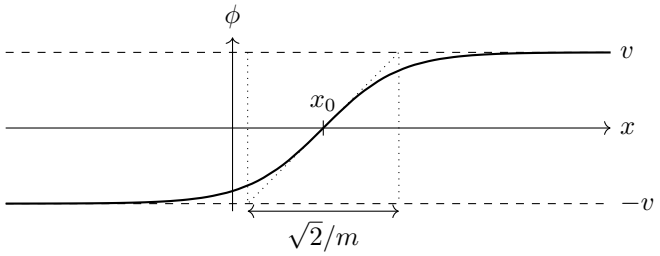
We want to extremize the energy functional

$$E[\phi] = \int_{-\infty}^{\infty} \mathrm{d}x\, U[\phi(x)] = \int_{-\infty}^{\infty} \mathrm{d}x \left(\frac{1}{2}\left(\frac{\mathrm{d}\phi}{\mathrm{d}x}\right)^2 + \frac{\lambda}{4}(\phi^2 - v^2)^2\right).$$



The solutions we want to find need to be different from $\phi_0 = \pm v$, and we want to assure finite energy. The function U has to vanish sufficiently fast when $x \to \pm\infty$. In the limits $x \to \pm\infty$ the field has to approach one of the minima of the potential (the values of the vacuum manifold $\{-v, v\}$). The possible solutions are given in the following table.

$x \to -\infty$	$x \to \infty$	name
$-v$	$-v$	vacuum
v	v	vacuum
$-v$	v	kink
v	$-v$	anti-kink



To explore the properties of the kink solution, we will introduce the scaling $\phi = \hat{\phi}v$ and $y = mx = \sqrt{\lambda}vx$. The energy functional becomes

$$E[\hat{\phi}] = \frac{m^3}{\lambda}\int_{-\infty}^{\infty} \mathrm{d}y \left[\frac{1}{2}\left(\frac{\mathrm{d}\hat{\phi}}{\mathrm{d}y}\right)^2 + \frac{1}{4}(\hat{\phi}^2-1)^2\right].$$

The kink has energy proportional to $E \propto m^3/\lambda$ which is much smaller than m in the limit $\lambda/m^2 \ll 1$. We can make the following estimate:

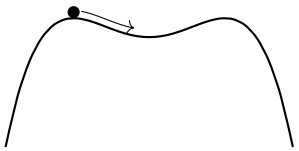
$$E[\phi] = \int_{-\infty}^{\infty} \mathrm{d}x \left[\frac{1}{2}\left(\frac{\mathrm{d}\phi}{\mathrm{d}x} - \sqrt{2V(\phi)}\right)^2 + \sqrt{2V(\phi)}\frac{\mathrm{d}\phi}{\mathrm{d}x}\right] \geq \int_{-v}^{v} \mathrm{d}\phi\sqrt{2V(\phi)} = \frac{2\sqrt{2}}{3}\frac{m^3}{\lambda}.$$

We can find the kink solution in two ways:

(1) Use Euler–Lagrange equations for $E[\phi]$:

$$\frac{\mathrm{d}^2\phi}{\mathrm{d}x^2} = \frac{\mathrm{d}V}{\mathrm{d}\phi}.$$

This describes the "motion" in an inverted potential $-V(\phi)$.



(2) The $E[\phi]$ is minimal for $\mathrm{d}\phi/\mathrm{d}x = \sqrt{2V(\phi)}$ and together with the boundary conditions $\phi(-\infty) = -v$, $\phi(\infty) = v$ the solution has one free parameter x_0:

$$\phi_{\mathrm{k}}(x) = v\tanh(m(x-x_0)/\sqrt{2}).$$

This solution has "size" $1/m$, which means it is localized around x_0, and it approaches exponentially fast to $\pm v$ on either side. The energy of the kink is

$$M = \frac{2\sqrt{2}}{3}\frac{m^3}{\lambda}.$$

This means that the size of the kink is much bigger than its Compton wavelength $1/m \gg 1/M$ (it is called an extended object). The same thing happens with a proton, which is much larger than its Compton wavelength.

Let's discuss some further properties of the kink.

- Stability of the kink solution. We perturb around the kink: $\phi(t,x) = \phi_{\mathrm{k}}(x) + \varphi(t,x)$. Plugging this into the equation of motion gives

$$\frac{\partial^2\varphi}{\partial x^2} = -\frac{\partial V^2}{\partial\phi^2}\bigg|_{\phi_{\mathrm{k}}}\varphi \equiv -W\varphi.$$

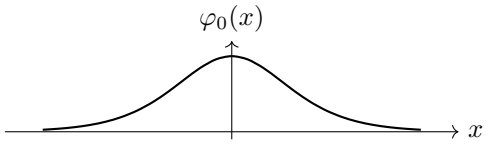
Since the proportionality factor is time-independent, we can use the Fourier ansatz

$$\varphi(t,x) = e^{i\omega t}\varphi_\omega(x).$$

The equation that follows is

$$-\frac{\mathrm{d}^2\varphi_\omega}{\mathrm{d}x^2} + W\varphi_\omega = \omega^2\varphi_\omega.$$

The stability of the kink requires that all eigenvalues ω^2 are non-negative. Since the kink is time independent, there is an eigenfunction with eigenvalue $\omega^2 = 0$, $\varphi_0(x) = \mathrm{d}\phi_\mathrm{k}(x)/\mathrm{d}x$. This solution is strictly positive and from the "node theorem" we know that $\varphi_0(x)$ must be the eigenfunction to the lowest eigenvalue, so all eigenvalues are non-negative and the kink is stable.

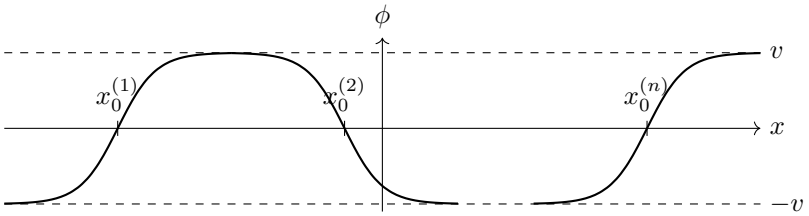


- If the action S is Lorentz invariant, we can produce "moving kinks" by Lorentz transformations of $\phi_\mathrm{k}(x)$:

$$\phi(t,x) = v\tanh\left(\frac{m}{\sqrt{2}}\frac{x - vt - x_0}{\sqrt{1-v^2}}\right),$$

where v is the velocity. In this case, $E = M_\mathrm{k}/\sqrt{1-v^2} = \sqrt{M_\mathrm{k}^2 + \mathbf{p}^2}$. Hence, it behaves like a relativistic particle with mass M_k (but it is not point-like).

- Multi-kink and antikink configurations exist, but they are not stable. Exact solutions exist only if their locations are infinitely far apart (*dilute instanton gas approximation*) $|x_0^{(i)} - x_0^{(i-1)}| \to \infty$. In this limit, their energies are additive. The kinks attract each other and annihilate.



- Kinks can be thought of as topological objects.

  (1) We cannot deform kink continuously to the standard vacuum solution $\phi_0(x) = \pm v$ with only going through finite energy configurations. For every finite energy configuration it holds $\phi(\infty) - \phi(-\infty) = 2vn$ for $n \in \{-1, 0, 1\}$.

  (2) We can define a conserved topological current (not from a Noether symmetry):

$$j_\mu \equiv \varepsilon_{\mu\nu}\partial^\nu\phi, \qquad \Longrightarrow \qquad \partial^\mu j_\mu = 0.$$

  The associated topological charge is

$$Q = \int_{-\infty}^{\infty}\mathrm{d}x\, j^0 = \int_{-\infty}^{\infty}\mathrm{d}x\,\frac{\mathrm{d}\phi}{\mathrm{d}x} = \phi(\infty) - \phi(-\infty) = 2vn.$$

4.1.2 Quantum corrections

So far we considered only classical solutions. To make sense out of them in the QFT as saddle/stationary point of the path integral, the quantum fluctuations must be small (we will justify this a posteriori):

$$\phi(t,x) = \phi_{\mathrm{k}}(x) + \varphi(t,x).$$

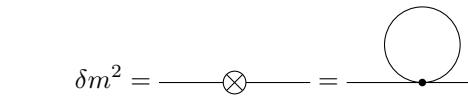
We expand the Lagrangian to second order in φ:

$$L = \int \mathrm{d}x \mathcal{L} = \int \mathrm{d}x \left(-\frac{1}{2} \left(\frac{\mathrm{d}\phi_{\mathrm{cl}}}{\mathrm{d}x} \right)^2 - V(\phi_{\mathrm{cl}}) \right) + \int \mathrm{d}x \left(\frac{1}{2}(\partial_\mu \varphi)(\partial^\mu \varphi) - \frac{1}{2} W(\phi_{\mathrm{cl}}) \varphi^2 \right) + \mathcal{O}(\varphi^3).$$

The linear terms in φ vanish, since classical solution is a stationary point. Plugging this into the path integral, we get a correction in the form of a determinant (for the quadratic correction), but here we will only focus on the correction to the kink mass. The theory is super-renormalizable, only the mass term has to be corrected (it is logarithmically divergent):

$$\mathcal{L} \to \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - V(\phi) + \mathcal{L}_{\mathrm{ct}}.$$

The counter-term Lagrangian is simply $\delta m^2 \phi^2/2$ and comes exclusively from the one-loop snail (tadpole) diagram. We choose the renormalization where the counter-term corresponds to the whole diagram:

$\delta m^2 =$  $= -3i\lambda \int \frac{\mathrm{d}k^2}{(2\pi)^2} \frac{1}{k^2 - 2m^2 + i\varepsilon} = \frac{3\lambda}{2\pi} \int_0^\infty \frac{\mathrm{d}k}{\sqrt{k^2 + 2m^2}}.$

The quantum correction for the standard vacuum is

$$E_0 = E_{\mathrm{cl}}(\phi_0) + \frac{1}{2} \sum_i \omega_i + E_{\mathrm{ct}}(\phi_0) + \delta E_{\mathrm{vac}}.$$

The second term corresponds to the zero-point energies, the third term is evaluated as

$$E_{\mathrm{ct}} = -\frac{1}{2} \delta m^2 \int \mathrm{d}x \, \phi_0^2$$

and the last term is the vacuum energy (which is the same for all vacuum solutions). The same relation holds for the kink solution, just the frequencies are different and the counter-term has to be evaluated on the kink solution. The actual correction to the kink mass will be

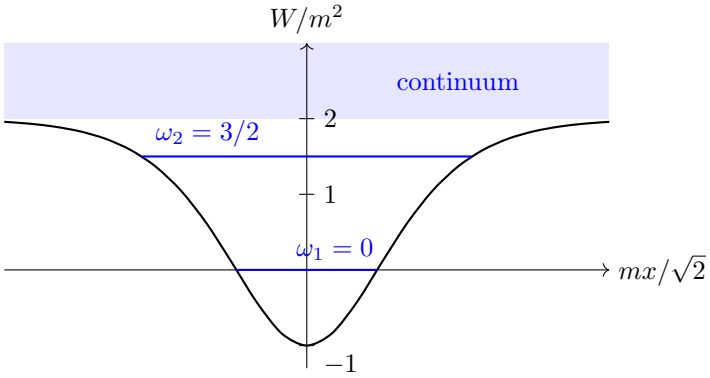
$$M_{\mathrm{kink}} = E_{\mathrm{kink}} - E_0 = E_{\mathrm{cl}}(\phi_{\mathrm{k}}) + \frac{1}{2} \sum_i (\omega_i^{\mathrm{kink}} - \omega_i^{\mathrm{vac}}) - \frac{1}{2} \delta m^2 \int_{-\infty}^{\infty} \mathrm{d}x \, (\phi_{\mathrm{k}}^2 - \phi_0^2).$$

The last integral evaluates to $-2\sqrt{2}m/\lambda$, the hard part is figuring out the kink eigenvalues

$$\left(-\frac{\mathrm{d}^2}{\mathrm{d}x^2} + W(\phi_{\mathrm{k}}) \right) \varphi_\omega(x) = \omega^2 \varphi_\omega(x).$$

The operator is

$$W(x) = -m^2 + 3\lambda \phi_{\mathrm{k}}^2 = m^2 \left(2 - \frac{3}{\cosh^2(m(x - x_0)/\sqrt{2})} \right).$$



There are two eigenvalues in the discrete spectrum $0, 3m^2/2$ and the continuum $\omega^2 \geq 2m^2$. Taking the continuum limit (very difficult and subtle, see homework), one obtains

$$\frac{1}{2}\sum_i \omega_i^{\text{kink}} - \frac{1}{2}\sum_i \omega_i^{\text{vac}} = \left(\frac{\sqrt{6}}{12} - \frac{3\sqrt{2}}{2\pi}\right) m - \frac{3\sqrt{2}}{2\pi}\int_0^\infty \frac{\mathrm{d}k}{\sqrt{k^2+2m^2}}.$$

The second (divergent) term cancels exactly the counter-term, so we get

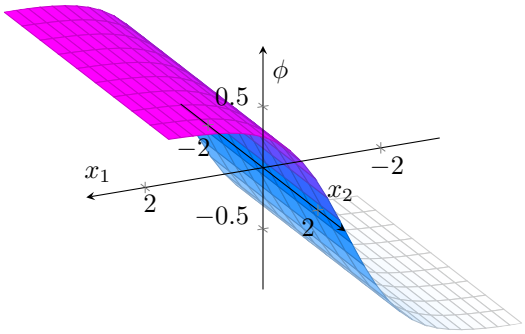
$$M_{\text{kink}} = \frac{2}{\sqrt{2}}\frac{m^3}{\lambda} + \left(\frac{\sqrt{6}}{12} - \frac{3\sqrt{2}}{2\pi}\right) m + \mathcal{O}(\lambda/m).$$

The correction is small when $\lambda/m^2 \ll 1$, which corresponds to the "weak coupling".

4.1.3 Derrick's theorem

The kink solution can be readily extended in d dimensions ($d = D + 1$):

$$\phi(t, x_1, \ldots, x_{d-1}) = \phi_k(x_1).$$



In two spatial dimensions, the kink at $x_1 = 0$ extends on the whole line of x_2 (kink line). In $d = 4$, the extended kink surface is called the *domain wall*. They are important in cosmology, because they can carry energy. However, they carry only a finite energy density and hence infinite total energy, which is not the solution we are after.

Theorem 4.1 (*Derrick*). *There are no static finite-energy solutions in $d > 1+1$ Minkowski space for scalar field theories with Lagrangian*

$$S[\phi] = \int \mathrm{d}^d x \left(\frac{1}{2} g_{ab}(\phi)(\partial_\mu \phi_a)(\partial^\mu \phi_b) - V(\phi) \right),$$

where g_{ab} is a positive-definite metric in field space and $V(\phi) \geq 0$.

Proof. Without loss of generality, we can assume that the potential vanishes at the vacuum $V(\phi_{\min}) = 0$. For any finite energy solution it will hold $\phi_a(x) \xrightarrow{|\mathbf{x}|\to\infty} \phi_{\min}$. The energy functional is

$$E[\phi] = T[\phi] + W[\phi] = \frac{1}{2} \int \mathrm{d}^D x \, g_{ab}(\phi) \nabla\phi_a \cdot \nabla\phi_b + \int \mathrm{d}^D x \, V(\phi).$$

The kinetic and potential part are separately non-negative. Suppose that $\phi(\mathbf{x})$ is a static finite-energy solution. Define $\phi_\lambda(\mathbf{x}) = \phi(\lambda\mathbf{x})$ for $\lambda > 0$. Changing the variable $\mathbf{x} \to \lambda\mathbf{x}$ in the energy functional gives

$$E[\phi_\lambda] = \lambda^{2-D} T[\phi] + \lambda^{-D} W[\phi].$$

We know that $\lambda = 1$ is the stationary point of the energy functional, so by differentiating with respect to λ and setting $\lambda = 1$ we get

$$(D-2)T[\phi] + DW[\phi] = 0.$$

- For $D = 1$, $T[\phi] = W[\phi]$ is satisfied by the virial theorem.
- For $D = 2$, $W[\phi] = 0$ and hence $\phi = \phi_{\min}$ everywhere (standard vacuum solution).
- In the case $D > 2$, we get $T[\phi] = W[\phi] = 0$, so again we only get the standard vacuum solutions.

□

Remark. We assumed that the kinetic term only has two derivatives. We could add an unconventional term T' with four (or six) derivatives:

$$(D-4)T'[\phi] + (D-2)T[\phi] + DW[\phi] = 0$$

For $d = 4$ ($D = 3$) this has a non-trivial solution (*skyrmions*). However, fourth (or higher) derivative terms lead to poles in the spectral function which is bad QFT-wise.

It seems as the end of the section (even including fermions doesn't work because fermions cannot have non-zero vacuum expectation values). However, there is a way out: the gauge fields:

$$S[\phi, A] = \int \mathrm{d}^d x \left(-\frac{1}{2} \operatorname{tr}(\mathbf{F}_{\mu\nu}\mathbf{F}^{\mu\nu}) + \frac{1}{2} g_{ab}(\phi)(\mathrm{D}_\mu \phi_a)^\dagger (\mathrm{D}^\mu \phi_b) - V(\phi) \right).$$

We call the three terms T_A, T_ϕ and W. We define the scaling

$$\phi_\lambda(\mathbf{x}) = \phi(\lambda\mathbf{x}), \quad \mathbf{A}^\mu_\lambda(x) = \lambda \mathbf{A}^\mu(\lambda\mathbf{x}), \quad \mathrm{D}^\mu_\lambda = \lambda(\partial_\mu - igA^\mu_\lambda(\lambda\mathbf{x})).$$

The condition from the Derrick's theorem becomes

$$0 = (D-4)T_A[\mathbf{A}] + (D-2)T_\phi[\phi, \mathbf{A}] + DW[\phi].$$

The solitons are allowed for $D = 2, 3$, but none for $D \geq 5$. For $D = 4$, T_ϕ and W have to vanish, but pure Yang–Mills theory is still possible.

4.2 Localized, static, topological, finite energy, classical field configurations in more than one space dimension

4.2.1 Vacuum manifold and finite-energy solutions

Let's start with a (continuous, Lie) gauge group G and a scalar field ϕ which scales linearly in the representation of G. The Lagrangian is

$$\mathcal{L} = -\frac{1}{2}\operatorname{tr}(\mathbf{F}_{\mu\nu}\mathbf{F}^{\mu\nu}) + \frac{1}{2}(\mathrm{D}_\mu\phi)^\dagger(\mathrm{D}^\mu\phi) - V(\phi).$$

In this chapter, we will use the bold notation for the gauge fields $\mathbf{A}_\mu = A^a_\mu T^a$ and the field strength tensor $\mathbf{F}_{\mu\nu} = F^a_{\mu\nu}T^a$, but on any other object (like $\mathbf{x}$) bold will still denote vectors in space. The vacuum manifold $\mathcal{M}_0$ is defined as the set of field values which minimize the potential. If $\phi_{\min} \neq 0$ minimizes $V(\phi)$, then so does $\mathcal{U}(g)\phi_{\min}$ for all $g \in G$. If $\phi_{\min}$ breaks G completely, then all $\mathcal{U}(g)\phi_{\min}$ correspond to different vacua and the vacuum manifold is isomorphic to the group manifold G. When there is an unbroken subgroup $H \subset G$, then $\mathcal{U}(h)\phi_{\min} = \phi_{\min}$ and $\mathcal{U}(gh)\phi_{\min} = \mathcal{U}(g)\phi_{\min}$. This means that the vacua form equivalence classes $\mathcal{U}(g)\phi_{\min} \sim \mathcal{U}(g')\phi_{\min}$ if there exists $h \in H$ such that $g' = gh$. The vacuum manifold is then isomorphic to the coset space $\mathcal{M}_0 \cong G/H$.

Example 3. Consider the potential $V(\phi) = \lambda(\phi^\dagger\phi - v^2)^2/4$.

- If ϕ is a complex scalar, then $G = \mathrm{U}(1)$ and the minimizing field is $\phi_{\min} = ve^{i\varphi}$. The unbroken subgroup is trivial $H = \{e\}$, so $\mathcal{M}_0 \cong \mathrm{U}(1) \cong S^1$.
- Let $\boldsymbol{\phi}$ be a real triplet and $G = \mathrm{SO}(3)$. The vacuum manifold is $\mathcal{M}_0 \cong S^2$. We can see this in two ways. From $V(\phi)$ we see the solution is $\boldsymbol{\phi} = v\hat{\mathbf{n}}$ ($\boldsymbol{\phi}^2 = v^2$). The possible values of $\hat{\mathbf{n}}(\theta, \varphi)$ parametrize a sphere S^2. Alternatively, the unbroken subgroup is $H = \mathrm{SO}(2)$ (rotations around the direction of $\boldsymbol{\phi}_{\min}$), so $\mathcal{M}_0 \cong \mathrm{SO}(3)/\mathrm{SO}(2) \cong S^2$.

Next, we will discuss the asymptotic behaviour of finite-energy solutions. For a scalar field, we know it has to approach any of the vacuum field configurations (in G/H) sufficiently fast when $|\mathbf{x}| \to \infty$. It can be parametrized as

$$\phi(\mathbf{x}) \approx \frac{1}{r^a} f(\mathbf{n}), \quad \mathbf{n} = \frac{\mathbf{x}}{|\mathbf{x}|},$$

where $f(\mathbf{n})$ is defined on a sphere S^{D-1}. Scalar field configurations with finite energy can be classified by the mappings

$$S^{D-1} \to \mathcal{M}_0 \cong G/H.$$

If $\phi \sim r^{-0}$, then $\nabla\phi \sim r^{-1}$ and the energy will diverge for $D > 0$ as expected from Derrick's theorem. But $\mathbf{D}\phi \sim \boldsymbol{\partial}\phi - ig\mathbf{A}\phi$ can be better behaved if we have a non-zero gauge field $\mathbf{A}$ that cancels r^{-1} behaviour.

The most common gauge that is used to study such configurations is the temporal gauge $A^0 = 0$ (the physics should be gauge invariant, so we just choose the most convenient one). In this gauge, it holds $\mathrm{D}_0\phi = \partial_t\phi \equiv 0$ for static solutions. Similarly, $\mathbf{F}_{0i} = \partial_t\mathbf{A}_i \equiv 0$. The energy functional simplifies to the Euclidean case:

$$E[\phi, \mathbf{A}] = \int \mathrm{d}^D x \left[\frac{1}{2}\operatorname{tr}(\mathbf{F}_{ij}\mathbf{F}_{ij}) + (\mathrm{D}_i\phi)^\dagger(\mathrm{D}_i\phi) + V(\phi)\right].$$

To ensure finite energy, the field strength tensor has to vanish asymptotically $\mathbf{F}_{ij} \xrightarrow{r\to\infty} 0$. This means that the gauge field is asymptotically a pure gauge:

$$\mathbf{A}_i(\mathbf{x}) \xrightarrow{r\to\infty} \frac{i}{g}(\partial_i \mathcal{U}(\mathbf{x}))\mathcal{U}(\mathbf{x})^{-1},$$

where $\mathcal{U}(\mathbf{x}) \in G$. These configurations can again be classified by the mappings $S^{D-1} \to G$. Of course, we also need to assure that $\mathbf{F}_{ij}$ approaches zero fast enough, but it happens that the condition that $\mathbf{A}$ cancels the leading r^{-1} behaviour of $\partial_i \phi$ is sufficient.

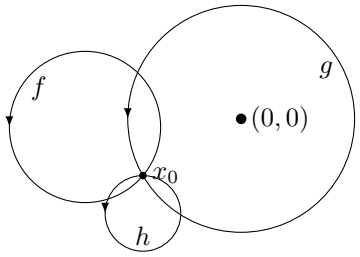
The topological classification of these mappings is presented in the following list.

- $D = 1$: the mapping is $\mathbb{Z}_2 \to \mathbb{Z}_2$ (equivalently $\{-\infty, \infty\} \to \{-v, +v\}$). This gives four distinct elements in the space of maps, all of which are topologically different: two standard vacua, kink and antikink.
- $D = 2$: choose $G = \mathrm{U}(1) \equiv S^1$. Topological configurations are related to maps $S^1 \to S^1$ that cannot be continuously deformed into each other. These maps are classified by the winding number $n \in \mathbb{Z}$, which counts how many times the domain circle wraps around the target circle. The mathematics behind this is the *homotopy theory*.

4.2.2 Homotopy

Consider closed paths f, g on a manifold M, starting at some point $x_0 \in M$. This means that $f, g : S^1 \to M$ and $f(0) = g(0) = x_0$. We say that f and g are *homotopic* (denoted $f \sim g$) if there exists a continuous deformation $K : S^1 \times [0,1] \to M$ such that $K(\theta, t = 0) = f(\theta)$ and $K(\theta, t = 1) = g(\theta)$.

Example 4. The simplest example is $M = \mathbb{R}^2 \setminus \{(0,0)\}$.



We see that f is homotopic to h but not to g.

Homotopy defines an equivalence relation on the space of closed paths on M. The set of equivalence classes is called the *fundamental group* or the *first homotopy group* $\pi_1(M, x_0)$. The identity element corresponds to a constant path and the group product is the concatenation of two paths. If M is connected, then the choice of x_0 is irrelevant, and we write simply $\pi_1(M)$. For simply connected manifolds (like $\mathbb{R}^n$, S^n for $n \geq 2$) it holds $\pi_1(M) \cong \{e\}$.

Space M	$\Pi_1(M)$
$\mathbb{R}^n$	$\{e\}$
S^1	$\mathbb{Z}$
S^n, $n \geq 2$	$\{e\}$
$\mathbb{R}^2 \setminus \{0\}$	$\mathbb{Z}$
$\mathrm{SO}(N)$, $N \geq 3$	$\mathbb{Z}_2$

We can generalize the homotopy to higher dimensions, to the *n-th homotopy group* $\pi_n(M)$, which classifies maps $S^n \to M$. This is an extremely technical mathematical subject, and we will only state one result: $\pi_n(S^n) \cong \mathbb{Z}$. This result is important for the classification of topological solitons in higher dimensions: the maps from S^{D-1} (spatial infinity) to S^{D-1} (vacuum manifold). Some other results are:

$$\begin{aligned}
\pi_n(S_1) &= 0 \quad n \neq 1, \\
\pi_n(S^m) &= 0 \quad m > n, \\
\pi_2(G) &= 0 \quad \text{any compact connected Lie group}, \\
\pi_3(G) &= \mathbb{Z} \quad \text{any compact connected simple Lie group}.
\end{aligned}$$

4.2.3 Vortices in $d = 1 + 2$ dimensions

Consider $D = 2$ Abelian Higgs model

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (\mathrm{D}_\mu\phi)^\dagger(\mathrm{D}^\mu\phi) - V(\phi), \quad V(\phi) = \frac{\lambda}{4}(\phi^\dagger\phi - v^2)^2.$$

The vacuum manifold is $\mathcal{M}_0 \cong G = \mathrm{U}(1) \cong S^1$ and topological field configurations are classified by the $\pi_1(S^1) = \mathbb{Z}$. The standard representative for the winding number n is the map $\theta \to e^{in\theta}$. The winding number of the field configuration $\phi(\mathbf{x}) \xrightarrow{r\to\infty} v\mathcal{U}(\theta)$ can be defined using the following integral representation

$$n = \frac{1}{2\pi i}\int_0^{2\pi} \mathrm{d}\theta\, \mathcal{U}^*(\theta)\partial_\theta\mathcal{U}(\theta) = \frac{1}{2\pi i}\frac{1}{v^2}\oint\limits_{r\to\infty} \mathrm{d}\mathbf{r}(\phi^*\nabla\phi)(\mathbf{x}).$$

For $\mathcal{U}_n(\theta) = e^{in\theta}$ the expression can be easily checked. We also need to check that the integral remains unchanged under small (continuous) deformations of $\mathcal{U} \to \mathcal{U} + \delta\mathcal{U}$. We get (using $\delta\mathcal{U}^* = -\mathcal{U}^{*2}\delta\mathcal{U}$ and $\delta\mathcal{U}(0) = \delta\mathcal{U}(2\pi) = 0$)

$$\delta n = \frac{1}{2\pi i}\int_0^{2\pi} \mathrm{d}\theta(\delta\mathcal{U}^*\partial_\theta\mathcal{U} + \mathcal{U}^*\partial_\theta\delta\mathcal{U}) = \frac{1}{2\pi i}\int_0^{2\pi} \mathrm{d}\theta\, \partial_\theta(\mathcal{U}^*\delta\mathcal{U}) = 0.$$

Let's look at some general properties of field configurations with winding number n. We start with the standard vacuum $\phi_0 = v$ and $A_0^\mu = 0$. Then we apply a *large* gauge transformation ($\mathcal{U}(r,\theta) \xrightarrow{r\to\infty} \mathcal{U}_n(\theta)$ for $\theta \neq 0$ which is not connected smoothly to $\mathcal{U} = 1$). We get $\phi(r,\theta) = \mathcal{U}(r,\theta)\phi_0$ and $A^\mu(r,\theta) = i\mathcal{U}\partial^\mu\mathcal{U}^\dagger/g$. Note that

$$A_i|_{r\to\infty} = \frac{i}{g}e^{in\theta}\mathbf{e}_\theta\frac{1}{r}\frac{\partial}{\partial\theta}e^{-in\theta} = \frac{n}{g}\frac{\mathbf{e}_\theta}{r}.$$

This forms a *vortex*. The minimum of the energy functional with $n = 1$ happens for $\lambda = e^2/2$ (when masses of the gauge boson and Higgs particle are equal) and is called the *Nielson–Olesen vortex*, but it doesn't have an analytical solution. A general ansatz is

$$\phi(r,\theta) = vf(r)\mathcal{U}_n(\theta), \quad A_i(r,\theta) = \frac{i}{g}a(r)\mathcal{U}_n^{-1}(\theta)\partial_i\mathcal{U}_n(\theta)$$

with the following boundary conditions

$$f(\infty) = a(\infty) = 1, \quad f(0) = a(0) = 0.$$

The last condition ensures that the fields are regular and single valued around $r = 0$. If we think about the plane as embedded in $\mathbb{R}^3$, then A_i generates a magnetic field in the perpendicular direction

$$\mathbf{B} = \nabla \times \mathbf{A} = \left(\frac{\partial}{\partial r}(rA_\theta) - \frac{\partial}{\partial \theta} A_r \right) \frac{\mathbf{e}_z}{r} = \frac{n}{g} a'(r) \frac{\mathbf{e}_z}{r}.$$

Magnetic flux through the plane is

$$\Phi_m = \int \mathrm{d}\mathbf{S} \cdot \mathbf{B} = \lim_{r\to\infty} \oint_{r\to\infty} \mathrm{d}\mathbf{r} \cdot \mathbf{A} = \frac{i}{g} \lim_{r\to\infty} a(r) \int_0^{2\pi} \mathrm{d}\theta \, \mathcal{U}^*(\theta) \partial_\theta \mathcal{U}(\theta) = \frac{2\pi}{g} n.$$

This means that the magnetic flux is quantized!

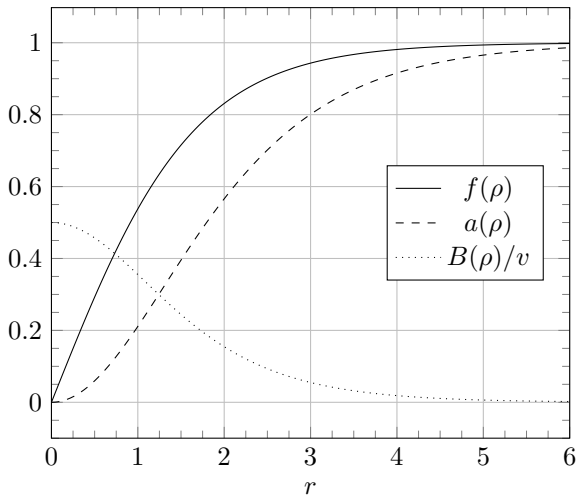
To get a vortex solution for $n = 1$ we minimize the energy functional. Since we are working in Abelian theory, the bold notation on A will again mean vectors in space. The energy functional is

$$E[\phi, \mathbf{A}] = \int \mathrm{d}^2 x \left(\frac{1}{2} \mathbf{B}^2 + |\nabla \phi - ig\mathbf{A}\phi|^2 + V(\phi) \right).$$

This is a coupled system of PDE's for $f(r)$ and $a(r)$.

- We define $\rho = grv$, and we get $E = \pi v^2 \cdot e(\lambda/g^2)$, where $e(x)$ is a dimensionless function. The first factor sets the energy scale (m^2/λ) where $\sqrt{2}m$ is the mass of the scalar excitation in the standard vacuum, and the second is a dimensionless function of the coupling constants.
- It always holds $E \geq \pi v^2 |n|$. This is called the *Bogomolny bound*.
- Qualitatively, the solutions for $f(\rho)$, $a(\rho)$ and $B(\rho)$ are shown in the following graph (for the minimal energy solution $\lambda = g^2/2$).

Nielson–Olesen vortex ($n = 1$, $\lambda = g^2/2$)



At large r both $f(r)$ and $a(r)$ approach 1 exponentially fast. At small r, it holds $f(r) \sim r$ and $a(r) \sim r^2$. The magnetic field $B(r)$ is localized around $r = 0$ at the radius $1/(gv)$.

- These types of topological solutions are used in Type II superconductivity, the Meissner effect and Abrikosov vortices.

4.2.4 Magnetic monopoles in $d = 3 + 1$

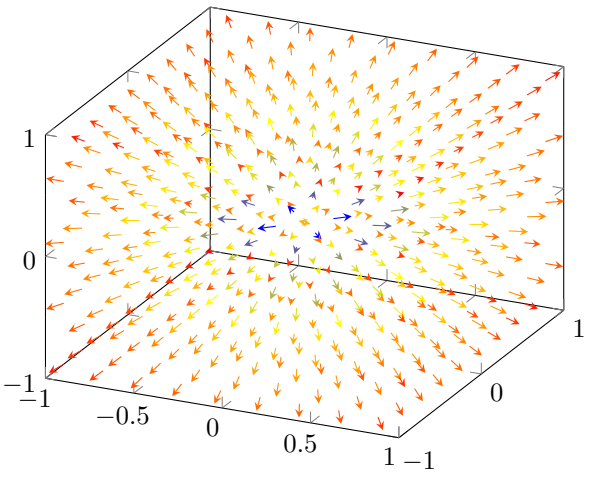
In $D = 3$ the spatial infinity is S^2, so we need a vacuum manifold with non-trivial $\pi_2(\mathcal{M}_0) = \pi_2(G/H)$. The example we will study is the *Georgi–Glashow model* with $G = \mathrm{SU}(2)$ gauge group:

$$\mathcal{L} = -\frac{1}{2}\,\mathrm{tr}(\mathbf{F}_{\mu\nu}\mathbf{F}^{\mu\nu}) + \frac{1}{2}(\mathrm{D}_\mu\phi_i)^\dagger(\mathrm{D}^\mu\phi_i) - \frac{\lambda}{4}(\phi_i\phi_i - v^2)^2.$$

The field ϕ_i is real-valued and in the adjoint representation, so we can represent it as a three component vector $\boldsymbol{\phi}$. The symmetry is broken down to $\mathrm{SO}(2) \cong \mathrm{U}(1)$. To find the vacuum manifold, first note that SU(2) can be parametrized as

$$\begin{pmatrix} a + ia' & b + ib' \\ -b - ib' & a + ia' \end{pmatrix}, \quad |a + ia'|^2 + |b + ib'|^2 = 1, \quad a^2 + b^2 + a'^2 + b'^2 = 1.$$

This represents S^3, so the vacuum manifold is $S^3/S^1 \cong S^2$. The second homotopy group is non-trivial $\pi_2(S^2) \cong \mathbb{Z}$, so topologically distinct finite energy solitons are characterized by an integer. An alternative way to find the vacuum manifold is to look at the field at large r: $\phi_i \to v n_i$. The normals n_i parametrize S^2 and the field configuration is referred to as *hedgehog configuration*.



In general, it holds $\phi(\mathbf{x}) \to v\hat{\phi}_a(\phi, \theta)$, and we can define the winding number in the same way as for the kink:

$$n = \frac{1}{8\pi}\varepsilon_{ijk}\int_{S^2} \mathrm{d}S_i\,\varepsilon_{abc}\hat{\phi}_a(\partial_j\hat{\phi}_b)(\partial_k\hat{\phi}_c).$$

The indices i, j, k are spatial and a, b, c are adjoint. A standard representation for the equivalence class with winding number n is $\hat{\phi}_a(\theta, \varphi) = n_a(\theta, n\varphi)$. The above integral is invariant under small changes $\hat{\phi}_a \to \hat{\phi}_a + \delta\hat{\phi}_a$. The topological current for the conserved topological charge is

$$j^\mu_{\mathrm{top}} = \frac{1}{8\pi}\varepsilon^{\mu\nu\rho\sigma}\varepsilon_{abc}\partial_\nu\hat{\phi}_a\partial_\rho\hat{\phi}_b\partial_\sigma\hat{\phi}_c.$$

The charge is conserved ($\partial_\mu j^\mu_{\mathrm{top}} = 0$), so

$$Q = \int \mathrm{d}^3x\, j^0_{\mathrm{top}} = \frac{1}{8\pi}\varepsilon_{ijk}\varepsilon_{abc}\int \mathrm{d}^3x\partial_i(\hat{\phi}_a\partial_j\hat{\phi}_b\partial_k\hat{\phi}_c) = n.$$

Suppose we choose $v\hat{\phi}_a(\theta,\varphi)$ in a particular direction in the field space ($a=3$). For $E[\phi,A^\mu]$ to be finite, the derivative term $|\mathrm{D}_i\phi_a|^2$ has to integrate to something finite, so $\mathrm{D}_i\phi_a \xrightarrow{r\to\infty} r^{-s}$ with $s>3/2$. Hence,

$$0=\frac{1}{8\pi}\varepsilon_{ijk}\int \mathrm{d}S_i\,\hat{\phi}\cdot(\mathrm{D}_j\hat{\phi}\times\mathrm{D}_k\hat{\phi})=\frac{1}{8\pi}\varepsilon_{ijk}\int \mathrm{d}S_i\left(\hat{\phi}\cdot(\partial_j\hat{\phi}\times\partial_k\hat{\phi})+e\hat{\phi}\cdot\hat{F}_{jk}\right).$$

The triple and scalar products are meant in the internal (adjoint) space. This implies the winding number is

$$n=\frac{g}{4\pi}\int \mathrm{d}S_i\,\hat{\phi}_a B_i^a,$$

where we used the relation $B_i^a=-\varepsilon_{ijk}F_{jk}^a/2$. Now take $\hat{\phi}_a=\delta_{a3}$ so that A_μ^3 corresponds to the "photon field" of the unbroken U(1) – $B_i=\hat{\phi}_a B_i^a$ is the U(1) magnetic field. Next, we write down the Gauss' law for the electromagnetism from the winding number expression ($n=1$):

$$Q_{\mathrm{M}}=\frac{4\pi}{g}=\int\limits_{r\to\infty}\mathrm{d}\mathbf{S}\cdot\mathbf{B}=\int \mathrm{d}^3\mathbf{x}\,\nabla\cdot\mathbf{B}.$$

Define

$$F_{\mu\nu}=\hat{\phi}_a F_{\mu\nu}^a-\frac{1}{g}\varepsilon_{abc}\hat{\phi}_a(\mathrm{D}_\mu\hat{\phi})_b(\mathrm{D}^\mu\hat{\phi})_c$$

so that for $\hat{\phi}_a=\delta_{a3}$, the field strength tensor $F_{\mu\nu}=\partial_\mu A_\nu^3-\partial_\nu A_\mu^3$ corresponds to the unbroken U(1) gauge field. Its dual's derivative gives exactly the topological current:

$$\partial_\mu\tilde{F}^{\mu\nu}=\frac{4\pi}{g}j^\nu_{\mathrm{top}}.$$

Let's look for the monopole solutions of the classical field equations with finite energy and $n=1$. The ansatz is

$$\phi_a(\mathbf{x})=vf(r)\frac{x_a}{r},\quad A_i^a(\mathbf{x})=a(r)\varepsilon_{aij}\frac{1}{g}\frac{x_j}{r^2}$$

where $f(\infty)=a(\infty)=1$ and $f(0)=a(0)=0$. To justify the condition $a(\infty)=1$ and the ansatz for A we go back to the requirement for the behaviour of the covariant derivative at infinity:

$$0=(\mathrm{D}_i\phi_a)|_{r\to\infty}=v\left(\partial_i\frac{x_a}{r}+g\varepsilon_{abc}A_i^b\frac{x_c}{r}\right)\Bigg|_{r\to\infty}.$$

When we multiply by $\varepsilon_{jda}rx_j$ we get

$$\varepsilon_{dij}x_j+g(x_dx_jA_i^j-r^2A_i^d)=0.$$

The ansatz solves this equation. There is no closed-form solution for $f(r)$ and $a(r)$, they obey the following system of coupled non-linear differential equations:

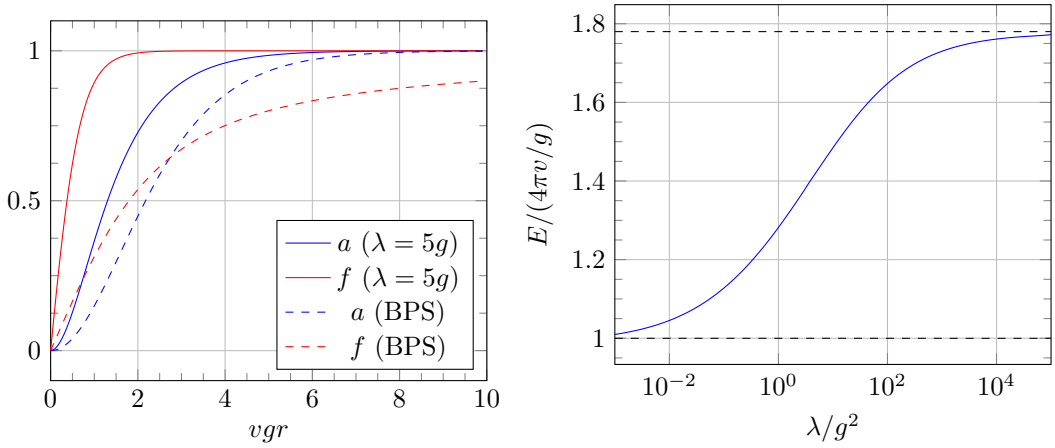
$$2f^2(\rho)(a(\rho)-1)-2a''(\rho)+\frac{2}{\rho^2}a(\rho)(2-a(\rho))(1-a(\rho))=0$$

$$2f(\rho)(a(\rho)-1)^2-(\rho^2f'(\rho))'+\rho^2\frac{\lambda}{g^2}f(\rho)(f^2(\rho)-1)=0$$

where $\rho=vgr$ and the derivatives are with respect to ρ. The solutions for $\lambda=0$ and $\lambda=5g^2$ are shown on the graph below (left). The mass can be estimated as

$$E=\frac{4\pi v}{g}\cdot e(\lambda/g^2).$$

For $\lambda = 0$ the energy is exactly $4\pi v/g$ and for $\lambda/g^2 \to \infty$ the function $e(\lambda/g^2)$ approaches approximately 1.78. The dependence of energy on λ/g^2 is shown on the graph below (on the right).



- The functions $a(r)$ and $f(r)$ approach 1 exponentially fast. The typical size of the monopole is $1/(gv)$, which is also the Compton wavelength of the massive gauge bosons.
- It always holds $E \geq 4\pi v/g$ (*Bogomolny bound*). In the limit $\lambda \to 0$ there is a closed form solution with $E = 4\pi v/g$ (*BPS monopole*, named after Bogomolny, Prasad and Sommerfield). For details, see the exercise sheet. The solution is

$$a(\rho) = 1 - \frac{\rho}{\sinh \rho}, \quad f(\rho) = \coth \rho - \frac{1}{\rho}.$$

- At large r, the magnetic field is

$$B_i^a = -\frac{1}{2}\varepsilon_{ijk} F_{jk}^a \xrightarrow{r\to\infty} \frac{1}{g}\frac{x_i x_a}{r^4}.$$

  This behaves like "direction" $\times$ "Coulomb field", so the solution is called a *magnetic monopole* (though the exact solution behaves like monopole only asymptotically). The $n = 1$ solution is called the *'t Hooft–Polyakov monopole*.
- The magnetic charge is quantized $Q_\mathrm{M} g = 4\pi n$. The smallest charge after $\mathrm{SU}(2) \to \mathrm{U}(1)$ breaking is the U(1) electric charge $Q_\mathrm{E} = g/2$ (from the covariant derivative). This gives the famous *Dirac quantization condition* $Q_\mathrm{M} Q_\mathrm{E} = 2\pi n$.
- In the Standard Model, the situation is different. The SSB is $\mathrm{SU}(2) \times \mathrm{U}(1)_Y \to \mathrm{U}(1)_\mathrm{em}$, so the vacuum manifold is S^3. Since $\pi_2(S^3) = 0$, there are no topological monopoles. The monopoles could exist from the breaking of some GUT group at very high energies, but they would be very heavy and difficult to produce.

4.3 Vacuum structure of Yang–Mills theory

The pure Yang–Mills theory has the Lagrangian

$$\mathcal{L} = -\frac{1}{2}\operatorname{tr}(\mathbf{F}_{\mu\nu}\mathbf{F}^{\mu\nu}).$$

We will consider $G = \mathrm{SU}(2)$ (and we will later generalize to any $\mathrm{SU}(N)$). In the temporal gauge $A_0^a = 0$, the Lagrangian simplifies to

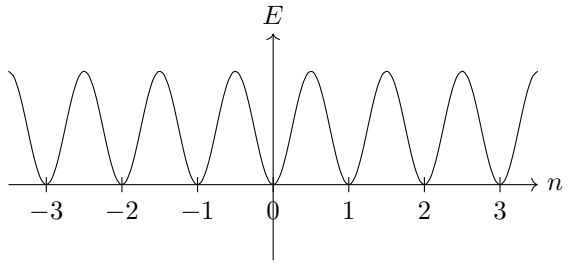
$$\mathcal{L} = \mathrm{tr}(\dot{\mathbf{A}}_i^2) - \frac{1}{2}\,\mathrm{tr}(\mathbf{F}_{ij}^2).$$

The equation of motion $\mathrm{D}_\mu \mathbf{F}^{\mu i} = 0$ gives field equations for A_i and $\mathrm{D}_\mu \mathbf{F}^{\mu 0} = 0$ gives $\mathrm{D}_i E_i = 0$ (non-Abelian Gauss constraint). In the temporal gauge, there are still time-independent gauge transformations $\mathcal{U}(\mathbf{x})$ that do not change A_0^a. Vacuum configurations have to satisfy $\mathbf{F}_{ij} = 0$, so they are pure time-independent gauges:

$$A_i = -\frac{i}{g}\mathcal{U}^\dagger \partial_i \mathcal{U}.$$

Now, $\mathcal{U}$ and A_i are function on $\mathbb{R}^3$. We consider $\mathcal{U}(\mathbf{x})$ which approach a constant SU(2) matrix as $r \to \infty$ (direction independent). This means that we can compactify $\mathbb{R}^3$ to S^3 by adding a point at infinity (an easier example is stereographic projection from S^2 to $\mathbb{R}^2$). Note that here we consider field configurations in space-time, which defines the vacuum manifold. This is different from the previous sections where the topological classification was for the static field configurations at spatial infinity at the restriction of finite energy.

Vacuum configurations are then maps $S^3 \to \mathrm{SU}(2) \sim S^3$, which are topologically non-trivial since $\pi_3(S^3) \cong \mathbb{Z}$. The vacuum manifold falls into equivalence classes labelled by an integer n. Members of different equivalence classes are related by "large" gauge transformations (not continuously connected to $\mathcal{U}(\mathbf{x}) = 1$ in the space $\mathbb{R}^3 \cong S^3$). Schematically, we have the following depiction.



The winding number is calculated as

$$n = \frac{1}{24\pi^2}\varepsilon_{ijk}\int \mathrm{d}^3\mathbf{x}\ \mathrm{tr}((\mathcal{U}^\dagger\partial_i\mathcal{U})(\mathcal{U}^\dagger\partial_j\mathcal{U})(\mathcal{U}^\dagger\partial_k\mathcal{U})).$$

The standard representatives are again denoted by $(\phi, \chi, \varphi) \to (\theta, \chi, n\varphi)$. For $n = 1$, we have

$$\mathcal{U}_1(\mathbf{x}) = \exp\left(\frac{i\pi\boldsymbol{\sigma}\cdot\mathbf{x}}{(a^2+\mathbf{x}^2)^{1/2}} + i\pi\sigma_3\right)$$

for some constant a. It also holds $\mathcal{U}_n = \mathcal{U}_1^n$ and the winding number is additive $n(\mathcal{U}\mathcal{V}) = n(\mathcal{U}) + n(\mathcal{V})$.

If the energy barrier is infinite, no tunnelling is possible and if the system is in one of the vacua, the other ones never affect physics. If there is a non-zero tunnelling amplitude, then we can form a superposition, called the *θ-vacuum* (for $0 \le \theta < 2\pi$):

$$|\theta\rangle = \sum_{n=-\infty}^{\infty} e^{-in\theta}\,|n\rangle\,.$$

These are Bloch states of the periodic potential. The tunnelling amplitude between two vacua depends only on the difference of their winding numbers:

$$\langle\theta'|H|\theta\rangle = \sum_{n,n'=-\infty}^{\infty} e^{in\theta'-in\theta}\langle n'|H|n\rangle = \sum_{n'=-\infty}^{\infty} e^{in'(\theta-\theta')}\sum_k e^{ik\theta}h(|k|) = 2\pi\delta(\theta'-\theta)f(\theta)$$

where $f(\theta)$ is some even function of θ and $h(|k|) = \langle k|H|0\rangle$.

- Since the energy of the vacuum depends on θ, the states are physically distinct for different θ.
- $\langle\theta'|e^{-iHt}|\theta\rangle = 0$ implies no transitions between different θ-vacua. θ is a parameter of the theory (a fundamental constant of nature) once a vacuum with some θ has been chosen.

Is the parameter θ observable? We can calculate it with

$$n = \frac{g^2}{16\pi^2}\int \mathrm{d}^4x\ \mathrm{tr}(\mathbf{F}_{\mu\nu}\tilde{\mathbf{F}}^{\mu\nu})$$

where the integral is in 4d Euclidean space. To see this, we note that the integral is a total divergence of the *Chern–Simons current*

$$j^\mu_{\mathrm{CS}} = \frac{g^2}{16\pi^2}\varepsilon^{\mu\nu\rho\sigma}\,\mathrm{tr}(\mathbf{A}_\nu\mathbf{F}_{\rho\sigma} + 2ig\mathbf{A}_\nu\mathbf{A}_\rho\mathbf{A}_\sigma/3).$$

Then the divergence is

$$\partial_\mu j^\mu_{\mathrm{CS}} = \frac{g^2}{16\pi^2}\int \mathrm{d}^3x\ \mathrm{tr}(\mathbf{F}\tilde{\mathbf{F}}), \quad \Longrightarrow \quad n = \int_{S^3}\mathrm{d}S_\mu\, j^\mu_{\mathrm{CS}}.$$

This holds for any $\mathbf{A}_\mu$. If $\mathbf{A}_\mu$ is a vacuum configuration, then $\mathbf{F}_{\rho\sigma} = 0$ and A_μ is pure gauge. Then there is the Chern–Simons charge

$$n = \int \mathrm{d}^3x\, j^0_{\mathrm{CS}} = Q_{\mathrm{CS}}$$

which is equal to the winding number of the vacuum configuration. Note that CS current is not a conserved current.

Now that we have established that there are distinct vacua and a possibility of tunnelling between them, we expand the expectation value of an operator $\mathcal{O}$ as a sum over all vacua (in a Euclidean space) and then we re-express it as a path integral with the θ-term (in Minkowski space):

$$\begin{aligned}\langle\theta|\mathcal{O}|\theta\rangle \equiv \langle\mathcal{O}\rangle_\theta &= \frac{1}{Z}\sum_{n=-\infty}^{\infty}P(n,\theta)\int\mathcal{D}[\mathbf{A}_\mu]_n\exp\left(-\int\mathrm{d}^4x\,\mathcal{L}^{\mathrm{E}}_{\mathrm{YM}}[\mathbf{A}]\right)\mathcal{O}[\mathbf{A}]\\ &= \frac{1}{Z}\int\mathcal{D}[\mathbf{A}_\mu]\exp\left(i\int\mathrm{d}^4x\,(\mathcal{L}_{\mathrm{YM}}+\mathcal{L}_\theta)\right)\mathcal{O}[\mathbf{A}],\end{aligned}$$

where $P(n,\theta) = e^{-in\theta}$ and

$$\mathcal{L}_\theta = -\theta\frac{g^2}{16\pi^2}\,\mathrm{tr}(\mathbf{F}_{\mu\nu}\tilde{\mathbf{F}}^{\mu\nu}).$$

If $\theta \neq 0$ then $\mathcal{L}$ violates CP symmetry. To see this, note that $\mathbf{A}_\mu$ is C-odd, and hence $\mathbf{F}_{\mu\nu}\tilde{\mathbf{F}}^{\mu\nu}$ is C-odd. Under parity, $\mathbf{F}$ is even and $\tilde{\mathbf{F}}$ is odd, so the product is P-odd. Thus, the θ-term

is CP-odd. But even if $\theta \neq 0$, would we be able to measure it? Not necessarily, since $\mathcal{L}_\theta$ is a total derivative, and it does not contribute to the field equation. However, $\mathbf{A}_\mu$ need not vanish at infinity, precisely when $n \neq 0$. These configurations do contribute to the path integral if tunnelling is possible. This question will be discussed in the next section.

4.4 Instantons and tunnelling in quantum mechanics

Quantum mechanics can be described as a $d = 0 + 1$ field theory. The Hamiltonian is $H = p^2/2m + V(x)$ and the action

$$S = \int \mathrm{d}t \left(\frac{1}{2} m \dot{x}^2 - V(x) \right).$$

For a double-well potential with minima at a and b, the classical solutions are $x(t) = a$ and $x(t) = b$. If no tunnelling occurred, there would be two quantum-mechanical ground state, $|a\rangle$ and $|b\rangle$, with energies $\omega_{a/b}^2 = V''(x_{a/b})$. However, with tunnelling, the eigenfunctions spread over both wells and there is a small energy shift and the tunnelling amplitude given by WKB approximation

$$|T|_{\mathrm{WKB}} \propto \exp\left(-\frac{1}{\hbar} \int_b^a \mathrm{d}x \, \sqrt{2m(V(x) - E)} \right).$$

We want to understand this in path-integral formalism because it generalizes to field theory.

4.4.1 Instantons

The transition amplitude in quantum mechanics is $\langle x_\mathrm{f} | e^{-HT/\hbar} | x_\mathrm{i} \rangle$, where $t = iT$ is imaginary time. This can be expressed as a path integral

$$\langle x_\mathrm{f} | e^{-HT/\hbar} | x_\mathrm{i} \rangle = \mathcal{N} \int_{x(-\tau/2)=x_\mathrm{i}}^{x(\tau/2)=x_\mathrm{i}} \mathcal{D}[x(\tau)] e^{-S_\mathrm{E}[x(\tau)]/\hbar}, \quad S_\mathrm{E} = \int_{-\tau/2}^{\tau/2} \mathrm{d}\tau \left(\frac{1}{2} \left(\frac{\mathrm{d}x}{\mathrm{d}\tau} \right)^2 + V(x(\tau)) \right).$$

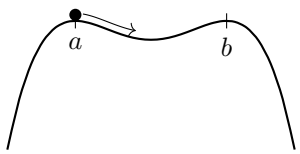
When there are multiple energy eigenstates, we get

$$\langle x_\mathrm{f} | e^{-HT/\hbar} | x_\mathrm{i} \rangle = \sum_{n-\infty}^{\infty} e^{-E_n T/\hbar} \langle x_\mathrm{f} | n \rangle \langle n | x_\mathrm{i} \rangle \xrightarrow{T\to\infty} e^{-E_0 T/\hbar} \langle x_\mathrm{f} | 0 \rangle \langle 0 | x_\mathrm{i} \rangle .$$

To get a semi-classical approximation, we expand S_E around the classical solutions with required boundary condition $x(\pm\tau/2) = x_{\mathrm{f/i}}$. Write $x(\tau) = x_\mathrm{cl}(\tau) + \delta x(\tau)$. The classical solution satisfies the Euclidean equation of motion

$$-m \frac{\mathrm{d}^2 x_\mathrm{cl}}{\mathrm{d}\tau^2} + V'(x_\mathrm{cl}(\tau)) = 0.$$

Once again, this represents a particle moving in the inverted potential $-V(x)$.



The classical action is

$$\begin{aligned} S_{\mathrm{cl}} &= \int_{-T/2}^{T/2} \mathrm{d}\tau \left(\frac{1}{2} \left(\frac{\mathrm{d}x_{\mathrm{cl}}}{\mathrm{d}\tau} \right)^2 + V(x_{\mathrm{cl}}) \right) \\ &= \int_{-T/2}^{T/2} \mathrm{d}\tau \left(2(V(x_{\mathrm{cl}}) - E + E) \right) \\ &\overset{\mathrm{d}\tau \to \mathrm{d}x}{=} \int_b^a \mathrm{d}x \, \sqrt{2m(V(x) - E)} + ET. \end{aligned}$$

This gives exactly the exponent in the WKB tunnelling amplitude (the last term ET only gives a complex phase when we go back to $t = -iT$). The full action is obtained by adding fluctuations around the classical solution

$$S_{\mathrm{E}}[x(\tau)] = S_{\mathrm{cl}} + \frac{1}{2} \int_{-T/2}^{T/2} \mathrm{d}\tau \, \delta x(\tau) \left[-m \frac{\mathrm{d}^2}{\mathrm{d}\tau^2} + V''(x(\tau)) \right] \delta x(\tau) + \mathcal{O}((\delta x)^3).$$

In the path integral formulation, the quadratic term gives a Gaussian integral, which can be evaluated using the eigenvalues of the operator in the square brackets. Given that the operator is nice enough, we can write the complete set of all normalized and orthogonal eigenfunctions:

$$\delta x(\tau) = \sum_n c_n x_n(\tau), \quad \int_{-\tau/2}^{\tau/2} \mathrm{d}\tau \, x_n(\tau) x_m(\tau) = \delta_{nm}, \quad \left(-m \frac{\mathrm{d}^2}{\mathrm{d}\tau^2} + V''(x_{\mathrm{cl}}) \right) x_n = \lambda_n x_n.$$

The integral measure becomes

$$\mathcal{D}[x(\tau)] = \prod_n \frac{\mathrm{d}c_n}{\sqrt{2\pi\hbar}}$$

and the Gaussian integral gives the product of inverse square roots of the eigenvalues:

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}c_n}{\sqrt{2\pi\hbar}} e^{-\frac{1}{2}\lambda_n c_n^2/\hbar} = \frac{1}{\sqrt{\lambda_n}}.$$

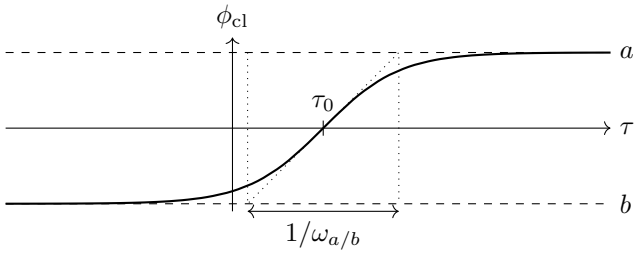
Hence,

$$\langle x_{\mathrm{f}} | e^{-HT/\hbar} | x_{\mathrm{i}} \rangle = \mathcal{N} e^{-S[x_{\mathrm{cl}}]/\hbar} \left(\det \left(-m \frac{\mathrm{d}^2}{\mathrm{d}\tau^2} - V''(x_{\mathrm{cl}}) \right) \right)^{-1/2} (1 + \mathcal{O}(\hbar)).$$

From now on, we will set $m = 1$. For the classical solution with $E = 0$ and $x_{\mathrm{i}} = a$, $x_{\mathrm{f}} = b$, we are solving

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = \sqrt{2V(x)} = \begin{cases} \omega_a(x - a), & x \approx a, \\ \omega_b(b - x), & x \approx b. \end{cases}$$

The solution is called an *instanton*, and it interpolates between the two minima. At large τ the solution exponentially approaches a and at small τ it approaches b. It looks exactly like a kink, only it is localized in the Euclidean time.



The corresponding Euclidean action for this mode is

$$S_{\text{cl}} = \int_b^a \mathrm{d}x \, \sqrt{2V(x)}$$

and there are other approximate solutions for widely separated multi-solitons $I - \tilde{I} + I - \tilde{I} + \ldots$

4.4.2 Calculation of the determinant

To calculate the determinant, we will use the *Gel'fand–Yaglom theorem*, which will allow us to calculate the ratio of two determinants without calculating the eigenvalues explicitly. Consider the operator $\mathcal{M} = -\partial_\tau^2 + V''(\tau)$ and the eigenvalue equation

$$\mathcal{M}\psi(\tau) = \lambda\psi(\tau)$$

with Dirichlet boundary conditions $\psi(\pm T/2) = 0$. Let $\psi_\lambda(\tau)$ be a solution of the above equation for a given λ with initial conditions $\psi_\lambda(-T/2) = 0$ and $\partial_\tau\psi_\lambda(-T/2) = 1$. In general λ is not the solution of the eigenvalue problem, but it is for those λ_n for which $\psi_{\lambda_n}(T/2) = 0$.

Theorem 4.2 (Gel'fand–Yaglom). *Define* $\det(-\partial_\tau^2 - V''(\tau)) = \prod_n \lambda_n$ *and let* $V_1''(\tau), V_2''(\tau)$ *be two different potentials with corresponding operators* $\mathcal{M}_1$ *and* $\mathcal{M}_2$. *Then the Gel'fand–Yaglom theorem states that*

$$\frac{\det(\mathcal{M}_1 - \lambda)}{\det(\mathcal{M}_2 - \lambda)} = \frac{\psi_\lambda^{(1)}(T/2)}{\psi_\lambda^{(2)}(T/2)}.$$

Proof. If λ is an eigenvalue of $\mathcal{M}_1$ then $\psi_\lambda^{(1)}(T/2)$ vanishes and both sides are meromorphic functions of complex λ with the same poles and zeros. Hence, their ratio is an entire function. For large $|\lambda|$, both sides approach 1, so by Liouville's theorem the ratio is identically 1. □

Taking $\lambda = 0$, we find that the following ratio is independent of the potential

$$\frac{\det(-\partial_\tau^2 + V''(\tau))}{\psi_0(T/2)} \equiv \pi\hbar\mathcal{N}^2.$$

The factor $\mathcal{N}$ is the normalization constant of the path integral measure, since

$$\mathcal{N}(\det(-\partial_\tau^2 + V''(\tau)))^{-1/2} = (\pi\hbar\Psi_0(T/2))^{-1/2}.$$

Example 5. Consider $V''(\tau) = \omega^2$ with $x_\mathrm{i} = x_\mathrm{f} = 0$. The classical solution in the inverted potential is constant $x_\mathrm{cl}(\tau) = 0$ (also $S_\mathrm{cl} = 0$). The quantum correction is

$$\langle 0|e^{-HT/\hbar}|0\rangle = (\pi\hbar\psi_0(T/2))^{-1/2}.$$

The solution of the eigenvalue equation $(-\partial_\tau^2 + \omega^2)\psi = 0$ is

$$\psi_0(\tau) = \frac{1}{\omega}\sinh(\omega\tau + \omega T/2), \quad \psi_0(T/2) = \frac{1}{\omega}\sinh(\omega T) \xrightarrow{T\to\infty} \frac{1}{\omega}e^{\omega T}.$$

The ground state becomes

$$e^{-E_0 T/\hbar}|\langle x|0\rangle|^2 = (\pi\hbar\psi_0(T/2))^{-1/2} = \sqrt{\frac{\omega}{\pi\hbar}}e^{-\omega T/2}, \quad \Longrightarrow \quad E_0 = \frac{\hbar\omega}{2}, \ |\langle x|0\rangle|^2 = \sqrt{\frac{\omega}{\pi\hbar}},$$

which is exactly the solution we expected for the harmonic oscillator. Note that this is not a prediction, but a specific choice of normalization constant $\mathcal{N}$.

4.4.3 Double-well potential

For the typical double-well potential $V(x) = g(x^2 - a^2)^2/4$ the energy is $\omega^2 = V''(\pm a) = 2a^2 g$ and the classical action is

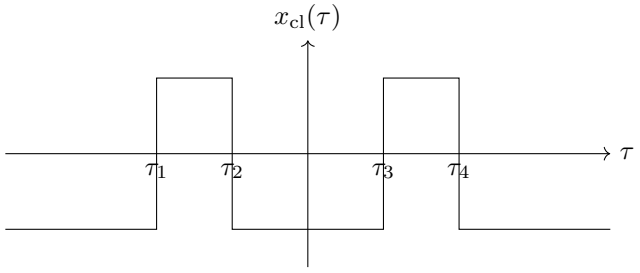
$$S_\mathrm{cl} = \int_{-a}^{a} \mathrm{d}x\, \sqrt{2V(x)} = \frac{\omega^3}{3g}.$$

The solution of the tunnelling between $-a$ and a is described by the movement of a particle in the inverted potential (the instanton) and is given by

$$\langle a|e^{-HT/\hbar}|-a\rangle = e^{-S_\mathrm{cl}/\hbar}\underbrace{\left[\frac{\det(-\partial_\tau^2 + V''(x_\mathrm{cl}))}{\det(-\partial_\tau^2 + \omega^2)}\right]^{-1/2}}_{K}\sqrt{\frac{\omega}{\pi\hbar}}e^{-\omega T/2}.$$

However, there could also be multiple instantons and anti-instantons. For widely separated instantons (*dilute instanton gas approximation*), the action is additive, and so we can approximate the total amplitude for an odd number of instantons n as

$$\langle a|e^{-HT/\hbar}|-a\rangle \approx \sqrt{\frac{\omega}{\pi\hbar}}e^{-\omega T/2}(e^{-S_\mathrm{cl}/\hbar}K)^n.$$



There is also the extra condition that the instantons and anti-instantons have to be ordered in time:

$$\langle \mp a|e^{-HT/\hbar}|\mp a\rangle = \sum_{n\ \text{even}} \int_{-T/2}^{T/2} \mathrm{d}\tau_1 \int_{-T/2}^{\tau_1} \mathrm{d}\tau_2 \cdots \int_{-T/2}^{\tau_{n-1}} \mathrm{d}\tau_n \sqrt{\frac{\omega}{\pi\hbar}}e^{\omega T/\hbar}(e^{-S_\mathrm{cl}/\hbar}K)^n$$

$$\langle \mp a|e^{-HT/\hbar}|\pm a\rangle = \sum_{n \text{ odd}} \int_{-T/2}^{T/2} \mathrm{d}\tau_1 \int_{-T/2}^{\tau_1} \mathrm{d}\tau_2 \cdots \int_{-T/2}^{\tau_{n-1}} \mathrm{d}\tau_n \sqrt{\frac{\omega}{\pi\hbar}} e^{\omega T/\hbar} (e^{-S_{\mathrm{cl}}/\hbar} K)^n.$$

Assuming that the integrand is independent of the τ_i (which we will justify later), the integrals give a factor of $T^n/n!$. Hence,

$$\begin{aligned}\langle \mp a|e^{-HT/\hbar}|\mp a\rangle &= \sqrt{\frac{\omega}{\pi\hbar}} e^{-\omega T/2} \frac{1}{2} \left(e^{K \exp(-S_{\mathrm{cl}}/\hbar)} \mp e^{-K \exp(-S_{\mathrm{cl}}/\hbar)} \right) \\ &\overset{T\to\infty}{=} \sum_{\pm} e^{-E_\pm T/\hbar} \langle \mp a|\pm\rangle \langle \pm| - a\rangle ,\end{aligned}$$

where $|\pm\rangle$ are the energy eigenstates with energy $E_\pm$ which correspond to the degenerate ground states without tunnelling:

$$E_\pm = \frac{1}{2}\hbar\omega \pm \hbar K e^{-S_{\mathrm{cl}}/\hbar}.$$

The second term vanishes to any order in g ("non perturbative"). It also holds $\langle a|-\rangle \langle -| - a\rangle = -\langle a|+\rangle \langle +| - a\rangle$. The lowest state $|-\rangle$ is parity even (the $\mathbb{Z}_2$ symmetry is restored by tunnelling). To justify the large separation of instantons (and consequently the independence of the integrand on τ_i), consider the largest term in the series $x^n/n!$. This is for $n \approx x$. Here, $x = KTe^{-S_{\mathrm{cl}}/\hbar}$. The value for T is then $T/n \sim e^{S_{\mathrm{cl}}}/K$, which is exponentially large in the weak-coupling limit.

The hardest task is to compute the value of K (for $T \to \infty$). There is a subtlety: when $T \to \infty$, if $x_{\mathrm{cl}}(\tau)$ is a solution, then so is $x_{\mathrm{cl}}(\tau - \tau_0)$ for any τ_0. Then, for the zero mode (with eigenvalue $\lambda = 0$) of the expansion of the fluctuation operator (see kink), it holds

$$x_0(\tau) = \frac{1}{\sqrt{S_{\mathrm{cl}}}} \frac{\mathrm{d}x_{\mathrm{cl}}}{\mathrm{d}\tau}$$

which an even function. The factor $1/\sqrt{S_{\mathrm{cl}}}$ comes from the normalization of the mode (and is a consequence of a virial theorem)

$$\int \mathrm{d}\tau \left(\frac{1}{\sqrt{S_{\mathrm{cl}}}} \frac{\mathrm{d}x_{\mathrm{cl}}}{\mathrm{d}\tau} \right)^2 = 1.$$

The issue is that normalization factor K is infinite in the limit $T \to \infty$. There is a flat direction along which the saddle point expansion is not justified, so we need to parametrize the fluctuation a bit more carefully,

$$\delta x(\tau) = c_0 x_0(\tau) + \sum_{n>0} c_n x_n(\tau).$$

The path integral measure becomes

$$\mathcal{D}[x(\tau)] = \frac{\mathrm{d}c_0}{\sqrt{2\pi\hbar}} \prod_{n>0} \frac{\mathrm{d}c_n}{\sqrt{2\pi\hbar}}.$$

For the first term, we change the variable to τ_0 in $x_0(\tau - \tau_0)$. Since the zeroth mode is the derivative of an instanton, it corresponds to the shift in the instanton position. So instead of integrating over the coordinate τ, we will integrate over the time location τ_0 of the instanton. Writing $\Delta x(\tau) = x_0(\tau)\mathrm{d}c_0$ and $\Delta x(\tau) = (\mathrm{d}x_{\mathrm{cl}}/\mathrm{d}\tau)\mathrm{d}\tau_0 = \sqrt{S_{\mathrm{cl}}} x_0(\tau)\mathrm{d}\tau_0$, we get the Jacobian

$$\frac{\mathrm{d}c_0}{\sqrt{2\pi\hbar}} = \sqrt{\frac{S_{\mathrm{cl}}}{2\pi\hbar}} \mathrm{d}\tau_0.$$

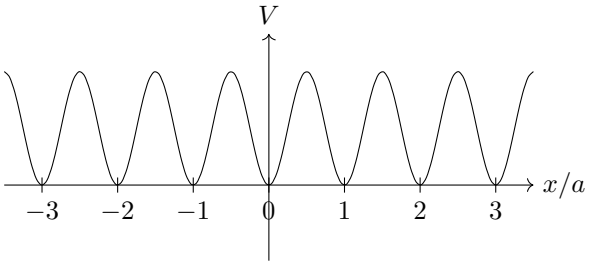
Instead of previous expression for K, we should compute

$$K = \sqrt{\frac{S_{\rm cl}}{2\pi\hbar}} \left[\frac{\det'(-\partial_\tau^2 + V''(x_{\rm cl}))}{\det(-\partial_\tau^2 + \omega^2)} \right]^{-1/2}$$

(with zero eigenvalue removed from the numerator determinant) and do the τ_0 integral exactly. The other τ_i integrals are approximated as before. So the previous result is correct with the new definition of K. The variables that we separate (here only τ_0) are called *collective coordinates*.

4.4.4 Periodic potential

Let $|n\rangle$ be the vacuum state localized at the n-th minimum of a periodic potential $V(x) = V(x+a)$ in the absence of tunnelling.



The θ-vacua are defined as

$$|\theta\rangle = \sum_{n=-\infty}^{\infty} e^{-in\theta} |n\rangle , \quad \theta \in [0, 2\pi).$$

Define the translation operator $\tau |n\rangle = |n+1\rangle$. Then $\tau |\theta\rangle = e^{-i\theta} |\theta\rangle$. Since the Hamiltonian is invariant under translations, $[\tau, H] = 0$ the θ-vacua can be chosen as energy eigenstates (but $|n\rangle$ cannot). Instantons connect neighbouring vacua $|n\rangle$ and $|n \pm 1\rangle$. In order to get from $|n\rangle$ to $|n+m\rangle$, we need at least $|m|$ instantons/anti-instantons, but we can have more, and we are much less constrained than before. For the transition amplitude between two position eigenstates $|n\rangle$ and $|m\rangle$ we need that $\#I - \#\overline{I} = m - n$. The transition amplitude can be written as

$$\langle m|e^{-HT/\hbar}|n\rangle = \sqrt{\frac{\omega}{\pi\hbar}} e^{-\omega T/2} \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{\overline{k}=0}^{\infty} \frac{1}{\overline{k}!} \delta(k - \overline{k} + n - m) (K e^{-S_{\rm cl}/\hbar} T)^{k+\overline{k}}.$$

The Kronecker delta can be expressed as an integral over θ:

$$\delta(k - \overline{k} - m + n) = \int_0^{2\pi} \frac{\mathrm{d}\theta}{2\pi} e^{i\theta(k-\overline{k}-m+n)}$$

and the resulting amplitude is

$$\begin{aligned} \langle m|e^{-HT/\hbar}|n\rangle &= \sqrt{\frac{\omega}{\pi\hbar}} e^{-\omega T/2} \int_0^{2\pi} \frac{\mathrm{d}\theta}{2\pi} e^{i\theta(n-m)} e^{2KT\cos\theta \, \exp(-S_{\rm cl}/\hbar)} \\ &= \int_0^{2\pi} \frac{\mathrm{d}\theta}{2\pi} e^{-E(\theta)T/\hbar} \langle m|\theta\rangle \langle\theta|n\rangle \end{aligned}$$

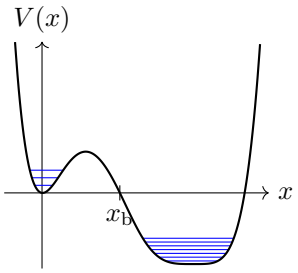
where $E(\theta) = \hbar\omega/2 - \hbar 2K\cos\theta\,\exp(-S_{\rm cl}/\hbar)$. This implies that the energies are exponentially close to $\omega\hbar/2$. It also holds

$$\langle\theta|n\rangle = \left(\frac{\omega}{\pi\hbar}\right)^{1/4} e^{in\theta},$$

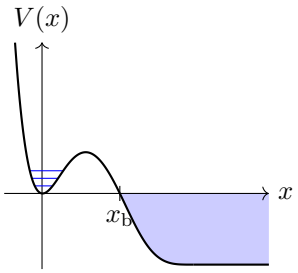
which are exactly the Bloch states of the periodic potential with energy band $E(\theta)$.

4.5 False vacuum decay and bounces

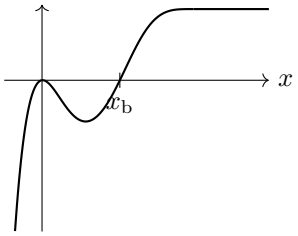
Consider a potential with a metastable minimum (false vacuum) at $x = 0$ and a true vacuum at some other point.



If we push the potential to become completely flat on the right, the spectrum on that side becomes continuous.



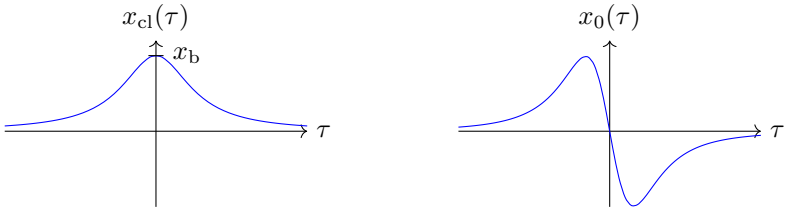
The tunnelling between the false vacuum and continuous true vacuum is described with the motion in the inverted potential.



There are two paths starting and ending at $x = 0$: the one with $x(\tau) = 0$ and the path from $0 \to x_{\rm b} \to 0$ (bounce). There could also be multiple bounces. Summing all options, we obtain

$$\langle 0|e^{-HT/\hbar}|0\rangle = \sqrt{\frac{\omega}{\pi\hbar}}e^{-\omega T/2}e^{KT\exp(S_{\rm cl}/\hbar)}$$

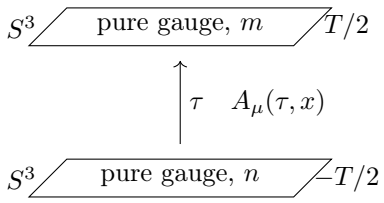
with $E_0 = \hbar\omega/2 + \hbar K \exp(S_{\rm cl}/\hbar)$. The non-trivial classical solution is the bounce, which starts at $x = 0$ at $\tau = -\infty$, goes to $x_{\rm b}$ at $\tau = 0$ and returns to $x = 0$ at $\tau = \infty$. The zero-eigenvalue mode is related to the derivative of the bounce with respect to τ.



Since x_0 has one node, there must exist exactly one eigenfunction with negative eigenvalue in the fluctuation operator around the bounce and the value of K is imaginary. Consequently, the energy E_0 has an imaginary part, which gives the decay rate of the false vacuum (called *metastable state*): $\Gamma = 2\mathrm{Im}\ (E_0)$ so that $|\Psi(x)|^2 \propto e^{-\Gamma t}$. There is again a subtlety with the saddle point approximation for the potential that doesn't grow at infinity (Coleman). One has to veer into the complex plane on the right side of the bounce solution, which only gives half of the expected result. So we have to replace $K \to K/2$ in the previous result. Then $\Gamma = 2\mathrm{Im}\,(E_0) = \hbar|K|e^{-S_{\rm cl}/\hbar}$.

4.6 The Yang–Mills instanton

We will work with SU(2) gauge field. To compute the tunneling amplitude between vacua with different winding numbers, we will look for a Euclidean-time dependent field configurations which solve the classical field equation and interpolate beween two time-independent vacuum states $\partial_i \mathcal{U}/g$ which are classified by winding number n.



From $\partial_\mu j^\mu_{\rm CS} = \frac{g^2}{16\pi^2}\,\mathrm{tr}(\mathbf{F}\tilde{\mathbf{F}})$, the field configuration $A_\mu(\tau, \mathbf{x})$ should satisfy

$$\int \mathrm{d}^4x\ \mathrm{tr}(\mathbf{F}\tilde{\mathbf{F}}) = \frac{16\pi^2}{g^2}(m - n).$$

The Derrick's theorem for static solutions in $d = D+1$ dimensional space gives $0 = (D-4)S_{\rm E}[A]$, so $D = 4$ is the only possible dimension where solutions exist. That's lucky for us, since we are interested in 4D Yang–Mills theory (3 spatial dimensions and one Euclidean time).

4.6.1 The instanton for $n - m = 1$

Without loss of generality, we can set $m = 0$ and $n = 1$. First, we note that

$$S_{\rm E} = \int \mathrm{d}^4x\ \frac{1}{2}\,\mathrm{tr}(\mathbf{F}\mathbf{F}) = \int \mathrm{d}^4x\ \frac{1}{2}\,\mathrm{tr}(\tilde{\mathbf{F}}\tilde{\mathbf{F}}).$$

With this, we can make the following estimate

$$0 \leq \int \mathrm{d}^4x \ \mathrm{tr}((\mathbf{F} \pm \tilde{\mathbf{F}})^2) = 4S_\mathrm{E} \pm \int \mathrm{d}^4x \ 2\,\mathrm{tr}(\mathbf{F}\tilde{\mathbf{F}})$$
$$S_\mathrm{E} \geq \frac{1}{2}\left|\int \mathrm{d}^4x \ \mathrm{tr}(\mathbf{F}\tilde{\mathbf{F}})\right| = \frac{8\pi^2}{g^2}|n|$$

This is another *Bogomolny bound*. Minimal action solutions to the field equation have $\mathbf{F} = \pm\tilde{\mathbf{F}}$ ((anti-)self-dual solutions). For $n = 1$, we have $\mathbf{F} = \tilde{\mathbf{F}}$ and the following ansatz

- $\mathbb{1}a + i\mathbf{b}\cdot\boldsymbol{\sigma}$ with $a^2 + \mathbf{b}^2 = 1$ is an SU(2) matrix.
- $$\mathcal{U}(\tau, \mathbf{x}) = \frac{\tau + i\boldsymbol{\sigma}\cdot\mathbf{x}}{r}, \quad r = \sqrt{\tau^2 + \mathbf{x}^2}$$
  is an SU(2) matrix which depends only on the direction of $x^\mu = (\tau, \mathbf{x})$. $\mathcal{U}$ is a map from S^3 (at infinity in $\mathbb{R}^4$) to SU(2) $\cong S^3$ with winding number 1.
- The ansatz is
  $$A_\mu(x) = -\frac{i}{g} f(r)\mathcal{U}^\dagger(\hat{x})\partial_\mu\mathcal{U}(\hat{x}), \quad \hat{x} = x_\mu/\sqrt{\tau^2 + \mathbf{x}^2}$$
  with $f(0) = 0$ and $f(\infty) = 1$.

In the coordinates (r, χ, ψ, φ) we can get a differential equation for $f(r)$ from $F_{r\chi} = -F_{\psi\varphi}$:

$$r\frac{\mathrm{d}f}{\mathrm{d}r} = 2f(1-f), \quad \Longrightarrow \quad f(r) = \frac{r^2}{r^2 + \rho^2}$$

where ρ is the size of the instanton. By using $\mathcal{U}^n$ in the ansatz, one can get any multi-instanton configuration. This is not the most general solution; to this solution we can apply the following symmetries:

- translations in $\mathbb{R}^4$: $x_\mu \to x_\mu - a_\mu$ ($r^2 = (\tau - a_0)^2 + (\mathbf{x} - \mathbf{a})^2$), which give 4 parameters,
- ρ: classical scale invariance (1 parameter),
- rotations in $\mathbb{R}^4$, i.e. SO(4) $\cong$ SU(2) $\times$ SU(2). Taking V from the first and W from the second SU(2), we have $A_\mu \to WA_\mu W^\dagger$ and $\mathcal{U} \to V\mathcal{U}W^{-1}$. The V drops out and W can be absorbed into a gauge transformation (which is already accounted in the last point), so there are no new parameters from this.
- Special conformal transformations (no new parameters).
- Global gauge transformations: $A_\mu \to UA_\mu U^\dagger$ with $U \in$ SU(2) (3 parameters).

This is an eight-parameter family of solutions called the *moduli space* of the instanton. The Jacobian for the collective coordinates will thus be $J \sim \sqrt{S_\mathrm{cl}}^8 \sim g^{-8}$. In order to write down the general solution most efficiently, we introduce the *'t Hooft symbols*:

$$\eta_{a\mu\nu} = \begin{cases} \varepsilon_{a\mu\nu}, & \mu,\nu = 1,2,3, \\ \delta_{a\nu}, & \mu = 0, \nu = 1,2,3 \\ -\delta_{a\mu}, & \nu = 0, \mu = 1,2,3, \\ 0 & \mu = \nu = 0, \end{cases}, \quad \overline{\eta}_{a\mu\nu} = \begin{cases} \varepsilon_{a\mu\nu}, & \mu,\nu = 1,2,3, \\ -\delta_{a\nu}, & \mu = 0, \nu = 1,2,3 \\ \delta_{a\mu}, & \nu = 0, \mu = 1,2,3, \\ 0 & \mu = \nu = 0, \end{cases}.$$

Then

$$A^a_\mu(x) = R_{ab}\frac{2}{g}\frac{\eta_{b\mu\nu}(x-a)_\nu}{(x-a)^2+\rho^2}, \quad \mathbf{A}_\mu = A^a_\mu\frac{\sigma^a}{2}, \quad F^a_{\mu\nu} = -R_{ab}\frac{4}{g}\frac{\rho^2\eta_{b\mu\nu}}{((x-a)^2+\rho^2)^2}$$

where R_{ab} is a global SU(2) rotation. Note that R_{ab} are essentially SO(3) matrices, so from their properties $\det R = 1$ and $RR^T = R^T, R = I$ one can show the following relations:

$$R_{ab}R_{ab} = 3, \quad R_{ab}R_{ac} = \delta_{bc}, \quad \varepsilon_{abc}R_{ax}R_{by}R_{cz} = \varepsilon_{xyz}, \quad \varepsilon_{abc}R_{ax}R_{by} = \varepsilon_{xyz}R_{cz}.$$

Next, we want to calculate the energy of the θ-vacuum in SU(2) Yang–Mills theory. We have to contain the integration region in a finite 4-dimensional box of size T in the τ direction and V in the spatial directions. With this, we replace $T \to TV$ and the formula from before becomes the energy density

$$E(\theta)/V = \text{const} - 2K\cos(\theta)e^{-8\pi^2/g^2}.$$

θ-vacua become physically distinct because they have different energy density.

4.6.2 Evaluation of the fluctuation determinant K

It is not enough to only consider the fluctuations around the gauge field, but also around the classical solution of the ghosts (which are classically zero):

$$A^a_\mu = A^a_{\mu,I} + \hat{A}^a_\mu, \quad c^a = 0 + \hat{c}^a, \quad a = 1,2,3.$$

To fix the gauge we pick the background gauge:

$$\partial^\mu \hat{A}^a_\mu + [A_{\mu,I}, \hat{A}^\mu]^a = 0.$$

The saddle point approximation of the classical action is

$$S_{\text{cl}}[A^a_\mu, c^a] = \frac{8\pi^2}{g} + \frac{1}{2}\int \mathrm{d}^4x \begin{pmatrix}\hat{A}^a_\mu & c^a & \overline{c}^a\end{pmatrix} \begin{pmatrix} L_A[A^a_{\mu,I}] & 0 & 0 \\ 0 & 0 & L_c[A^a_\mu] \\ 0 & L_c[A^a_\mu] & 0 \end{pmatrix} \begin{pmatrix}\hat{A}^a_\mu \\ c^a \\ \overline{c}^a\end{pmatrix}$$

and the determinant can be written as

$$K = \text{normalization} \times \text{Jacobian for collective coordinates} \times [\det{}'(L_A)]^{-1/2}[\det(L_c)].$$

In order to calculate it, we have to first take a few steps.

(1) We need to add a UV regulator to the gauge field (*Pauli–Villars regularization*):

$$\frac{1}{p^2} \to \frac{1}{p^2} - \frac{1}{p^2+M^2_{\text{UV}}}, \quad \det{}'(L_A) \to \frac{\det{}'(L_A)}{\det(L_A+M^2_{\text{UV}})}.$$

(2) Jacobian for the zero modes gives a factor of $\sqrt{S_{\text{cl}}} \sim 1/g$ for each collective coordinate. Additionally, every zero mode brings a factor of M_{UV} from $[\det(L_A + M^2_{\text{UV}})]^{1/2}$.

We write

$$\begin{aligned}
&\mathcal{N}\int\mathcal{D}[A^a_\mu,c^a,\overline{c}^a]e^{-S_{\rm cl}}\to\\
&\quad\to\text{const.}\times e^{-S_{\rm cl}}(M_{\rm UV})^8S_{\rm cl}^4\int\mathrm{d}^4a\int\mathrm{d}\,\mathrm{vol}(\mathrm{SU}(2))\int\mathrm{d}\rho\,\rho^3\times\text{non-zero mode determinant}\\
&\quad\overset{(1)}{=}\text{const.}\times\int\mathrm{d}^4a\,\mathrm{d}\,\mathrm{vol}(\mathrm{SU}(2))\,\mathrm{d}\rho\,\frac{1}{\rho^5}\frac{1}{g_0^8}e^{-8\pi^2/g^2(\rho^{-1})}e^{(-b_0+8)\ln(M_{\rm UV}\rho)}\times\text{non-zero modes}\\
&\quad\overset{(2)}{=}\text{const.}\times\int\mathrm{d}^4a\int_0^\infty\frac{\mathrm{d}\rho}{\rho^5}\frac{1}{g(\rho^{-1})^8}e^{-8\pi^2/g^2(\rho^{-1})}
\end{aligned}$$

With the above two equalities, we perform the following steps:

(1) introduce renormalized coupling

$$g_0\overset{\text{dimreg}}{=}Z_g\tilde{\mu}^\varepsilon g(\mu)\overset{\text{PV}}{=}\left(1-\frac{\alpha_{\rm s}}{4\pi}b_0\ln\frac{M_{\rm UV}}{\mu}\right)g(\mu)$$

where (in dim-reg)

$$Z_g=1-\frac{\alpha_{\rm s}}{4\pi}\frac{b_0}{2\varepsilon},\quad b_0=\frac{11}{3}N-\frac{2}{3}n_{\rm f}\xrightarrow{\mathrm{SU}(2)}\frac{22}{3}.$$

From this, it follows

$$\frac{1}{g_0^2}=\frac{1}{g(\mu)^2}+\frac{b_0}{8\pi^2}\ln\frac{M_{\rm UV}}{\mu}.$$

We choose $\mu=1/\rho$ and use $\rho^3M_{\rm UV}^8=e^{8\ln(M_{\rm UV}\rho)}/\rho^5$.

(2) Non-zero mode determinant is very difficult to calculate. The result has to be such that the final expression is gauge invariant. Hence, we know the structure must be $(M_{\rm UV}\rho)^{-2/3}\times$ const. The integral over SU(2) space is also performed in this step.

The behaviour in the ρ-integral

- for $\rho\to0$ is finite due to asymptotic freedom,
- for $\rho\to\infty$ the exponential suppression is lifted and infrared events occur $\rho^{-5}e^{-b_0\ln(M_{\rm UV}\rho)}\sim\rho^{7/3}$.

We will now generalize this construction from SU(2) to SU(N) Yang–Mills theory. We know that $\pi_3(\mathrm{SU}(N))\cong\mathbb{Z}$ for $N\geq2$ and that SU(2) $<$ SU(N) (subgroup). This means that the instanton solution is the same, but there are additional zero modes from different ways we can embed SU(2) into SU(N). The generators of SU(N) that are outside SU(2) subgroup don't generate new solutions, so we only have $2(N-2)\cdot2$ new generators that give new solutions. The final result is simple:

- $e^{-8\pi^2/g(\rho^{-1})}$ but with $b_0=11N/3$,
- $g^{-8}\to g^{-4N}$,
- the constant changes.

4.7 Instantons, fermions and chiral anomaly

We consider SU(N) Yang–Mills theory with n_f massless Dirac fermions in the fundamental representation:

$$\mathcal{L} = -\frac{1}{2}\operatorname{tr}\mathbf{F}^2 + \sum_{f=1}^{N_\mathrm{f}} (\overline{\psi}_{\mathrm{L}f} i\not{D}\psi_{\mathrm{L}f} + \overline{\psi}_{\mathrm{R}f} i\not{D}\psi_{\mathrm{R}f} - m_f[\overline{\psi}_{\mathrm{L}f}\psi_{\mathrm{R}f} + \overline{\psi}_{\mathrm{R}f}\psi_{\mathrm{L}f}]).$$

Without the mass terms, there is $\mathrm{U}(N_\mathrm{f}) \otimes \mathrm{U}(N_\mathrm{f})$ global symmetry, which we can (in case $N_\mathrm{f} = 3$) decompose into

$$\mathrm{U}(3) \otimes \mathrm{U}(3) = \mathrm{U}(1)_V \otimes \mathrm{U}(1)_A \otimes \mathrm{SU}(3)_V \otimes \mathrm{SU}(3)_A.$$

The last symmetry is spontaneously broken by the chiral condensate, $\mathrm{SU}(3)_A$ gives 8 Goldstone bosons $\pi^{\pm,0}$, $K^\pm$, $\overline{K}^\pm$ and η. The $\mathrm{U}(1)_V$ is exact and corresponds to the quark number conservation and $\mathrm{U}(1)_A$ is called the *chiral symmetry* which is anomalous. Its Noether current is

$$j_5^\mu = \sum_f \overline{\psi}_f \gamma^\mu \gamma_5 \psi_f, \quad \partial_\mu j_5^\mu = \frac{N_\mathrm{f} g_\mathrm{s}^2}{32\pi^2} \varepsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} = \frac{N_\mathrm{f} g_\mathrm{s}}{16\pi^2} \operatorname{tr}(\mathbf{F}\tilde{\mathbf{F}}) = 2N_\mathrm{f} \partial_\mu j_\mathrm{CS}^\mu.$$

Anomaly arises form non-invariance of $\mathcal{D}[\Psi, \overline{\Psi}]$ under a $\mathrm{U}(1)_A$ change of the integration variables. Under a local $\mathrm{U}(1)_A$ transformation with $e^{i\alpha(x)\gamma_5}\Psi(x)$, we have

$$\begin{aligned} Z &= \int \mathcal{D}[\Psi, \overline{\Psi}] \exp\left(i \int \mathrm{d}^4x\, \overline{\Psi} i\not{D}\Psi + \dots \right) \\ &\to \int \mathcal{D}[\Psi, \overline{\Psi}] \exp\left(\overline{\Psi} i\not{D}\Psi + \dots + \alpha(x) \left(\partial_\mu j_5^\mu - \frac{N_\mathrm{f} g_\mathrm{c}^2}{8\pi^2} \operatorname{tr}(\mathbf{F}\tilde{\mathbf{F}}) \right) \right). \end{aligned}$$

Although the anomaly term is a total divergence, it can produce physical effects, because gauge fields don't vanish at ∞. This exactly corresponds to configurations with non-zero *Pontryagin index*. Since the divergences of Chern–Simmons and chiral current are related, it holds that $j_5^\mu - 2N_\mathrm{f} j_\mathrm{CS}^\mu$ is conserved. For the charges this implies

$$n_\mathrm{R} - n_\mathrm{L} - 2nN_\mathrm{f} = \text{const.}$$

In an instanton tunneling process chirality is violated according to this equation. In the following, we will try to understand how this happens and whether we can give a quantitative amplitude for such process.

Recall θ-vacua structure is accounted for by adding

$$\exp\left(i \int \mathrm{d}^4x \, \left(-\frac{\theta g^2}{16\pi^2} \operatorname{tr}(\mathbf{F}\tilde{\mathbf{F}}) \right) \right)$$

to the path integral. Take $\alpha(x) = \text{const}$, i.e., global $\mathrm{U}(1)_A$. The effect of this $\mathrm{U}(1)_A$ transformation is $\theta \to \theta + 2N_\mathrm{f}\alpha$, hence θ cannot have a physical effect if $\mathrm{U}(1)_A$ is an exact classical

symmetry ($m_\mathrm{f} = 0$) and anomalous. It is enough in fact that one of the fermions is massless,

$$\begin{aligned}\mathcal{L}_m &= -m\overline{\psi}_\mathrm{L}\psi_\mathrm{R} - m^*\overline{\psi}_\mathrm{R}\psi_\mathrm{L} = -\begin{pmatrix}\overline{\psi_\mathrm{L}} & \overline{\psi}_\mathrm{R}\end{pmatrix}\begin{pmatrix}0 & m\\ m^* & 0\end{pmatrix}\begin{pmatrix}\psi_\mathrm{L}\\ \psi_\mathrm{R}\end{pmatrix}\\ &\xrightarrow{\mathrm{U}(1)_A} -|m|\begin{pmatrix}\overline{\psi_\mathrm{L}} & \overline{\psi}_\mathrm{R}\end{pmatrix}e^{-i\alpha\gamma^5}\begin{pmatrix}0 & e^{-i\phi}\\ e^{i\phi} & 0\end{pmatrix}e^{i\alpha\gamma_5}\begin{pmatrix}\psi_\mathrm{L}\\ \psi_\mathrm{R}\end{pmatrix}\\ &= -|m|\begin{pmatrix}\overline{\psi_\mathrm{L}} & \overline{\psi}_\mathrm{R}\end{pmatrix}\begin{pmatrix}0 & e^{-i(\phi-2\alpha)}\\ e^{i(\phi-2\alpha)} & 0\end{pmatrix}\begin{pmatrix}\psi_\mathrm{L}\\ \psi_\mathrm{R}\end{pmatrix}.\end{aligned}$$

We used $m = |m|e^{i\phi}$ and the chiral representation for γ^5

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

In short, a global chiral transformation just changes the phase of the Dirac mass $\phi \to \phi + 2\alpha$. This is also the reason why Dirac mass is always real – it can always be made real by such transformation. For N_f quarks, the determinant of the mass matrix M changes as

$$\det M \to e^{2iN_\mathrm{f}\alpha}\det M$$

and consequently $e^{-i\theta}\det M$ is invariant. Hence, the physical quantities can only depend on $e^{-i\theta}\det M$ or equivalently

$$\overline{\theta} = \theta - \operatorname{argdet} M.$$

The value of $\overline{\theta}$ is minuscule (smaller than detection limits) and the question of why is it so small is called the *strong CP problem*. The value of $\operatorname{argdet} M$ is non-zero due to CP violation and CKM phases. If one fermion mass is zero, $\det M = 0$ and θ is not observable.

Yet another way to see that θ is unobservable if (at least one) $m = 0$ is to look at the path integral with a specific winding number n

$$\int \mathcal{D}[A_\mu,\Psi,\overline{\Psi}]_n e^{in\theta}e^{i\int \mathrm{d}^4x\,\mathcal{L}_\mathrm{YM}}e^{i\int \mathrm{d}^4x\,\overline{\Psi}i\not{\mathrm{D}}\Psi} = \int \mathcal{D}[A_\mu]_n\, e^{i\theta n}e^{i\int \mathrm{d}^4x\,\mathcal{L}_\mathrm{YM}}\det(i\not{\mathrm{D}}[A_{\mu,n}]) = 0.$$

The expression is zero if $\not{\mathrm{D}}[A_{\mu,n}]$ has a zero mode (for the Euclidean Dirac operator). It turns out that also the violation of chirality is connected to zero modes of $i\not{\mathrm{D}}[A_\mu]$. To calculate zero modes of the Dirac operator on the (anti-)instanton background we first take $n = \pm 1$, $N_\mathrm{f} = 1$ and $m = 0$ (massless fermion). Since we are in Euclidean space, we have to use the Euclidean gamma matrices:

$$\gamma^0_\mathrm{E} = \gamma^0 = \begin{pmatrix}0 & \mathbb{1}\\ \mathbb{1} & 0\end{pmatrix},\quad \gamma^i_\mathrm{E} = -i\gamma^i = \begin{pmatrix}0 & -i\sigma^i\\ i\sigma^i & 0\end{pmatrix}.$$

From now on, we will drop the subscript E. The bispinor is written as

$$\Psi = \begin{pmatrix}\psi_\mathrm{L}\\ \psi_\mathrm{R}\end{pmatrix},\quad \overline{\Psi} = \Psi^\dagger\gamma^0 = \begin{pmatrix}\psi^\dagger_\mathrm{R} & \psi^\dagger_\mathrm{L}\end{pmatrix} = \begin{pmatrix}\overline{\psi}_\mathrm{R} & \overline{\psi}_\mathrm{L}\end{pmatrix}.$$

The Euclidean Dirac term reads

$$\overline{\Psi}i\not{\mathrm{D}}\Psi = \overline{\psi}_\mathrm{L}i\overline{\sigma}^\mu\mathrm{D}_\mu\psi_\mathrm{L} + \overline{\psi}_\mathrm{R}i\sigma^\mu\mathrm{D}_\mu\psi_\mathrm{R}.$$

We note that $i\overline{\sigma}^\mu \mathrm{D}_\mu$ and $i\sigma^\mu \mathrm{D}_\mu$ are not self-adjoint, but rather conjugate to each other $(i\overline{\sigma}^\mu \mathrm{D}_\mu)^\dagger = i\sigma^\mu \mathrm{D}_\mu$. To see this, write

$$\int \mathrm{d}^4x\, \overline{\psi}_\mathrm{L} i\overline{\sigma}^\mu \mathrm{D}_\mu \psi_\mathrm{L} = \int \mathrm{d}^4x\, \overline{\psi}_\mathrm{R} (i\sigma^\mu \mathrm{D}_\mu)^\dagger \psi_\mathrm{R}.$$

So the properly self-adjoint operator is

$$i\not{\mathrm{D}} = \begin{pmatrix} 0 & i\sigma^\mu \mathrm{D}_\mu \\ i\overline{\sigma}^\mu \mathrm{D}_\mu & 0 \end{pmatrix}.$$

Note. Let $\lambda \neq 0$ be an eigenvalue of $\mathrm{D}_\mathrm{L} \equiv i\overline{\sigma}^\mu \mathrm{D}_\mu$, i.e., there exists $\psi_\mathrm{L} \neq 0$ such that $\mathrm{D}_\mathrm{L}\psi_\mathrm{L} = \lambda\psi_\mathrm{R}$ and $\mathrm{D}_\mathrm{R}\psi_\mathrm{R} \equiv i\sigma^\mu \mathrm{D}_\mu \psi_\mathrm{R} = \lambda\psi_\mathrm{L}$. Hence, λ is a non-zero eigenvalue of $i\not{\mathrm{D}}$ with and therefore real. Since

$$\mathrm{D} = \begin{pmatrix} 0 & \mathrm{D}_\mathrm{R} \\ \mathrm{D}_\mathrm{L} & 0 \end{pmatrix}, \quad \mathrm{D}^2 = \begin{pmatrix} \mathrm{D}_\mathrm{R}\mathrm{D}_\mathrm{L} & 0 \\ 0 & \mathrm{D}_\mathrm{L}\mathrm{D}_\mathrm{R} \end{pmatrix},$$

λ^2 is simultaneously eigenvalue of $\mathrm{D}_\mathrm{R}\mathrm{D}_\mathrm{L}$ and $\mathrm{D}_\mathrm{L}\mathrm{D}_\mathrm{R}$. By assumption $\lambda \neq 0$, $\psi_\mathrm{L} \neq 0$ it follows $\psi_\mathrm{R} \neq 0$ (otherwise $\psi_\mathrm{L} = \mathrm{D}_\mathrm{R}\psi_\mathrm{R}/\lambda = 0$). For every $\lambda \neq 0$ there is a non-vanishing left- and right-handed eigenfunction, $\psi_\mathrm{L}, \psi_\mathrm{R}$. Then $(\psi_\mathrm{L}, -\psi_\mathrm{R})^T$ is an eigenfunction of $i\not{\mathrm{D}}$ with eigenvalue $-\lambda$. Non-zero eigenvalues thus come in pairs $\pm\lambda$ for $i\overline{\sigma}^\mu \mathrm{D}_\mu$ and $i\sigma^\mu \mathrm{D}_\mu$, but zero eigenvalues need not come in pairs. We define the index of the Dirac operator as

$$\begin{aligned} \mathrm{index}(i\overline{\sigma}^\mu \mathrm{D}_\mu) &= \dim\ker(i\overline{\sigma}^\mu \mathrm{D}_\mu) - \dim\ker(i\sigma^\mu \mathrm{D}_\mu) \\ &= \text{no. of zero EVs of } (i\overline{\sigma}^\mu \mathrm{D}_\mu) - \text{no. of zero EVs of } (i\sigma^\mu \mathrm{D}_\mu). \end{aligned}$$

There is a lot of mathematical development in the field of such operators, we only note the *Atiyah–Singer index theorem*:

$$\mathrm{index}(i\overline{\sigma}^\mu \mathrm{D}_\mu) = \frac{g^2}{16\pi^2} \int \mathrm{d}^4x\ \mathrm{tr}(\mathbf{F}\tilde{\mathbf{F}}) = n.$$

From this, we would expect that the Dirac operator in the one-instanton background has one left-handed and no right-handed modes. To check this, we will calculate the zero mode. Leaving the constant i, we want to find the solution for the zero mode of the instanton ($n = 1$):

$$\overline{\sigma}^\mu \mathrm{D}_\mu \psi_\mathrm{L} = (\mathrm{D}_0 + i\boldsymbol{\sigma}\cdot\mathbf{D})\psi_\mathrm{L} = 0.$$

From the ansatz form of A^a_μ and $F^a_{\mu\nu}$, we try the following ansatz for the zero mode (using α as spinor index and i as SU(2) index):

$$\Psi^\mathrm{L}_{\alpha,i}(x) = \frac{c_{\alpha i}}{(x^2+\rho^2)^s},$$

where $c_{\alpha i}$ is a constant spinor. Without loss of generality, we took $R_{ab} = 1$ and $a^\mu = 0$. The covariant derivative is

$$\mathrm{D}_\mu = \partial_\mu - igT^a A^a_\mu = \partial_\mu - i\frac{2\eta^a_{\mu\nu}x_\nu}{x^2+\rho^2}\frac{\sigma^a}{2}.$$

The object $(\overline{\sigma}_\mu \mathrm{D}_\mu)^{ij}_{\alpha\beta}$ is a matrix both in ij and in $\alpha\beta$. Acting on the ansatz, we get

$$(\overline{\sigma}_\mu \mathrm{D}_\mu)^{ij}_{\alpha\beta}\psi^\mathrm{L}_{\beta,j} = \frac{1}{(x^2+\rho^2)^{s+1}}\left(-2sx^0\delta_{\alpha\beta}\delta_{ij} - i\eta^a_{0\nu}x_\nu\sigma^a_{ij}\delta_{\alpha\beta} - i\sigma^k_{\alpha\beta}2sx^k + \eta^a_{k\nu}\delta_{ij}x^\nu\sigma^k_{\alpha\beta}\sigma^a_{ij}\right)c_{\beta j}.$$

The expression consists of 4 terms, which all have to vanish. Since $\eta^a_{00} = 0$, the second term is already 0 for x^0. Combining 1 and 4 (both factors with x^0), we get

$$(-2s\delta_{ij}\delta_{\alpha\beta} + \eta^a_{k0}\sigma^k_{\alpha\beta}\sigma^a_{ij})c_{\beta j} = 0, \quad \Longrightarrow \quad -(\boldsymbol{\sigma}_{\alpha\beta}\cdot\boldsymbol{\sigma}_{ij} + 2s\delta_{\alpha\beta}\delta_{ij})c_{\beta j} = 0.$$

Note that for the sum of two spin-1/2 operators ($S_1 = S_2 = 1/2$) it holds that

$$\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2 = \frac{1}{2}(\boldsymbol{\sigma}_1+\boldsymbol{\sigma}_2)^2 - \frac{1}{2}\boldsymbol{\sigma}_1^2 - \frac{1}{2}\boldsymbol{\sigma}_2^2 = 2\mathbf{S}^2 - 2\cdot S_1(S_1+1) - 2\cdot S_2(S_2-1) = 2\mathbf{S}^2 - 3.$$

On symmetric $c_{\beta j}$, $\mathbf{S}^2$ has eigenvalues $S(S+1) = 2$ (spin $S = 1$) and $S(S+1) = 0$ (spin $S = 0$). Hence, we have the equation

$$2S(S+1) - 3 + 2s = 0$$

with solutions $s = 3/2$ for $S = 0$ and $s = 1/2$ for $S = 1$. The "spin 1" solution is non-normalizable, but "spin 0" is:

$$\psi^{\mathrm{L}}_{\alpha i}(x) = \frac{\varepsilon_{\alpha i}}{(x^2+\rho^2)^{3/2}}$$

where $\varepsilon_{\alpha i}$ is antisymmetric 2×2 matrix. We repeat the same steps for $\sigma^\mu \mathrm{D}_\mu\psi_{\mathrm{R}} = (\mathrm{D}_0 - i\boldsymbol{\sigma}\cdot\mathbf{D})\psi_{\mathrm{R}} = 0$. The only difference in the solution is $s \to -s$, so the two possible exponents are $s = 1/2, -3/2$. There are thus no renormalizable right-handed zero modes. It remains to check that $\Psi^{\mathrm{L}}_{\alpha i}$ above also gives zero for the spatial components x^i.

In the $n = 1$ instanton, there is a left-handed zero mode and the index of $i\overline{\sigma}^\mu \mathrm{D}_\mu$ is 1 and of $i\sigma^\mu \mathrm{D}_\mu$ is -1:

$$\Psi_0(x) = \frac{\sqrt{2}}{\pi}\frac{\rho}{(x^2+\rho^2)^{3/2}}\begin{pmatrix}\varepsilon_{\alpha i}\\ 0\end{pmatrix}.$$

Returning to the fermionic path integral, we decompose the left-handed (and right-handed) solution into the eigenmodes

$$\Psi_{\mathrm{L}}(x) = c_0\Psi^{\mathrm{L}}_0(x) + \sum_{i>0} c_i\psi^{\mathrm{L}}_i(x), \quad \psi_{\mathrm{R}}(x) = \sum_{i>0} d_i\psi^{\mathrm{R}}_i$$

where c_0, c_i, d_i are Grassmann coefficients. Similar decomposition works for $\overline{\psi}_{\mathrm{L}}$ and $\overline{\psi}_{\mathrm{R}}$ with coefficients $\overline{c}_0, \overline{c}_i, \overline{d}_0, \overline{d}_i$ (note that $\overline{d}_0$ is not necessarily zero). The term in the Lagrangian decomposes as

$$\int \mathrm{d}^4x\ \overline{\Psi} i\not{\mathrm{D}}\Psi = \sum_i \lambda_i(\overline{c}_i c_i + \overline{d}_i d_i)$$

and the path integral measure is

$$\int \mathcal{D}[\Psi,\overline{\Psi}] = \int \mathrm{d}c_0\,\mathrm{d}\overline{d}_0 \prod_i \mathrm{d}c_i\,\mathrm{d}\overline{c}_i \prod_i \mathrm{d}d_i\,\mathrm{d}\overline{d}_i.$$

Hence,

$$\int \mathcal{D}[\Psi,\overline{\Psi}]\exp\left(i\int \mathrm{d}^4x\ \overline{\Psi} i\not{\mathrm{D}}\Psi\right) = \det{}'(i\not{\mathrm{D}}[A])\int \mathrm{d}c_0\mathrm{d}\overline{d}_0.$$

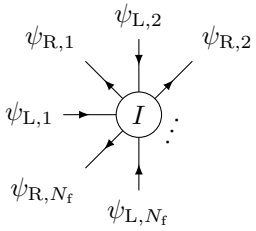
To evaluate it, we first note the behaviour of Grassmann variables

$$\int \mathrm{d}x = 0, \quad \int \mathrm{d}c\, c = 1.$$

The above path integral vanishes, unless there are field insertions into the integrand that give a factor of $c_0\overline{d}_0$. In particular, consider

$$\int \mathcal{D}[\Psi,\overline{\Psi}]\exp\left(i\int \mathrm{d}^4x\;\overline{\Psi}i\not{D}\Psi\right)\overline{\psi}_{\mathrm{R}}(x)\Psi_{\mathrm{L}}(x) = \det{}'(i\not{D}[A])\overline{\psi}_0^{\mathrm{R}}(x)\psi_0^{\mathrm{L}}(x).$$

This corresponds to $\langle 0|\overline{\psi}_{\mathrm{R}}(x)\Psi_{\mathrm{L}}(x)|0\rangle \neq 0$, a chirality violation. This effect is entirely non-perturbative and appears only through instantons. One instanton ($n=1$) destroys a left-handed mode and creates a right-handed fermion, which corresponds to $n_{\mathrm{R}} - n_{\mathrm{L}} = 2N_{\mathrm{f}}n = 2$. This can be depicted as a vertex (in a more general setting with N_{f} flavours)



$$= \text{const.}\times\int\frac{\mathrm{d}\rho}{\rho^5}\frac{1}{g(\rho^{-1})^8}e^{-8\pi^2/g^2(\rho^{-1})}\prod_{f=1}^{N_{\mathrm{f}}}(\overline{\psi}_f^{\mathrm{R}}w)(\overline{w}\psi_f^{\mathrm{L}})e^{i\theta}$$

where w is a constant spinor. This is called the *'t Hooft vertex*. The product on the right equals the determinant of the mass matrix M of fermions

$$\prod_{f=1}^{N_{\mathrm{f}}}(\overline{\psi}_f^{\mathrm{R}}w)(\overline{w}\psi_f^{\mathrm{L}}) = \frac{1}{N_{\mathrm{f}}!}\det M,\quad M_{fg} = (\overline{\psi}_g^{\mathrm{R}}w)(\overline{w}\psi_f^{\mathrm{L}}).$$

The effective Lagrangian $\mathcal{L}_{\mathrm{eff}}$ is invariant under $\mathrm{SU}(N_{\mathrm{f}})_{\mathrm{L}}\times\mathrm{SU}(N_{\mathrm{f}})_{\mathrm{R}}$, since $M\to\mathcal{U}_{\mathrm{R}}^\dagger M\mathcal{U}_{\mathrm{L}}$ but $\det M\to\det M$ because $\det\mathcal{U}_{\mathrm{R}} = \det\mathcal{U}_{\mathrm{L}} = 1$. Under $\mathrm{U}(1)_V\times\mathrm{U}(1)_A$, $\mathcal{L}_{\mathrm{eff}}$ breaks $\mathrm{U}(1)_A$ but remains $\mathrm{U}(1)_V$ invariant.

From this discussion we can infer the CP violation of the strong interaction. We will have to show for the homework that $\theta\neq 0$ implies an electric dipole moment for the neutron. The sketch of the CP violation goes as follows. Assume $N_{\mathrm{f}} = 2$ (so we only consider up and down quarks). The chiral Lagrangian to the second order reads (see Theoretical particle physics notes)

$$\mathcal{L}^{(2)} = \frac{f_\pi^2}{4}\operatorname{tr}(\partial_\mu\mathcal{U}^\dagger\partial^\mu\mathcal{U}) + \frac{f_\pi^2}{2}B_0\operatorname{tr}(M\mathcal{U}+M^\dagger\mathcal{U}^\dagger)$$

where M is the quark mass matrix and

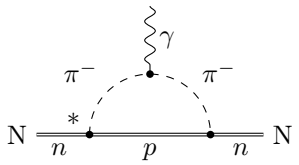
$$\mathcal{U} = \exp\left(-i\frac{\pi^a\sigma^a}{f_\pi}\right),\quad \pi^a = \begin{pmatrix}\pi^0 & \sqrt{2}\pi^+\\ \sqrt{2}\pi^- & -\pi^0\end{pmatrix}.$$

If $\theta\neq 0$, the mass matrix transforms as $M\to M^{-i\theta/2}$. In order to do chiral perturbation theory, we need to find the minimum of the potential $V(\mathcal{U})$ and then expand in π-field around the minimum for $\theta\ll 1$. This induces a CP-violating interaction of π with nucleon field

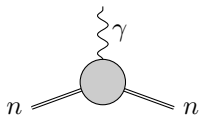
$$\mathcal{L} = -\frac{g_A m_N}{f_\pi}\pi^a\overline{N}\tau^a i\gamma^5 N - \frac{h}{f_\pi}\frac{m_u m_d}{m_u+m_d}\theta\pi^a\overline{N}\tau^a N,\quad N = \begin{pmatrix}p\\ n\end{pmatrix}.$$

The first term is CP conserving and the second term is CP violating. Then the electric dipole

moment of the neutron can be calculated from the following Feynman diagram.



The starred vertex is from the CP-violating term (it could also be on the other side). The interaction between the photon and neutron can be expressed with form factors



$$= i\overline{u}(p')\Gamma^\mu(q^2)u(p)\varepsilon^*_\mu(q),$$

$$\Gamma^\mu = \gamma^\mu F_1(q^2) + \frac{i\sigma^{\mu\nu}q_\nu}{2m_N}F_2(q^2) + \frac{i\sigma^{\mu\nu}q_\nu}{2m_N}\gamma^5 F_3(q^2) + \frac{k^2\gamma^\mu - \not{k}k^\mu}{m_N^2}\gamma^5 F_A(q^2).$$

The electric dipole moment is given by the interaction $\boldsymbol{\sigma}\cdot\mathbf{E}$, so $d_N = F_3(0)/2m_N$. The calculation gives (see homework for details)

$$d_N = \frac{eg_A\theta h m_{ud}}{4\pi^2 f_\pi^2}\ln\left(\frac{4\pi f_\pi}{m_\pi}\right) \approx 4.02\cdot 10^{-16}\,\theta e\,\text{cm}.$$

4.8 Baryon and lepton number violation in the SM

The Standard Model (SM) is a chiral gauge theory. This means that $\mathrm{SU(2)_L}$ and $\mathrm{U(1)}_Y$ don't have matched left- and right-handed fermion representations. There are accidental symmetries in the classical SM Lagrangian, namely baryon number B (from $\mathrm{U(1)}_B$) and lepton number L (from $\mathrm{U(1)}_L$) with (classically) conserved currents

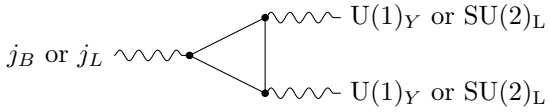
$$j^\mu_B = \sum_{I=1}^{3}\frac{1}{3}(\overline{Q}\gamma_\mu Q - \overline{u}^c\gamma^\mu u^c - \overline{d}^c\gamma^\mu d^c),$$

$$j^\mu_L = \sum_{I=1}^{3}(\overline{L}_I\gamma^\mu L - \overline{e}^c\gamma^\mu e^c).$$

The *anomaly equation* relates the divergence of these currents to the field strengths:

$$\partial_\mu j^\mu_B = \partial_\mu j^\mu_L = \frac{n_\mathrm{G}}{2}\frac{g^2}{16\pi^2}\,\mathrm{tr}(\mathbf{F}\tilde{\mathbf{F}})|_{\mathrm{SU(2)_L}} - \frac{n_\mathrm{G}}{2}\frac{g'^2}{32\pi^2}F_{\mu\nu}\tilde{F}^{\mu\nu}|_{\mathrm{U(1)}_Y}$$

where $n_\mathrm{G} = 3$ is the number of fermion generations. This equation comes from the triangle diagram with one B or L current and two gauge currents:



The symmetry $B - L$ is anomaly free (conserved, $\partial_\mu(j_B^\mu - j_l^\mu) = 0$), but $B + L$ is anomalous. There are no instanton effects related to $\mathrm{U}(1)_Y$, since $\pi_3(\mathrm{U}(1)_Y) = \varnothing$. However, there exist $\mathrm{SU}(2)_\mathrm{L}$ instantons and there are possible processes that violate B and L by n_G units each (one per generation):

$$\Delta B = \Delta L = n_\mathrm{G} n, \quad n \in \mathbb{Z}.$$

There is no factor of 2, because fields are left-handed ($P_\mathrm{L} = (1-\gamma^5)/2$) and the currents already come with a factor of $1/2$. The $\mathrm{SU}(2)_\mathrm{L}$ is spontaneously broken by the Higgs mechanism, so how does it affect instantons? The vacuum expectation value (VEV) of the Higgs field sets a scale $v = 246\,\mathrm{GeV}$ and the only exact solution is the one with $\rho = 0$ as it minimizes the action. The additional term in the action is

$$\int \mathrm{d}^4x\ (\mathrm{D}_\mu\phi)^\dagger(\mathrm{D}^\mu\phi)|_{\phi=(0,v/\sqrt{2})^T} = 2\pi^2 v^2\rho^2,$$

which is clearly minimized by $\rho = 0$. The instanton solution developed in previous chapters for $v = 0$ remains fairly valid even for non-zero v as long as $g \ll 1$. The reasoning is as follows. The classical action gets the additional term

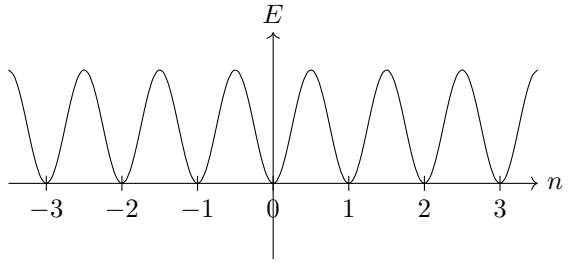
$$e^{-S_\mathrm{cl}} \to \exp\left(-\frac{8\pi^2}{g^2(\rho)} - 2\pi^2 v^2\rho^2\right).$$

When $\rho \geq 1/v$, there is an additional exponential suppression and large instantons are suppressed. Thus, we focus on $\rho < 1/v$. Then we can assume $2\pi^2 v^2\rho^2 \ll 8\pi^2/g^2$, so the symmetry breaking affects the instanton configuration/shape by very little, only at $|x - a| > 1/(gv)$, far away from the core. The IR divergence for large instantons is regulated by v and the amplitude for the instanton process is

$$\exp\left(-\frac{2\pi}{\alpha_{\mathrm{SU}(2)_\mathrm{L}}(v)}\right) \approx e^{-180}.$$

Hence, $B + L$ (and also B and L) violation is non-vanishing, but negligibly small for practical purposes. There is a 't Hooft vertex with 12 fermions: Q_1^i, L_1 for the first generation (index i denotes colour), and 8 more for the second and third.

As the last highlight of this chapter we will again consider the energy barrier between different vacua.



The unbroken theory is classically scale-invariant, so there is not much sense in talking about the energy barrier between vacua. However, with spontaneous symmetry breaking, we can make sense out of it, as it provides a scale v. There exists an approximate, static solution/field configuration with finite energy that corresponds to the maximum of the barrier. This is called the *sphaleron* (from Greek shpaleros ($\sigma\varphi\alpha\lambda\varepsilon\acute{o}\varsigma$) meaning unstable or slippery, the same origin as for sphalerite, a zinc iron sulphide). The sphaleron energy is of order $E_\mathrm{sph} \approx 4\pi v/g \sim 10\,\mathrm{TeV}$. At high enough energies (temperatures), the barrier can be passed by thermally excited particles (not just by

tunnelling). There is unsuppressed $B+L$ violation at high temperatures in the SM (while $B-L$ is still conserved). The gauge group $\mathrm{SU}(2)_\mathrm{L}$ is restored back at energy $T_\mathrm{c} \approx 150\,\mathrm{GeV}$ ($v = 0$) and thermal mass takes the role of the energy scale (and the use of thermal QFT calculations is required). This phenomenon is very important for the SM cosmology: baryon number violation is turned off as the universe cools and the symmetry is broken. Unlike the strong gauge interactions where semi-classical instanton calculations are not reliable, instanton calculus is on solid footing in the broken electroweak gauge theory.

Chapter 5

Large orders and large N

Weak coupling expansions are the basis of most analytic computations in QFT and quantum mechanics. Consider a slight perturbation of exactly solvable system

$$H = H_0 + gH_{\text{int}}$$

where H_0 is exactly solvable and g is small. Any observable will be computed as a series in g:

$$\mathcal{O} = \sum_{n=0}^{\infty} \mathcal{O}_n g^n.$$

Take, for example, the harmonic oscillator, perturbed by a potential of q^4:

$$\mathcal{L} = \frac{1}{2}\dot{q}^2 - V(q), \quad V(q) = \frac{1}{2}q^2 + \frac{\lambda}{4}q^4.$$

We set $\hbar = \omega = 1$. This system is not exactly solvable, however, the ground state energy can be expanded in the parameter $\lambda/4$:

$$E_0(\lambda) = \frac{1}{2} + \frac{3}{4}\left(\frac{\lambda}{4}\right) - \frac{21}{8}\left(\frac{\lambda}{4}\right)^2 + \frac{333}{16}\left(\frac{\lambda}{4}\right)^3 + \dots$$

What is the radius of convergence of such series? To answer this question, consider $E_0(\lambda)$ as a function of complex λ. Is it holomorphic/analytic? Well, for $\lambda < 0$, the hamiltonian is unbounded and the system is unstable. Hence, we don't expect $E_0(\lambda)$ to be analytic function at $\lambda = 0$ and the series is not a regular Taylor series. The classic result for this specific problem was given by Bender and Wu (1972)

$$E_0(\lambda) = \sum_{n=0}^{\infty} e_n \lambda^n, \quad e_n = (-1)^{n+1}\sqrt{\frac{6}{\pi^3}}\left(\frac{3}{4}\right)^n \Gamma(n+1/2)\left(1 - \frac{95}{72n} + \mathcal{O}(n^{-2})\right).$$

We will derive this result later with a different method than Bender and Wu that generalizes to field theory. We see that e_n grows factorially with n, so the *radius of convergence is zero*! Does such series then even make any sense? Mathematically it can always exist as a formal series, but can we extract any physics out of it? This will be the topic of the following sections.

5.1 Asymptotic series and Borel summation

Definition 5.1. A formal series

$$\sum_{n=0}^{\infty} f_n z^{n+1}$$

is said to be *asymptotic* to the function $f(z)$ if and only if for all $z \in \mathbb{C}$ and all $N \in \mathbb{N}$ there exists a constant $C_N > 0$ such that

$$\left| f(z) - \sum_{n=0}^{N-1} f_n z^{n+1} \right| \leq C_N |z|^{N+1} \quad \text{as } z \to 0.$$

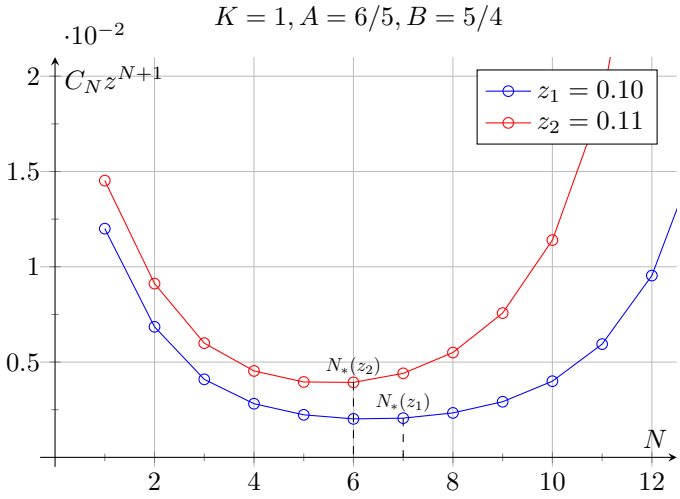
Remark.

(1) Without loss of generality, assume $f(0) = 0$, otherwise replace $f(z) \to f(z) - f(0)$.

(2) This definition includes the case when the series is the Taylor series of some $f(z)$ with a finite radius of convergence r_c. We are interested in the case $r_c = 0$ and C_N of the form

$$C_N = K \cdot A^{-N} N! N^B.$$

This is called *Gevrey-1 class* of asymptotic series (and $E_0(\lambda)$ from before is an element of it). Gevrey classes are spaces of functions between analytic functions and smooth functions. For a fixed N, the approximation is good at $\mathcal{O}(z^{N+2})$ and the error goes down as $z \to 0$. However, for fixed $z \neq 0$, $C_N |z|^{N+1}$ increases for large enough N (when $N > N_*(z)$) when z is outside the radius of convergence. Schematically, the large N behaviour is depicted on the sketch below.



The point where the error is minimal depends on z and for Gevrey-1 class (see expression for C_N above) it is approximately $N_*(z) \approx |A|/z \gg 1$ for small z (the value of z on the sketch above is not small enough for this approximation to hold). We want to expand the function $f(z)$ around $z = 0$ and truncate the series at $N = N_*(z)$ to minimize the error, which is then of order

$$K e^{-A/|z|} \times \text{powers of } z.$$

This is exponentially small in z, thus an asymptotic series can approximate a function very well, depending on the size of A and z.

(3) A formal series can be asymptotic to more than one function (if they differ by exponentially small terms).

For the series expansion of Gevrey-1 class functions, it is useful to define the *Borel transformation* of the series as

$$\sum_{n=0}^{\infty} f_n z^{n+1} \to B[f](t) \equiv \sum_{n=0}^{\infty} \frac{f_n}{n!} t^n.$$

Definition 5.2. The integral

$$\hat{f}(z) = \int_0^{\infty} e^{-t/z} B[f](t)\, \mathrm{d}t$$

is called the *Borel integral* and the *Borel sum* of the series if the integral is well-defined (convergent).

Remark.

(1) The series expansion of $\hat{f}(z)$ is the same as the original series.

(2) The series expansion of $B[f](t)$ has a finite radius of convergence r_c. To see this, suppose $f_n = KA^{-n}\Gamma(n+1+b)$ for a non-negative integer b. Then

$$B[f](t) = \frac{K\Gamma(1+b)}{(1-t/A)^{1+b}}.$$

With the singularity at $t = A$, this has the radius of convergence $r_c = |A|$.

(3) If the series is alternating ($A < 0$), then the Borel integral is well-defined and gives an unambiguous sum. If the series is not alternating ($A > 0$), then $B[f](t)$ has a pole on the positive real axis and the Borel integral is not well-defined.

Suppose $f(z)$ has an asymptotic expansion of Gevrey-1 class with $A > 0$. When is the Borel sum equal to $f(z)$?

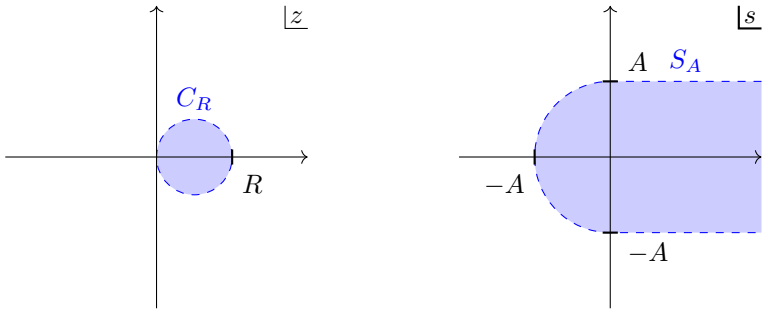
Theorem 5.1 (*Nevalinna–Sokal*). *If $f(z)$ is an analytic function in the sector $C_R = \{z \in \mathbb{C} : \mathrm{Re}\,(1/z) > 1/R\}$ and satisfies the bound*

$$\left| f(z) - \sum_{n=0}^{N-1} f_n z^{n+1} \right| \leq KA^{-N} N! N^B |z|^{N+1}$$

for all $N \in \mathbb{N}$ and all $z \in C_R$, then the series

$$\sum_{n=0}^{\infty} \frac{f_n}{n!} t^n$$

converges for $|t| < A$ and can be analytically continued to the region $S_A = \{t \in \mathbb{C} : \mathrm{dist}(t, \mathbb{R}_+) < A\}$, i.e., $B[f](t)$ is analytic in S_A and $|B(t)| < Ke^{|t|/R}$ uniformly. Then the Borel integral converges absolutely for $z \in C_R$ and equals $f(z)$ in C_R.



It is very hard to show that a field theory is Borel summable and in fact, there is no example in 4D where this has been rigorously shown.

Let's look at some examples of Borel summable divergent series.

(1) The *Striling formula*:

$$\ln\Gamma(z+1) - \left[(z+1/2)\ln(z) - z + \frac{1}{2}\ln(2\pi)\right] = \sum_{n=1}^{\infty} \frac{B_{2n}}{2n(2n-1)z^{2n-1}}$$

where B_{2n} are Bernoulli numbers

$$B_{2n} = (-1)^{n+1}\frac{2(2n)!}{(2\pi)^{2n}}\zeta(2n) \xrightarrow{n\to\infty} (-1)^{n+1}\frac{2(2n)!}{(2\pi)^{2n}}.$$

This is a sign-alternating asymptotic expansion for $z \to \infty$. The left-hand side can be expressed as a Borel integral

$$\ln\Gamma(z+1) - \left[(z+1/2)\ln(z) - z + \frac{1}{2}\ln(2\pi)\right] = \int_0^{\infty} \mathrm{d}t\, e^{-t/z}\frac{1}{t}\left(\frac{1}{e^t-1} - \frac{1}{t} + \frac{1}{2}\right).$$

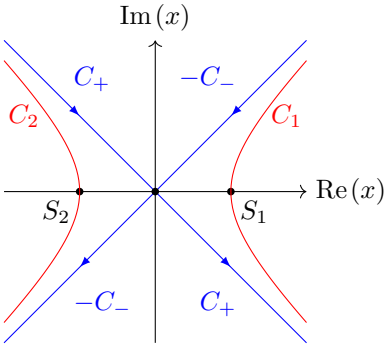
In this case, $A = \pm 2\pi i$.

(2) Consider the partition function of the anharmonic oscillator in zero dimensions

$$Z(\lambda) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{d}x\, \exp\left(-\frac{1}{2}x^2 - \frac{\lambda}{4}x^4\right) = 1 + \mathcal{O}(\lambda).$$

One can analytically continue $Z(\lambda)$ to complex λ by rotating the contour of integration in the complex x plane such that $\mathrm{Re}\,(\lambda x^4/4) > 0$. To satisfy this, it has to hold $\arg(x) = -\arg(\lambda)/4$. Our goal is to use the dispersion relation to find the large order behaviour of the perturbative expansion coefficients. The function $Z(\lambda)$ is analytic in the complex λ plane with a cut along the negative real axis. To calculate the discontinuity across the cut, we take contour C_+ (a straight line at angle $-\pi/4$ to the real axis) for $\lambda = -|\lambda| + i\varepsilon$ and contour C_- (line at angle $+\pi/4$ to the real axis) for $\lambda = -|\lambda| - i\varepsilon$. The contours are shown on the sketch below. We need to calculate

$$\mathrm{Disc}(Z) = Z(\lambda + i\varepsilon) - Z(\lambda - i\varepsilon) = \frac{1}{\sqrt{2\pi}}\int_{C_+ - C_-} \mathrm{d}x\, \exp\left(-\frac{x^2}{2} - \frac{\lambda x^4}{4}\right).$$



To evaluate the integral, we will perform the saddle point approximation. The exponent has three saddle points ($x+\lambda x^3=0$): $S_0=0$ and $S_{1,2}=\pm 1/\sqrt{-\lambda}$. The contribution from S_0 cancels out in $C_+ - C_-$, so we can deform the contours to C_1 and C_2 (see sketch above) that pass through S_1 and S_2, respectively. The value of the exponent at the saddle points is $-1/(4\lambda)$, so the leading contribution to the discontinuity for $\lambda<0$ is

$$\mathrm{Disc}\, Z(\lambda) = 2i\mathrm{Im}\; Z(\lambda) \propto i e^{1/(4\lambda)}.$$

For small λ, the discontinuity is exponentially small. Now we can use the once-subtracted dispersion relation (see chapter on On-shell methods from QFT2 notes)

$$Z(\lambda) = 1 + \frac{\lambda}{\pi}\int_{-\infty}^{0} \mathrm{d}\lambda' \, \frac{\mathrm{Im}\; Z(\lambda')}{\lambda'(\lambda'-\lambda)} = \sum_{n=0}^{\infty} z_n \lambda^n,$$

where

$$z_n = \frac{1}{\pi}\int_{-\infty}^{0} \mathrm{d}\lambda' \, \frac{\mathrm{Im}\; Z(\lambda')}{(\lambda')^{n+1}} \propto (-4)^n (n-1)!(1+\mathcal{O}(1/n)).$$

The $\mathcal{O}(1/n)$ correction comes from the $\mathcal{O}(\lambda)$ error of the saddle point approximation. This expression is Borel summable and gives us a recipe for computing factorial divergences.

(3) The anharmonic oscillator in QM is also Borel summable.

If the factorial divergence is not alternating ($f_n = KA^{-n}\Gamma(n+1+b)$ and $A>0$), then Borel summation is not well-defined. However, we can define the Borel sum by deforming the contour of integration above or below the pole on the positive real axis and define the *lateral Borel integrals*

$$\hat{f}_{\pm}(z) = \lim_{\theta\to 0^{\pm}} \int_0^{e^{i\theta}\cdot\infty} \mathrm{d}t \, e^{-t/z} B[f](t).$$

For real f and $z>0$, the singularity at $t=A$ results in an imaginary part of the integral

$$\mathrm{Im}\; \hat{f}_{\pm}(z) = \mp(\pi K A) e^{-A/z} (z/A)^b,$$

which means that there is an exponentially small "ambiguity" in the lateral Borel sums.

Stokes discontinuities

Let $f(z)$ be a function of complex $z=|z|e^{i\varphi}$. The asymptotic expansion (denoted by tilde $\sim$) depends on the phase φ and changes discontinuously for some values of φ. This requires the asymptotic expansion to include exponentially small terms. Suppose

$$f(z) \sim \sum_n f_n z^{n+1} + G(z) \sum_n g_n z^{n+1}$$

where $G(z)$ is exponentially small in z. Then the series $g_n z^{n+1}$ changes discontinuously at those phases φ, where f_n have equal sign asymptotically. These phase lines in the complex z-plane are called *Stokes lines*.

Example 6. Let

$$f(z) = \sum_{n=0}^{\infty} (-1)^n \frac{\psi(n)}{n! z^n}, \quad \psi(x) = \frac{\mathrm{d}}{\mathrm{d}x} \ln \Gamma(x).$$

The function $\psi(x)$ is called the *digamma function* and is the logarithmic derivative of the Gamma function. The above defines an analytic function in $\mathbb{C} \setminus \{0\}$. The series can be related to the exponential integral function:

$$f(z) = e^{-1/z} \left(\mathrm{Ei}(-1/z) - \ln(-1/z) - \gamma_{\mathrm{E}} \right), \quad \mathrm{Ei}(z) = \int_{-\infty}^{z} \frac{e^t}{t} \, \mathrm{d}t.$$

For small z, the complete asymptotic expansion is

$$f(z) \sim \sum_{n=0}^{\infty} (-1) n! z^{n+1} + e^{-1/z} \underbrace{\ln(-1/z)}_{-\ln(z) \pm i\pi}, \quad \arg(z) \neq 0.$$

- Stokes line is at $\arg(z) = 0$ with an exponentially small discontinuity $\mp \pi e^{-1/z}$ from the second part of the expression above.
- The first part is a non-alternating factorially divergent series, so we can define it via lateral Borel summation. Then the $\mathrm{Im}\, \hat{f}_{\pm}(z)$ exactly cancels the Stokes discontinuity.

The lateral Borel ambiguity and the Stokes discontinuity cancel each other out, giving an unambiguous result which coincides with the original function $f(z)$ in $\mathbb{C} \setminus \{0\}$.

The systematization of these ideas leads to the field of *resurgence theory* and *transseries*. The exploration of these topics is one of the problems on the homework sheet.

5.2 Large order behaviour in quantum mechanics

5.2.1 Anharmonic oscillator

We are interested in the large-order behaviour of the ground state energy of an anharmonic oscillator in one-dimensional QFT (i.e., quantum mechanics). We will employ the path-integral methods. The Hamiltonian corresponding to the Lagrangian introduced at the beginning of the chapter is

$$H = -\frac{1}{2} \left(\frac{\mathrm{d}}{\mathrm{d}q} \right)^2 + \frac{1}{2} q^2 + \frac{\lambda}{4} q^4.$$

- The ground state energy $E_0(\lambda)$ is an analytic function in the cut complex plane (it is similar to the zero-dimensional field theory studied at the beginning of the chapter).
- For large $|\lambda|$, the ground state energy behaves as $E_0(\lambda) \propto \lambda^{1/3}(1 + \mathcal{O}(\lambda^{-2/3}))$. To show this, introduce $q' = q\lambda^{1/6}$. Then

$$H = \lambda^{1/3} \left(-\left(\frac{\mathrm{d}}{\mathrm{d}q'} \right)^2 + \frac{1}{2} \frac{1}{\lambda^{2/3}} q'^2 + \frac{1}{4} q'^4 \right).$$

Denote what is in outer parenthesis with H' and treat the quadratic term as a perturbation for $|\lambda| \to \infty$. It follows

$$E_0(\lambda) = \lambda^{1/3} E_0'(\lambda) \propto \lambda^{1/3} \quad \text{for } |\lambda| \to \infty.$$

This allows us to write down a once-subtracted dispersion relation

$$E_0(\lambda) = \frac{1}{2} + \frac{\lambda}{\pi} \int_{-\infty}^{0} \mathrm{d}\lambda' \, \frac{\operatorname{Im} E_0(\lambda')}{\lambda'(\lambda' - \lambda)}.$$

For details, see QFT2 notes. In short, we use a Cauchy theorem to express the value of $E_0(\lambda)$ as the integral of a contour around λ. Then we deform the contour to go around the branch cut (which gives the discontinuity – or equivalently the imaginary part) and then closes the complex plane with a big circle at infinity. Ideally, we don't want the big arc to contribute to the integral. This would happen, if the asymptotic growth of $E_0(\lambda)$ was slower than constant. However, the behaviour is $\lambda^{1/3}$, so we perform the subtraction $E_0(\lambda) - E_0(0)$. This improves the convergence properties of the big circle part of the contour to asymptotic growth below λ^1 (which holds for us). The $1/2$ in the expression above is exactly the subtracted energy at $\lambda = 0$.

Now we expand both sides of the dispersion relation in a series and compare the coefficients

$$e_n = \frac{1}{\pi} \int_{-\infty}^{0} \mathrm{d}\lambda \, \frac{\operatorname{Im} E_0(\lambda)}{\lambda^{n+1}}.$$

This expression is exact; for large n the integral will be dominated by small λ. Hence, to calculate the large order behaviour we need $\operatorname{Im} E_0(\lambda)$ at small, negative λ.

- We can obtain the ground-state energy from the Euclidean-time evolution operator or partition function (thermal QFT):

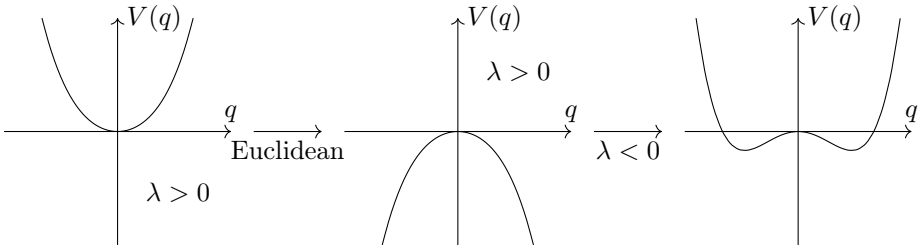
$$e^{-\beta E_0} \xleftarrow{\beta \to \infty} \sum_i e^{-\beta E_i} = \operatorname{tr}(e^{-\beta H}) = \int \mathrm{d}q \, \langle q | e^{-\beta H} | q \rangle = \mathcal{N} \int\limits_{q(-\beta/2) = q(\beta/2)} \mathcal{D}[q[\tau]] e^{-S_{\mathrm{E}}[q(\tau)]},$$

where

$$S_{\mathrm{E}}[q] = \int_{-\beta/2}^{\beta/2} \mathrm{d}\tau \, \left(\frac{1}{2} \dot{q}^2 + \frac{1}{2} q^2 + \frac{\lambda}{4} q^4 \right).$$

The dot denotes the derivative with respect to the Euclidean time τ. The potential can be thought of as

$$V(q) = \frac{1}{\lambda} \tilde{V}(\sqrt{\lambda} q), \quad \tilde{V}(Q)0 = \frac{1}{2} Q^2 + \frac{1}{4} Q^4.$$



For $|\lambda| \ll 1$, the path integral is dominated by its extrema (the classical solutions of motion), that is the motion in the potential on the right, subject to periodic boundary conditions. There are three constant solutions to $-\ddot{q} + q + \lambda q^3 = 0$:

$$q(\tau) = 0, \quad q(t) = \pm\frac{1}{\sqrt{-\lambda}}$$

and the oscillatory solution around the left or right minimum. The first static solution is independent of λ and so it doesn't contribute to the branch cut. The other two static solutions have $S_{\text{cl}} = -\beta/(4\lambda)$. Hence,

$$e^{-S_{\text{cl}}} = e^{\beta/(4\lambda)} \xrightarrow{\beta\to 0} 0.$$

We are left with oscillatory solutions, which have to satisfy

$$\frac{1}{2}\dot{q}^2 - \frac{1}{2}q^2 - \frac{\lambda}{4}q^4 = C < 0.$$

The solution is

$$\beta = 2\int_{q_<}^{q_>} \frac{\mathrm{d}q}{\sqrt{q^2 + \lambda q^4/2 + 2C}}$$

where β is the period of oscillation and $q_>, q_<$ are the turning points. For $\beta \to \infty$, the motion corresponds to a particle staring with infinitesimal velocity in the central local maximum and bouncing from right or left. This requires $C = 0$, $q_<0$ and $q_> = \pm\sqrt{-2/\lambda}$. The classical solution is

$$q_{\text{cl}}(\tau) = \pm\sqrt{\frac{2}{-\lambda}}\frac{1}{\cosh(\tau - \tau_0)}.$$

The free parameter τ_0 is a collective coordinate from the translation invariance. At $\tau = \tau_0$, the generalized coordinate q assumes its extremal value of $q_>$. This is exactly the solution from the discussion of vacuum decay and bounce (see homework 6). For the same reasons as there, this produces a negative eigenvalue and hence a non-vanishing Im $E_0(\lambda)$ for negative λ. The calculation follows exactly as in the chapter on Instantons and in one of the previous homework problems. The saddle point approximation works, as always, by expanding the action around the classical solution:

$$S_{\text{E}} = S[q_{\text{cl}}] + \frac{1}{2}\int \mathrm{d}\tau\,\mathrm{d}\tau'\,(q(\tau) - q_{\text{cl}}(\tau))\frac{\delta^2 S_{\text{E}}}{\delta q(\tau)\delta q(\tau')}(q(\tau') - q_{\text{cl}}(\tau')) + \dots$$

Then we perform the Gaussian path integral except for the zero mode, which must be treated separately. The normalization is given by the simple harmonic oscillator ($\omega = 1, \lambda = 0$). The result is

$$\frac{\text{Im tr}\, e^{-\beta H}}{\text{tr}\, e^{-\beta H_0}} = \underbrace{\frac{1}{i}}_{1} \cdot \underbrace{\beta\sqrt{\frac{S_{\text{cl}}}{2\pi}}}_{2} \cdot \underbrace{e^{-S_{\text{cl}}}}_{3} \underbrace{\left(\frac{\det' M[q_{\text{cl}}]}{\det M_0}\right)^{-1/2}}_{4} \cdot(1 + \mathcal{O}(\lambda)).$$

The four terms are explained as follows.

1 From Im $(\bullet) = ((\bullet)|_{\lambda+i\varepsilon} - (\bullet)|_{\lambda-i\varepsilon})/(2i)$ and a factor 2 to account for the two bounces (left and right).

2 From the zero mode integration over the collective coordinate τ_0.

3 This term comes from the classical action at the bounce

$$S_{\text{cl}} = \lim_{\beta\to\infty}\int_{-\beta/2}^{\beta/2} \mathrm{d}\tau\,\dot{q}_{\text{cl}}^2 = -\frac{4}{3\lambda} > 0.$$

4 From the Gaussian path integral around the classical solution. The operator M_0 is the fluctuation operator around the trivial vacuum

$$M_0 = -\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} + 1,$$

and $M[q_{\mathrm{cl}}]$ is the fluctuation operator around the bounce solution (take $\tau_0 = 0$)

$$M[q_{\mathrm{cl}}] = -\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} + 1 - \frac{6}{\cosh^2\tau}.$$

The determinant can be solved using the Gel'fand–Yaglom theorem (see chapter on Instantons). This is a specific example of *Pöschl–Teller potential*

$$V_{\mathrm{PT}}(x) = m^2 - \frac{a(a+1)}{\cosh^2 x}.$$

The Schrödinger problem with this potential is exactly solvable and the functional determinant (denote $M_a^m = -\mathrm{d}_\tau^2 + m^2 - a(a+1)\operatorname{sech}^2\tau$) has a rather compact form

$$\frac{\det(M_a^m)}{\det(M_0^m)} = \frac{\Gamma(1+m)\Gamma(m)}{\Gamma(1+a+m)\Gamma(m-a)}.$$

Notice that we cannot simply plug $m = 1$ and $a = 2$ into this expression as we would get a zero for answer (gamma function of a negative integer in the denominator). This is exactly expected, because the operator has a zero eigenvalue. If M_a^m has a zero eigenvalue, then $M_a^m + \varepsilon$ has an eigenvalue ε. Since we know that M_2^1 has exactly one zero-eigenvalue and M_0^1 has none, we can remove the zero eigenvalue by shifting the potentials by ε, dividing by ε and taking the limit $\varepsilon \to 0$:

$$\begin{aligned}\frac{\det'(M_2^1)}{\det(M_0^1)} &= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\frac{\det(M_2^1+\varepsilon)}{\det(M_0^1+\varepsilon)} = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\frac{\det(M_2^{\sqrt{1+\varepsilon}})}{\det(M_0^{\sqrt{1+\varepsilon}})} = \lim_{\varepsilon\to 0}\frac{\Gamma(1+\sqrt{1+\varepsilon})\Gamma(\sqrt{1+\varepsilon})}{\Gamma(3+\sqrt{1+\varepsilon})\Gamma(\sqrt{1+\varepsilon}-2)}\\ &= -\lim_{\varepsilon\to 0}\frac{1}{\varepsilon\Gamma(\varepsilon)\Gamma(4)} = -\frac{1}{12}.\end{aligned}$$

The value is negative, as expected.

Combining all four terms, one obtains

$$\frac{\operatorname{Im}\ \operatorname{tr} e^{-\beta H}}{\operatorname{tr} e^{-\beta H_0}} = -\frac{4\beta}{\sqrt{-2\pi\lambda}}e^{4/(3\lambda)}(1+\mathcal{O}(\lambda)).$$

When we take the limit $\beta \to \infty$, the left-hand side becomes

$$\begin{aligned}\frac{\operatorname{Im}\ e^{-\beta E_0}}{e^{-\beta/2}} &= e^{\beta/2}e^{-\beta\mathrm{Re}\,(E_0)}\mathrm{Im}\,(e^{-i\beta\mathrm{Im}\,(E_0)})\\ &= e^{\beta/2}e^{-\beta(1/2+\mathcal{O}(\lambda))}\mathrm{Im}\,(e^{-i\beta\mathrm{Im}\,(E_0)})\\ &= -\beta\mathrm{Im}\,(E_0(\lambda))(1+\mathcal{O}(\lambda)).\end{aligned}$$

Thus, it holds

$$\operatorname{Im}\ E_0(\lambda)|_{\lambda<0} = \frac{4}{\sqrt{-2\pi\lambda}}e^{4/(3\lambda)}(1+\mathcal{O}(\lambda))$$

and finally

$$e_n = \frac{1}{\pi}\frac{4}{\sqrt{2\pi}}\int_{-\infty}^{0}\mathrm{d}\lambda\,\frac{e^{4/(3\lambda)}}{\sqrt{-\lambda}\lambda^{n+1}} = (-1)^{n+1}\left(\frac{6}{\pi^3}\right)^{1/2}\left(\frac{3}{4}\right)^n\Gamma(n+1/2)(1+\mathcal{O}(1/n)).$$

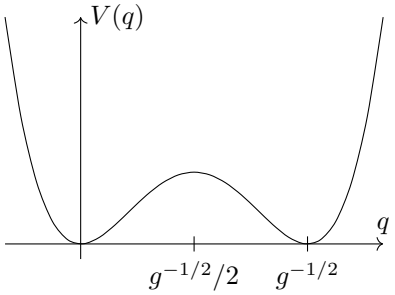
The result is the same as quoted in the beginning of the chapter (Bender and Wu, 1972).

5.2.2 Degenerate ground state – the double well potential

Consider the following Hamiltonian

$$H = -\frac{1}{2}\left(\frac{\mathrm{d}}{\mathrm{d}q}\right)^2 + V(q) = -\frac{1}{2}\left(\frac{\mathrm{d}}{\mathrm{d}q}\right)^2 + \frac{1}{2}q^2(1-\sqrt{g}q)^2.$$

This is essentially the same potential as in section Section 4.4.3, with a slight change of conventions.



Section 4.4.3	here
a	$g^{-1/2}/2$
g	$2g$
$\omega^2 = 2a^2 g$	1

Two degenerate ground states in the absence of tunnelling spontaneously break $\mathbb{Z}_2$ symmetry. Tunnelling leads to an exponentially small energy splitting and a unique ground state which is $\mathbb{Z}_2$ symmetric. As we have seen in Section 4.4.3, the classical action of a bounce is $S_{\mathrm{cl}} = 1/(6g)$ and the energies of two (nearly) degenerate ground states are

$$E_\pm(g) = \underbrace{\sum_{n=0}^{\infty} e_n g^n}_{E_{\pm,\mathrm{pert.}}} \mp \frac{1}{\sqrt{\pi g}} e^{-1/(6g)}(1+\mathcal{O}(g)).$$

Since we expect large-order behaviour to be related to exponentially small terms, the periodic paths are related to instanton-anti-instanton configurations (or even multi-instanton). Define

$$\operatorname{tr} e^{-\beta H} = e^{-\beta E_+} + e^{-\beta E_-} + \text{exited oscillator states' configuration}$$

and the same with parity operator P

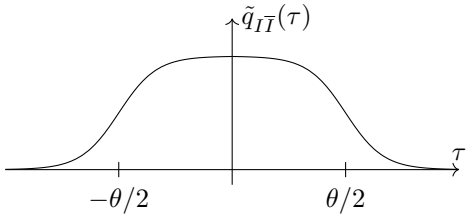
$$\operatorname{tr} P e^{-\beta H} = e^{-\beta E_+} - e^{-\beta E_-} + \dots$$

The contributions to the first expression come from periodic paths $q = 0 \to q = 0$ or $q = 1/\sqrt{q} \to q = 1/\sqrt{g}$. These contributions have to start and finish at the same point, so they have to be a combination of equal number of instantons and anti-instantons. For the parity ordered expression, the paths with different endpoints contribute ($0 \to 1/\sqrt{g}$ or $1/\sqrt{g} \to 0$). As we have already calculated, the (anti-)instanton centred at $\tau = 0$ has the following form

$$q_I(\tau) = \frac{1}{\sqrt{g}}\frac{1}{1+e^{-\tau}}, \quad q_{\bar{I}}(\tau) = \frac{1}{\sqrt{g}}\frac{1}{1+e^{\tau}}.$$

The instanton-anti-instanton configuration is not stable, but in the limit of infinite separation it approaches solves the classical equations of motion. Consider an instanton at $\tau = -\theta/2$ and an anti-instanton at $\tau = \theta/2$:

$$\tilde{q}_{I\bar{I}}(\tau) = \begin{cases} \frac{1}{\sqrt{g}}\frac{1}{1+e^{-(\tau+\theta/2)}}, & \tau < 0; \\ \frac{1}{\sqrt{g}}\frac{1}{1+e^{\tau-\theta/2}}, & \tau > 0. \end{cases}$$



The action of this configuration is

$$\tilde{S}_{I\bar{I}} = \frac{1}{g}(1/3 - e^{-\theta} + \mathcal{O}(e^{-2\theta})).$$

Since this is not stationary for finite θ, we try to construct a family of configurations with parameter θ that is as stationary as possible, i.e., a "valley" in configuration space in the sense that variations in orthogonal directions are larger:

$$q_{I\bar{I}}(\tau) = \tilde{q}_{I\bar{I}}(\tau) + \delta q_{I\bar{I}}(\tau).$$

We determine δq by requiring the linear term in the action expanded around $\tilde{q}_{I\bar{I}}(\tau)$ to vanish. The result is

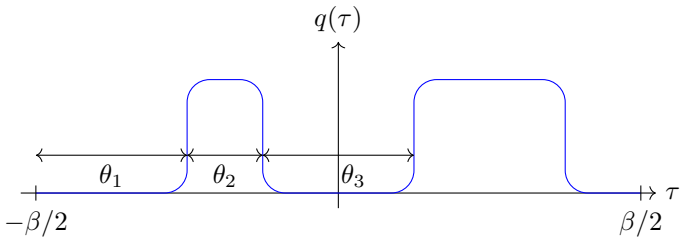
$$q_{I\bar{I}} = q_I(\tau + \theta/2) + q_{\bar{I}}(\tau - \theta/2) - 1/\sqrt{g},$$

that is the sum of instanton and anti-instanton configuration. The corresponding action is

$$S_{I\bar{I}} = \frac{1}{g}(1/3 - 2e^{-\theta} + \mathcal{O}(e^{-2\theta})).$$

Since the term $e^{-\theta}$ contains a factor of 2, this new solution is better as the action is smaller. We can generalize this calculation to n (anti-)instantons on a finite time interval β. If we denote the distance between the centres of two adjacent (anti-)instantons by θ_i, the action reads

$$S_n = \frac{1}{g}\left(\frac{n}{6} - 2\sum_{i=1}^{n} e^{-\theta_i}\right), \quad \sum_{i=1}^{n} \theta_i = \beta.$$



In the case of one instanton and one anti-instanton, the action on finite interval becomes $1/3 - 2e^{-\theta} - 2e^{-\beta+\theta}$.

The path integral is evaluated in the dilute instanton gas approximation, which assumes that all θ_i are large and the determinant of the fluctuation operator is the product of its values on a

single (anti-)instanton. This leads to the following expression:

$$\operatorname{tr}\left(\frac{1+\varepsilon P}{2}e^{-\beta H}\right)_{n\text{ jumps}} = e^{-\beta/2}\cdot\frac{\beta}{n}\cdot\left(\frac{\varepsilon e^{-1/(6g)}}{\sqrt{\pi g}}\right)^n\cdot\mathcal{J}_n(g);$$

$$\mathcal{J}_n(g) = \iiint\limits_{\theta_i\geq 0}\prod_{i=1}^n \mathrm{d}\theta_i\,\delta\left(\beta-\sum_{i=1}^n\theta_i\right)\exp\left(\frac{g}{2}\sum_{i=1}^n e^{-\theta_i}\right)$$

where $\varepsilon=\pm 1$. The first term $e^{-\beta/2}$ comes from the harmonic oscillator's zero point energy and will be cancelled by the normalization $\operatorname{tr} e^{-\beta H_0}$. The factor β in the second term is from global time translation and time integral. The $1/n$ comes from n equivalent cyclic permutations of instanton-anti-instanton that were over-counted (remember that time is compacted into a loop). The last term, $\mathcal{J}_n(g)$, is the integral over collective coordinates. The Jacobian $\sqrt{S_I}^n$ is already included in the third term. The above expression is a compact way of writing that only even n contribute to $\operatorname{tr} e^{-\beta H}$ and only odd n to $\operatorname{tr} Pe^{-\beta H}$. If we ignore the instanton-anti-instanton interaction and take

$$\exp\left(\frac{2}{g}\sum_{i=1}^n e^{-\theta_i}\right)\to 1,$$

then the last term simplifies to $\mathcal{J}_n(g)=\beta^{n-1}/(n-1)!$ and the sum over all n of the energies gives $\exp(-\beta E_\varepsilon(g))$, where

$$E_\varepsilon(g)=E_\pm(g)=\frac{1}{2}+\frac{\varepsilon}{\sqrt{\pi g}}e^{-1/(6g)}.$$

Hence, in the absence of interaction, the dilute instanton gas is simply the exponential of the one instanton contribution (and the energy is not affected).

The $I\bar{I}$ interaction is attractive because the action decreases with θ. All multi-instanton configurations are unstable. The way we derived $\mathcal{J}_n(g)$ is valid only if the integral is dominated by large θ_i. This is true for negative g (repulsive interaction), therefore we define $\mathcal{J}_n(g)$ by analytic continuation from $g<0$ to $g>0$. The θ_i integrals can be factorized by writing

$$\delta\left(\beta-\sum_{i=1}^n\theta_i\right)=\frac{1}{2\pi i}\int_{-c-i\infty}^{-c+i\infty}\mathrm{d}s\,\exp\left(-s\beta+s\sum_{i=1}^n\theta_i\right)$$

for $c>0$ and closing the contour in the right half plane. The expression for $\mathcal{J}_n(g)$ becomes

$$\mathcal{J}_n(g)=\frac{1}{2\pi i}\int_{-c-i\infty}^{-c+i\infty}\mathrm{d}s\,e^{-\beta s}I(g,s)^n,\quad I=\int_0^\infty \mathrm{d}\theta e^{s\theta+(2/g)\exp(-\theta)}.$$

For small g the integral in I can be evaluated to

$$I=\left(-\frac{2}{g}\right)^s\Gamma(-s)(1+\mathcal{O}(e^{2/g})).$$

Then

$$\sum_{n=1}^\infty\operatorname{tr}\left(\frac{1+\varepsilon P}{2}e^{-\beta H}\right)_n=-\frac{\beta e^{-\beta/2}}{2\pi i}\int_{-c-i\infty}^{-c+i\infty}\mathrm{d}s\,e^{-\beta s}F(g,s)$$

where

$$F(g,s)=\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{\varepsilon e^{-1/(6g)}}{\sqrt{\pi g}}\right)^n I^n(g,s)=\ln\left(1-\frac{\varepsilon e^{-1/(6g)}}{\sqrt{\pi g}}\left(-\frac{2}{g}\right)^s\Gamma(-s)\right)\equiv\ln(f(s)).$$

The integral expression above is the inverse *Mellin transform* of the function $F(g,s)$ (closing the contour on the right). It seems to me more like inverse Laplace transform or inverse two-sided Laplace transform . After the integration by parts, the Cauchy integral picks up the contribution from the poles of $f'(s)/f(s)$ which are located at the zeros of $f(s)$. For small g, they are close to the poles of $\Gamma(-s)$ (to compensate for exponentially small $e^{-1/(6g)}$), at $s_N(g)=N+\#e^{-1/(6g)}/\sqrt{\pi g}$ for $N=0,1,\dots$. Thus, we get

$$\frac{e^{-\beta/2}}{2\pi i}\int_{-c-i\infty}^{-c+i\infty}\mathrm{d}s\,e^{-\beta s}F(g,s)=\sum_{N=0}^{\infty}e^{-\beta(1/2+s_N(g))}=\sum_{N=0}^{\infty}e^{-\beta E_n(g)}.$$

The ground state $E_\pm(g)$ corresponds to $N=0$. To get the N-th energy, write

$$E_N(g)=\sum_{k=0}^{\infty}E_N^{(k)}(g)(e^{-1/(6g)})^k$$

and solve for $f(s_N)=0$ iteratively in powers of $e^{-1/(6g)}$. This leads to the following result (for $N=0$, $\varepsilon=\pm$):

$$E_\pm(g)=\sum_{n=0}^{\infty}e_n g^n\mp\frac{1}{\sqrt{\pi g}}e^{-1/(6g)}(1+\mathcal{O}(g))+\frac{1}{\pi g}e^{-1/(3g)}\left(\ln(-2/g)+\gamma_{\mathrm{E}}+\mathcal{O}(g\ln g)\right)+\dots+$$
$$+\left(\frac{1}{\sqrt{\pi g}}e^{-1/(6g)}\right)^k P_k(\ln(-2/g))\cdot\sum_{l=0}^{\infty}f_{kl}g^l+\dots$$

The term proportional to $e^{-1/(3g)}$ corresponds to instanton-anti-instanton contribution and because the logarithm, it has a branch cut at positive g. This is because $\overline{I}I$ interaction is attractive. Large-order behaviour of $E_{\pm,\mathrm{pert}}$ will be related with the dispersion relation to the imaginary part of the instanton-anti-instanton contribution across the branch cut

$$\mathrm{Im}\;E_\pm(g)=-\frac{1}{g}e^{-1/(3g)}(1+\mathcal{O}(g))$$

and so

$$e_n=\frac{1}{\pi}\int_0^\infty\mathrm{d}g\frac{\mathrm{Im}\;E_\pm(g)}{g^{n+1}}=\frac{-1}{\pi}3^{n+1}n!.$$

This expansion has a fixed sign, so it is not Borel summable. This means we need to take the lateral Borel transforms of $E_{\pm,\mathrm{perp}}(g)$ and interpret the expansion of $E_+(g)$ in $g,\ln(-2/g),e^{-1/(6g)}$ as a transseries, where all series should be interpreted as lateral Borel sums, whose Stokes discontinuity together with the one from $P_k(\ln(-2/g))$ make $E_+(g)$ unambiguous and well-defined for positive g. The same holds for $E_-(g)$.

Lesson we can learn from this example is that whenever we see a fixed sign divergence of perturbative series, it indicates we should explicitly add exponentially small non-perturbative terms – their size must be of order of the Stokes discontinuity.

5.3 Extension to quantum field theory

5.3.1 ϕ^4 theory

The action of ϕ^4 theory in d dimensions is

$$S[\phi] = \int \mathrm{d}^d x \left(\frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2} m^2 \phi^2 - \frac{\lambda}{4} m^{4-d} \phi^4 \right).$$

Consider the Euclidean action for $\lambda < 0$. The non-trivial minimal action configuration is described by the following ansatz

$$\phi(x) = \frac{m^{d/2-1}}{\sqrt{-\lambda}} f(mr), \quad r = |x - x_0|$$

where $f(r)$ satisfies

$$\left(-\frac{\mathrm{d}^2}{\mathrm{d}\hat{r}^2} + \frac{d-1}{\hat{r}} \frac{\mathrm{d}}{\mathrm{d}\hat{r}} + 1 \right) f(\hat{r}) - f^3(\hat{r}) = 0, \quad \hat{r} = mr.$$

This can be only solved numerically and from this S_{cl} can be calculated

$$S_{\mathrm{cl}} = \frac{1}{-\lambda} \frac{1}{4} \int \mathrm{d}u \, f_{\mathrm{cl}}(u)^4.$$

It has imaginary part for $\lambda < 0$ which is exponentially small in $-1/\lambda$. The singularity in the Borel transform is located at $-1/S_{\mathrm{cl}}$ and one can show mathematically that perturbation series are Borel summable (they satisfy Nevanlinna–Sokal requirements) for $d < 4$. At $d = 4$ the theory passes from super-renormalizable to renormalizable, but we can still say something. For $m = 0$ the theory is scale invariant. The ansatz is

$$\phi(x) = \frac{1}{\sqrt{-\lambda}} f(r), \quad \left(-\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{3}{r} \frac{\mathrm{d}}{\mathrm{d}r} \right) f(r) - f^3(r) = 0.$$

This now has analytic solution

$$\phi_{\mathrm{cl}}(x) = \pm \frac{1}{\sqrt{-\lambda}} \frac{2\sqrt{2}}{\rho} \frac{1}{1 + (x - x_0)^2/\rho^2}$$

where ρ is the size of the instanton (a collective coordinate as a consequence). The action is $S[\phi_{\mathrm{cl}}] = 8\pi^2/(-3\lambda)$. The running of the parameter λ follows

$$\lambda(1/\rho) = \frac{\lambda(\mu_0)}{1 + \beta_0 \lambda(\mu_0) \ln(\mu_0^2 \rho^2)}.$$

In ϕ^4, $\beta_0 > 0$ and λ grows for large instantons. The correlation functions $\langle \Omega | \phi(x_1) \dots \phi(x_n) | \Omega \rangle$ are still calculable due to $1/\rho$ in ϕ_{cl} which make the integral over ρ convergent (this is not guaranteed for the ground-state energy). The theory in $d = 4$ looks Borel-summable, however there is another obstruction to Borel summability, which we will discuss later (renormalons).

5.3.2 Yang–Mills theory, QCD

Non-trivial, non-perturbative vacuum structure of Yang–Mills theory leads to instantons as classical solutions of the Euclidean field equations with finite action. The situation resembles the double well potential. Hence, we expect non-Borel summability due to fixed sign an $n!$'s related to $S_{I\overline{I}} = 2S_I = 16\pi^2/g^2$ and transseries structure with exponentially small correction in powers of $e^{-16\pi^2/g^2}$. However, there are other, more important sources of factorial divergences in renormalizable QFTs.

5.4 Large-N expansion on two-dimensional QFTs

Let N be the number of field components, together with $\mathrm{O}(N)$ or $\mathrm{U}(N)$ internal symmetry. The idea is to expand in $1/N$ but not in small coupling. We will do this in various models, starting with two-dimensional QFT and finishing with QCD.

5.4.1 Gross–Neveu model

The *Gross–Neveu model* is a two-dimensional ($d = 1 + 1$) QFT of N massless fermions with quartic interaction and $\mathrm{U}(N)$ symmetry. The Lagrangian density is

$$\mathcal{L} = \overline{\psi}_i i\partial\!\!\!/ \psi_i + \frac{g}{2}(\overline{\psi}_i\psi_i)^2, \quad i = 1, 2, \ldots, N.$$

There are two gamma matrices

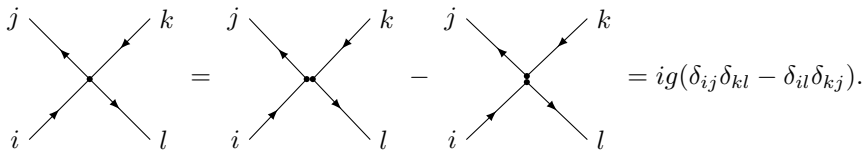
$$\gamma^0 = \sigma^1, \quad \gamma^1 = i\sigma^2, \quad \{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}, \quad \gamma_5 = \gamma^0\gamma^1 = -\sigma^3, \quad \{\gamma_5, \gamma^\mu\} = 0, \quad \gamma_5^2 = \mathbb{1}.$$

- In perturbation theory, there are N massless fermions.
- There is a $\mathbb{Z}_2$ "chiral symmetry" $\psi_i \to \gamma_5\psi_i$, $\overline{\psi}\psi \to \psi^\dagger\gamma_5\psi^0\gamma^5\psi = -\overline{\psi}\psi$ and $\overline{\psi}\partial\!\!\!/\psi \to \overline{\psi}\partial\!\!\!/\psi$.

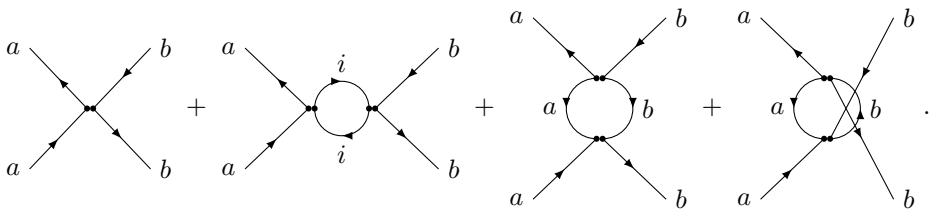
The dimension of the field is $[\psi] = 1/2$, the coupling is dimension-less $[g] = 0$ and the full model is perturbatively renormalizable.

Gross–Neveu model in perturbation theory

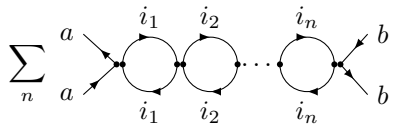
The Feynman rules are as follows



Consider the process $\psi_a\overline{\psi}_a \to \psi_b\overline{\psi}_b$ for $a \neq b$. The next-to-leading order diagrams are

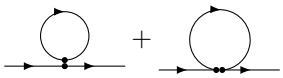


The second diagram comes with a factor of N while the other two loop diagrams only have a factor of 1. That means that in the large-N limit, the second diagram will contribute most. For large N, we have

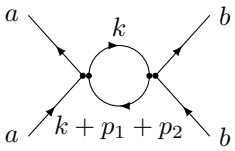


$$\sim \sum_n g(gN)^n \sim \frac{1}{N}\sum_n \lambda^{n+1}.$$

When we take $N \to \infty$, we keep $\lambda = gN$ fixed. The coupling λ is called the *'t Hooft coupling*, which does not need to be small in the large-N limit. In the perturbation theory (and renormalization) we do require $\lambda \ll 1$. There is no field renormalization at one loop because the corresponding graphs have momentum independent divergences.



There is also no mass renormalization because of chiral symmetry. The coupling renormalization comes from UV divergence of the four-point function:



$$= (-1)(ig)^2 N i^2 \tilde{\mu}^{2\varepsilon} \int \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{\mathrm{tr}(\not{k}(\not{k}+\not{p}_1+\not{p}_2))}{k^2(k+p)^2}$$

$$\overset{\text{UV}}{=} -2g^2 N \tilde{\mu}^{2\varepsilon} \int \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{1}{k^2 - m^2 + i\varepsilon}$$

$$\overset{\text{pole}}{=} ig\frac{gN}{2\pi\varepsilon} \equiv -ig\delta_g.$$

We use the bare coupling $g_0 = \tilde{\mu}^{2\varepsilon} g Z_g$, where

$$Z_g^{\text{1-loop}} = 1 + \delta_g = 1 - \frac{gN}{2\pi\varepsilon}.$$

To all orders in g and for large N, we extract the result $1/Z_g = 1 + gN/(2\pi\varepsilon)$ and the beta function is

$$\beta(g) = \mu^2 \frac{\mathrm{d}g(\mu)}{\mathrm{d}\mu^2} = -\varepsilon g - \frac{Ng^2}{2\pi} = -\varepsilon g + \beta_0 g^2.$$

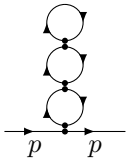
Hence, $\beta_0 = -N/(2\pi)$ which is asymptotically free. The coupling runs according to the equation

$$g(\mu) = \frac{g(\mu_0)}{1 + \frac{g(\mu)N}{2\pi}\ln(\mu^2/\mu_0^2)} \equiv \frac{1}{-\beta_0 \ln(\mu^2/\Lambda^2)},$$

where $\Lambda^2 = \mu^2 e^{-1/(\beta_0 g(\mu))}$.

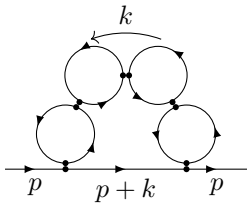
Large-order behaviour

The leading set of graphs of the ψ-field is the set of "snowman diagrams" which all vanish due to the scaleless loops

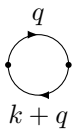


$$= 0.$$

The full set of graphs at $\mathcal{O}(1/N)$ which is not scaleless is the set of "chain graphs".



A single loop is



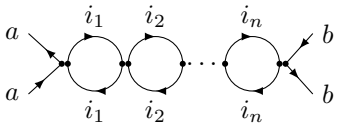
$$= ig\Pi(k^2), \quad \Pi(k^2) = (-1)\frac{gN}{4\pi}\left(\frac{-k^2+i0_+}{\mu^2}\right)^{-\varepsilon}\frac{\Gamma(1+\varepsilon)\Gamma(-\varepsilon)^2}{\Gamma(-2\varepsilon)}e^{\varepsilon\gamma_{\mathrm{E}}} + \text{counterterms}.$$

When we send $\varepsilon \to 0$, we obtain

$$\Pi(k^2) = -\frac{gN}{2\pi}\ln\left(-\frac{k^2}{\mu^2}\right).$$

Summing all chain graphs gives the following $2 \to 2$ amplitude ($s = (p_1+p_2)^2$)

$$iT_{2\to 2} = \sum_{n=0}^{\infty}$$



$$\sim \sum_{n=0}^{\infty} g(gN)^n \sim \frac{1}{N}\sum_n \lambda^{n+1} = \frac{ig(\mu)}{1-\Pi(s)} = \frac{ig(\mu)}{1-\frac{gN}{2\pi}\ln(-s/\mu^2)} = ig(\mu = -\sqrt{s}),$$

which is just the running of the coupling. Returning to the chain diagram with n loops, we have

$$I_n = ig\left(-\frac{gN}{2\pi}\right)^n \int \frac{\mathrm{d}^d k}{(2\pi)^d}\frac{\not{p}+\not{k}}{(p+k)^2}\ln^n\left(-\frac{k^2}{\mu^2}\right).$$

The logarithms become large for $k_{\mathrm{E}}^2 \gg p^2$ and $k_{\mathrm{E}}^2 \ll p^2$. Hence, we split the integral at $k_{\mathrm{E}}^2 = \mu^2$. In the IR, expand in k^2/p^2:

$$I_n^{\mathrm{IR}} = ig\left(-\frac{gN}{2\pi}\right)^n \frac{\not{p}}{p^2} i\int_0^{\mu^2}\frac{\mathrm{d}^2 k_{\mathrm{E}}}{(2\pi)^2}\ln^n(k_{\mathrm{E}}^2/\mu^2) = \not{p}\frac{\mu^2}{p^2}g\left(\frac{gN}{2\pi}\right)^n n! = \not{p}\frac{\mu^2}{p^2}(-\beta_0)^n g^{n+1} n!$$

where we used $x = -\ln(k_{\mathrm{E}}^2/\mu^2)$, $\mathrm{d}k_{\mathrm{E}} = \mu^2 e^{-x}\mathrm{d}x$ and obtained a factorial from the integrand $(e^{-x}x^n)$.

- This is not Borel summable. There is a pole in Borel transform at $t = -1/\beta_0$. Higher orders in k^2/p^2 give terms such as $e^{-mx}x^n$ which give poles at $t = -m/\beta_0$. Singularities appear for all $m \in \mathbb{N}$, and are called *IR renormalon* singularities.
- From UV expansion (p^2/k^2) one gets similar factorials but with sign alternation, so these are Borel summable. Singularities appear at $t = m/\beta_0$ and are called *UV renormalon* singularities.

Instanton singularities are related to factorial growth in the number of Feynman diagrams. However, renormalon singularities come from a single diagram, and are related to logarithms in the integrands. Singularities involve renormalizing group function coefficients, in particular β_0, then $\beta_1, \gamma_0, \beta_2, \gamma_1, \ldots$ Such singularities appear only in renormalizable theories (not super-renormalizable) – which is exactly the theories relevant to nature.

Since infrared renormalons are not Borel summable, they come with natural exponentially small ambiguity. This implies that the perturbative expansion is incomplete by exponentially small terms (use $p = \mu$)

$$e^{m/(\beta_0 g(p))} = e^{-m \ln(p^2/\Lambda^2)} = \left(\frac{\Lambda^2}{p^2}\right)^m$$

where Λ is the strong interaction scale. Since $g(p) \ll 1$ it holds $\lambda^2/p^2 \ll 1$, so this is non perturbative-power expansion.

5.4.2 $1/N$ expansion, non-perturbative

To perform the non-perturbative expansion, we introduce an auxiliary field σ in the path integral

$$Z[J] = \mathcal{N} \int \mathcal{D}[\psi, \overline{\psi}, \sigma] \exp\left(i \int \mathrm{d}^d x (\mathcal{L}_\sigma(\psi, \overline{\psi}, \sigma) + J\sigma)\right) \quad \mathcal{L}_\sigma = \overline{\psi} i \partial\!\!\!/ \psi_i - \frac{1}{2g}\sigma^2 + \sigma \overline{\psi}_i \psi_i.$$

The path integral over σ gives the equation of motion that sets $\sigma = g\overline{\psi}_i \psi_i$ which restores the four-fermion interaction. However, we will instead integrate over $\mathcal{D}[\psi_i, \overline{\psi}_i]$:

$$Z[J] = \mathcal{N} \int \mathcal{D}[\sigma] \, \exp\left(i S_{\text{eff}}[\sigma] + i \int \mathrm{d}^d x \, J\sigma\right), \quad S_{\text{eff}} = -\int \mathrm{d}^d x \, \frac{\sigma^2}{2g} + iN \operatorname{tr}(\ln(i\partial\!\!\!/ - \sigma)).$$

To calculate extremum of the effective action, we write the stationarity condition

$$\left.\frac{\delta S_{\text{eff}}}{\delta \sigma}\right|_{\sigma=\sigma_{\text{cl}}} = 0$$

which gives (in momentum space)

$$\frac{1}{g_0 N} = \frac{1}{\sigma_{\text{cl}}} \int \frac{\mathrm{d}^d k}{(2\pi)^d} \operatorname{tr} \frac{1}{k\!\!\!/ - \sigma_{\text{cl}}}.$$

The integral on the right is divergent, and the coupling on the left is bare. For the integral, we have

$$\int \frac{\mathrm{d}^d k}{(2\pi)^d} \operatorname{tr} \frac{1}{k\!\!\!/ - \sigma_{\text{cl}}} = i\sigma_{\text{cl}} \int \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{2}{k^2 - \sigma_{\text{cl}}} = \frac{1}{2\pi} \Gamma(\varepsilon) \left(\frac{\sigma_{\text{cl}}^2}{\mu^2}\right)^{-\varepsilon} e^{\varepsilon \gamma_{\text{E}}} = \frac{1}{2\pi\varepsilon} - \ln \frac{\sigma_{\text{cl}}^2}{\mu^2}.$$

The $1/\varepsilon$ pole cancels exactly with the coupling renormalization at large N ($1/(g_0 N) = 1/(gN) + 1/(2\pi\varepsilon)$), giving

$$\sigma_{\text{cl}}^2 = \mu^2 e^{-2\pi/(N g(\mu))}.$$

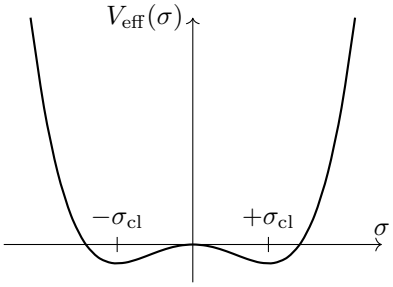
This is exponentially small even in 't Hooft coupling $\lambda = gN = \text{const}$. The effective potential at leading order in $1/N$ is

$$V_{\text{eff}} = \frac{\sigma_{\text{cl}}^2}{2g_0} + iN \int \frac{\mathrm{d}^d p}{(2\pi)^d} \ln(-p^2 + \sigma_{\text{cl}}^2) \sim \frac{\sigma_{\text{cl}}^2}{2g_0} + N\sigma_{\text{cl}}^2 \ln \frac{\sigma_{\text{cl}}^2}{\mu^2}.$$

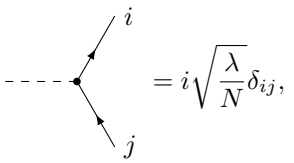
This means that $\langle\sigma\rangle$ has a non-vanishing vacuum expectation value. Chiral symmetry is thus dynamically broken, $\langle\overline{\psi}_i\psi_i\rangle \neq 0$ (like in QCD) and $\sigma_{\rm cl}\overline{\psi}_i\psi_i \to m\overline{\psi}_i\psi_i$ (when we expand $\sigma = \sigma_{\rm cl} + \hat{\sigma}$). The fermions thus acquire a mass

$$m = \mu e^{-\pi/(Ng(\mu))} = \Lambda$$

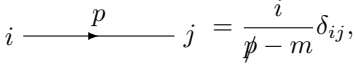
which is exactly the interaction scale. The $\overline{\psi}\psi$ interaction, no matter how small g_0 will generate fermion mass and break chiral symmetry dynamically.



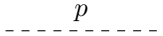
We want to find the Feynman rules in the $1/N$ expansion. We will work with $\mathcal{L}_\sigma$ and shift $\sigma \to \sigma_{\rm cl} + \hat{\sigma}$. The fluctuation is rescaled by $\hat{\sigma} \to \sqrt{g}\hat{\sigma} = \sqrt{\lambda/N}\hat{\sigma}$. There is one vertex and two types of propagators: the fermion propagator and the auxiliary field propagator:



$$= i\sqrt{\frac{\lambda}{N}}\delta_{ij},$$



$$= \frac{i}{\not{p} - m}\delta_{ij},$$



$$= D(p).$$

We need to construct the auxiliary field propagator by taking the second derivative of the effective action

$$iD^{-1}(x,y) = \frac{1}{N}\frac{\delta^2 S_{\rm eff}}{\delta\sigma(x)\delta\sigma(y)}\bigg|_{\sigma=\sigma_{\rm cl}} = i\,\langle x|\,\mathrm{tr}(1/(i\not{\partial} - m))^2|y\rangle - \frac{1}{gN}\delta^{(2)}(x-y).$$

Again, it is more convenient to work in the momentum space

$$iD^{-1}(p) = -\frac{1}{gN} - \frac{1}{2\pi\varepsilon} + i\int\frac{\mathrm{d}^d k}{(2\pi)^d}\frac{\mathrm{tr}((\not{k}+m)(\not{k}+\not{p}+m))}{(k^2-m^2+i\epsilon)((p+k)^2-m^2+i\epsilon)} = -\frac{1}{gN} + B(p)$$

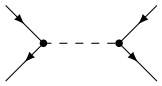
where $B(p)$ is the renormalized massive bubble graph. The propagator is then equal to the sum over all chain diagrams for the $2 \to 2$ scattering

$$D(p) = \frac{i}{-1/(gN) + B(p)} \overset{d\to 2}{=} \frac{2\pi}{i}\frac{1}{\xi\ln\frac{\xi+1}{\xi-1}}, \quad \xi = \sqrt{1 - \frac{4m^2}{p^2}}.$$

Note that $-1/(g(\mu)N) = -1/(2\pi)\ln(\mu^2/m^2)$, so for $m^2 \to 0$ it holds

$$D(p) \to \frac{1}{i}\frac{2\pi}{\ln(-p^2/m^2)} = ig(\sqrt{-p^2}) = iT_{2\to 2}.$$

The scattering $\psi\overline{\psi} \to \psi\overline{\psi}$ with the auxiliary field is more than the perturbative bubble sum, because the fermions now have a non-perturbative mass $m = \mu \exp(-\pi/\lambda(\mu))$. In the following calculation, we will adopt $\mu^2 = -p^2$ and $\xi = \sqrt{1+4u}$, $u = e^{-2\pi/\lambda(\rho)}$. We expand the $2 \to 2$ scattering amplitude for small u:



$$= iT_{2\to 2} = i\left(\frac{2\pi}{\ln u} + \sum_{k=1}^{\infty} u^k \sum_{l=1}^{k+1} \frac{c_{kl}}{\ln^l(1/u)}\right).$$

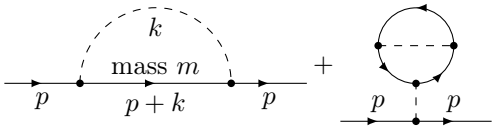
The first term $2\pi/\ln u$ is $\lambda(\sqrt{-p^2})$, and the following two terms are

$$u^k = \left(e^{-2\pi/(\lambda(\sqrt{-p^2}))}\right)^k, \quad \frac{1}{\ln^l(1/u)} = (\lambda(\sqrt{-p^2}))^l.$$

This expansion does not grow factorially!

Self-energy non-perturbatively and transseries

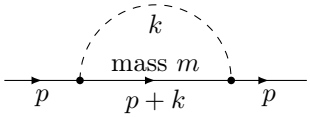
At $\mathcal{O}(1/N)$ there are two diagrams that contribute to the self energy



The second graph only contributes to the mass part of the self-energy Σ_m (since the tadpole is independent of p)

$$\Sigma(p, m) = m\Sigma_m(p^2, m^2) + \not{p}\Sigma_p(p^2, m^2).$$

We only focus on Σ_p (because we found renormalons in the perturbative expansion only for Σ_p). The first diagram gives



$$= \frac{\lambda}{N} \int \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{i(\not{p} + \not{k} + m)}{(k+p)^2 - m^2} \frac{2\pi i}{\xi(k) \ln \frac{\xi(k)+1}{\xi(k)-1}}.$$

We expand this expression for small λ. This is very technical and involves Mellin integral representation for σ-propagator in terms of $(m^2/p^2)^s$ and pick up the singularities. The structure of the result is

$$\Sigma_p(p^2, m) = \sum_{k=0}^{\infty} \left(-\frac{m^2}{p^2}\right)^k \int_0^{\infty} \mathrm{d}t \left(e^{-t/\lambda(p)} \left[\frac{F_k(t)}{\lambda(p)} + G_k(t)\right] + H_k(t)\right),$$

where F_k, G_k and H_k are some functions that are not relevant for understanding the structure.

- The term $(-m^2/p^2)^k$ correspond to the exponentially small terms $(e^{-2\pi/\lambda(p)})^k$.
- The part with F_k and G_k is the Borel integral which can be expanded in powers of $\lambda(p)$ giving perturbative series around each non-perturbative exponential term. Both F_k and G_k have poles at $t = -n/\beta_0 > 0$, which give rise to IR renormalons.
- The integral over H_k is independent of λ, but it also has singularities at $t = -n/\beta_0$.

However, the entire expression has no singularities on the t-contours and is well-defined. Defining all t integrals as lateral Borel transforms

$$\lim_{\theta\to 0_\pm}\int_0^{e^{i\theta}\cdot\infty}\mathrm{d}t,$$

the imaginary poles/ambiguities in the individual terms cancel. This happens between terms with different k, e.g. for $t=-1/\beta_0$ will be cancelled between F_0, G_0 and H_2.

5.4.3 Two-dimensional theories with scalar fields

The $\mathrm{O}(N)$ non-linear sigma model

The Lagrangian of the $\mathrm{O}(N)$ non-linear sigma model is

$$\mathcal{L}=\frac{1}{2g}\partial_\mu\sigma^i\partial^\mu\sigma^i,\quad \sum_{i=1}^N\sigma^i\sigma^i=N.$$

The field space is S^{N-1} and has $\mathrm{O}(N)$ symmetry. We can introduce an auxiliary field $\alpha(x)$ to impose the constraint (as a Lagrange multiplier) and rescale $\sigma^i\to\sqrt{g}\sigma^i$:

$$\mathcal{L}[\sigma^i,\alpha]=\int\mathrm{d}^dx\,\frac{1}{2}((\partial_\mu\sigma^i)^2-\alpha(\sigma^i\sigma^i-N/g)).$$

We are working in dimensional regularization $d=2-2\varepsilon$.

- The theory is asymptotically free, $\beta_0<0$.
- In $1/N$ expansion $\langle\alpha\rangle\neq 0$, which gives a mass term for σ that is exponentially small in g.
- The field equation $(\partial^2+\alpha)\sigma^i=0$ has the solution $\alpha=(g/N)\sigma_i\partial^2\sigma_i$ which gives rise to non-perturbative contributions (transseries)

The $\mathbb{C}P^{N-1}$ model

This is the model with N complex scalar fields z_i with constraint

$$\sum_{i=1}^N z_i^\dagger z_i=N.$$

This is captured the same as in the $\mathrm{O}(N)$ model by introducing an auxiliary field α. However, there is an additional constraint that the overall phase of the fields is unphysical. Two vectors $\mathbf{z}=(z_1,\ldots,z_n)$ and $\mathbf{z}'$ are considered the same if they differ by a phase. This defines a complex $(N-1)$-dimensional projective space. This constraint is captured by constructing a gauge symmetry

$$A_\mu=\frac{i}{2}\left(z^\dagger\partial_\mu z-(\partial_\mu z^\dagger)z\right),\quad \mathcal{L}=\frac{1}{g}(\mathrm{D}_\mu z_i)^\dagger(\mathrm{D}_\mu z_i).$$

5.5 Large-N QCD

In this section we are talking about $\mathrm{SU}(N)$ gauge theory with n_f flavours of Dirac fermions in the fundamental representation. The beta function is

$$\mu^2\frac{\mathrm{d}\alpha_\mathrm{s}}{\mathrm{d}\mu^2}=\beta(\alpha_\mathrm{s})=\beta_0\alpha_\mathrm{s}^2+\ldots,\quad \beta_0=-\frac{1}{4\pi}\left(\frac{11N}{3}-\frac{2n_\mathrm{f}}{3}\right).$$

When we take the large-N limit, we keep $\hat{\alpha} = \alpha_s N$ constant, i.e., we rescale the coupling as $g \to g/\sqrt{N}$. Then

$$\beta_0 = -\frac{1}{4\pi}\left(\frac{11}{3} - \frac{2}{3}\frac{n_f}{N}\right) \xrightarrow{N\to\infty} -\frac{11}{12\pi} < 0.$$

5.5.1 Four dimensions

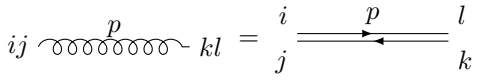
The (rescaled) Lagrangian in four dimensions is

$$\mathcal{L}_{\text{QCD}} = \frac{N}{g^2}\left(\frac{1}{4}\operatorname{tr}(\mathbf{F}_{\mu\nu}\mathbf{F}^{\mu\nu}) + \overline{\Psi}(i\partial\!\!\!/ + A\!\!\!/)\Psi - m\overline{\Psi}\Psi\right).$$

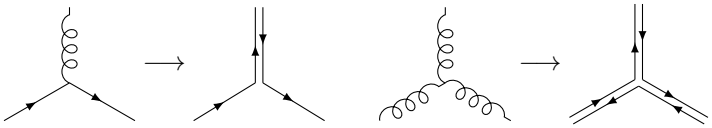
Keep in mind that here we use a rather unconvetional normalization $\operatorname{tr}(T^A T^B) = \delta^{AB}$. When the coupling parameter was rescaled to $g/\sqrt{N}$, fields were also rescaled, with the factor of $\sqrt{N}/g$. The propagators can then be written as follows:

$$\begin{aligned}\Psi_i(x)\overline{\Psi}_j(y) &= \delta_{ij}S(x-y),\\ (A_\mu)^i_j(x)(A_\nu)^k_l(y) &= A^a_\mu(x)A^b_\nu(y)T^{ai}{}_j T^{bk}{}_l = (\delta^i_l\delta^k_j - \delta^i_j\delta^k_l/N)D_{\mu\nu}(x-y).\end{aligned}$$

The gauge field propagator was rewritten using the Fierz identities and second term is subleading in $1/N$, so it can be neglected. Gauge fields can be thought of as carrying two colours (so one would have red-green gluon, in reality they can be linear combinations of two-colour combinations) and the propagator can in general swap the two colours. In large-N limit, the mixing term is suppressed, so gluons can be represented by a double line connecting i to l and j to k.



The vertices become



When we count the power of N in the diagram, we follow these rules:

- each vertex (V) gets a factor of N (this comes from one N in front of the whole Lagrangian),
- each propagator (P) gets a factor of $1/N$ (because it is, in a sense, inverted term in a Lagrangian),
- each index loop (L) gets a factor of N (for N different colours, the number of index loops might not be equal to the number of momentum loops).

Thus, the exponent of N can be calculated as $L + V - P$. This exactly equals the *Euler characteristic* χ of the topological space on which can the diagram be embedded. Generally, graphs can be embedded on tori of some genera – a torus of genus h (having h holes) has Euler characteristic of $\chi = 2(1-h)$. The main results are

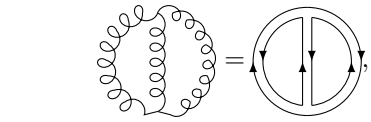
(1) Leading (large-N) connected vacuum graphs with only gluon lines are of $\mathcal{O}(N^2)$ (because of the double line representation of a gluon there will always be an index loop around a vacuum gluon bubble, which forces us to count the exterior of the loop as a face and topologically this makes the plane a sphere; Euler characteristic of a sphere is $\chi(S^2) = 2$, so all planar vacuum gluon bubbles contribute at order N^2).

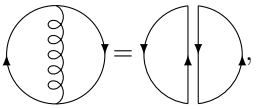
(2) A replacing a gluon loop with a quark loop costs a factor of $1/N$. This effectively removes one index loop (and hence one face), so the embedding space becomes a disc. Disc has Euler characteristic $\chi(D^2) = 1$, so leading graphs with one quark loop are of $\mathcal{O}(N)$. We have to be careful though, as we can allow gluons only on one side of the fermion circle (most commonly inside).

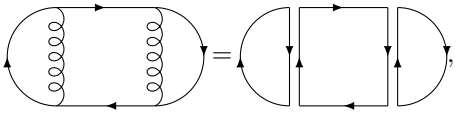
(3) The leading graphs are planar. Once we identify one leading graph (which is usually simple and planar), we know that all graphs embedded in the same space (usually the plane) will contribute equally.

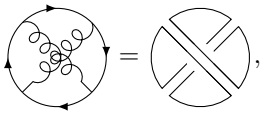
(4) Any sub-leading graph will actually be suppressed by at least $1/N^2$. This is because to change the topology of the embedding space we have to either add a handle (which reduces χ by 2) or add another quark loop (which reduces χ by 1, but also forces us to add gluons on both sides of the fermion loop, which reduces χ by another 1).

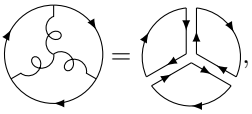
Let's look at some examples.


$V - P + L = 2 - 3 + 3 = 2,$


$V - P + L = 2 - 3 + 2 = 1$


$V - P + L = 4 - 6 + 3 = 1,$


$V - P + L = 4 - 6 + 1 = -1$ (non-planar).


$V - P + L = 4 - 6 + 1 = 1$

To talk about mesons, we introduce quark bilinears $M_s(x) \sim \overline{q}\Gamma_s q$, which we add to the Lagrangian with sources $J_s(x)$:

$$\mathcal{L}_{\text{QCD}} \to \mathcal{L}_{\text{QCD}} + N \sum_{s=1} J_s(x) M_s(x).$$

The factor N was added so that this part contributes at the same order as the rest of the Lagrangian $\mathcal{O}(N)$, but the normalization will change this factor anyway. The correlation function of n mesons is

$$\langle \Omega | M_{s_1}(x_1) \dots M_{s_n}(x_n) | \Omega \rangle \equiv \frac{1}{N^n} \frac{1}{i^n} \frac{\delta W}{\delta J_{s_1}(x_1) \dots \delta J_{s_n}(x_n)} \bigg|_{J=0}$$

where $W[J] = -i \ln Z[J]$ is the generating functional of connected graphs. The leading contribution to this correlation function seems to be of order $\mathcal{O}(1/N^{n-1})$, however a properly normalized

meson state comes with a factor of $\sqrt{N}$, $\tilde{M}_s = \sqrt{N} M_s$, such that $\langle \tilde{M}_{s_1} \tilde{M}_{s_2} \rangle \sim 1$. Hence, the n-point meson amplitude comes with

$$\langle \Omega | \tilde{M}_{s_1}(x_1) \dots \tilde{M}_{s_n}(x_n) | \Omega \rangle \sim N^{1-n/2}.$$

From this we can conclude that

(1) mesons are weakly interacting with strength $\mathcal{O}(1/\sqrt{N})$,

(2) mesons are stable at $N \to \infty$ since their decay width is $\mathcal{O}(1/N)$ (decay amplitude of two mesons into one contributes $N^{-1/2}$, so the decay time is $\sim N$),

(3) for every spin/flavour quantum number there must be an infinite number of meson bound states of arbitrarily large mass. The two point function can be decomposed by Källén–Lehmann spectral decomposition

$$\langle \Omega | \tilde{M}_s \tilde{M}_s | \Omega \rangle (k) = \sum_{\ell = \text{meson states}} \frac{f_\ell^2}{k^2 - M_\ell^2},$$

where f_ℓ are some constants related to the field strength renormalization of the field. There's no continuum, as all mesons are stable. To support the last claim, suppose that the spectrum of M_s was bounded. Then there exists some scale Λ such that $M_\ell^2 < \Omega^2$ for all ℓ. Then the two-point function would behave as $\langle \tilde{M}_s \tilde{M}_s \rangle \propto 1/k^2$. This would be in conflict with asymptotic freedom, which requires $\langle \tilde{M}_s \tilde{M}_s \rangle \propto \ln(-k^2/\mu^2)$ for large k^2.

(4) Similar results hold for glueballs, which are bound states of gluons. However, the real QCD is different, because mesons interact strongly and are not stable.

5.5.2 Two dimensions

The two-dimensional QCD is called *'t Hooft model*. In this model, the gauge bosons don't have propagating degrees of freedom (but rather they mediate an instantaneous force). The theory is confining, even for Abelian gauge group. Consider the field strength tensor in the axial gauge ($A_1 = 0$, $\mu = 1$ is the spatial direction)

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

The non-linear term vanishes ($f^{abc} A^b_0 A^c_0 = 0$) so even non-Abelian gauge field comes without self-interaction. For any vector (including x^μ) define $a^\pm = (a^0 \pm a^1)/\sqrt{2}$. The *light-cone gauge* is $(A^a)^+ = 0$. In the light-cone coordinates, the metric is

$$\mathrm{d}s^2 = g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu = 2\,\mathrm{d}x^+\,\mathrm{d}x^-, \quad g_{\mu\nu} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \equiv \begin{pmatrix} 0 & g_{+-} \\ g_{-+} & 0 \end{pmatrix}.$$

The gamma matrices obey $(\gamma^+)^2 = (\gamma^-)^2 = 0$ and $\{\gamma^+, \gamma^-\} = 2$. For the gauge field it holds $A_- = 0$ and $A_+, A^- \neq 0$. In $d = 2$, the light-cone gauge is covariant gauge: $\Omega^\mu_\nu = \delta^\mu_\nu + \omega^\mu_\nu$ where $\omega_{\mu\nu}$ is antisymmetric. Since light-cone gauge is covariant, we will work with it rather than axial gauge. Hence, $\delta A^+ = \omega^{10} A_- = 0$. The Lagrangian in light-cone gauge is

$$\mathcal{L}_{\text{QCD}} = \frac{1}{2} \operatorname{tr}\big((\partial_- A_+)^2\big) + \overline{\Psi}(i\gamma^\mu \partial_\mu - m + \frac{g}{\sqrt{N}} \gamma^+ A_+)\Psi.$$

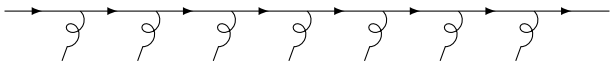
From this Lagrangian the following Feynman rules apply.

- Gluon propagator:
$$D^{ab}_{\mu\nu}(k) = i\delta_{ab}\delta_{\mu+}\delta_{\nu-}\frac{P}{k_-^2} \to i\frac{P}{k_-^2},$$
where P denotes the principal value prescription which follows from the Fourier transform of the instantaneous linear potential in position space $\overline{\Psi}\Psi(x)|x_- - y_-|\overline{\Psi}\Psi(x)$. The expression on the right is a simplified version that will be explained later. Where does this come from? In axial gauge, the only non-zero component of the gauge field is temporal part, and in the equations of motion only the spatial derivative remains: $\partial_x^2 A_0(x,t) = gJ(x,t)/\sqrt{N}$. This is 1D Poisson equation, which can be solved using Green's function:
$$A_0(x,t) = \frac{g}{2\sqrt{N}}\int \mathrm{d}x'\,|x-x'|J(x',t).$$
The force acts instantaneously, without the retarded time (as there are no propagating gluons). The Hamiltonian is
$$H = \frac{g}{2\sqrt{N}}\int \mathrm{d}x\,\mathrm{d}x'\,J(x,t)|x-x'|J(x',t),$$
so the potential is $V(r) = |r|$ and the force is confining.

- Quark propagator:
$$i\delta_{ij}\frac{p_+\gamma^+ + p_-\gamma^- + m}{2p_+p_- - m^2 + i\varepsilon} \to \frac{ip_-}{2p_+p_- - m^2 + i\varepsilon}.$$

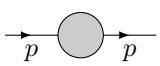
- Quark-gluon vertex:
$$\frac{ig}{\sqrt{N}}\gamma^+ T^a_{ij} \to -\frac{2ig}{\sqrt{N}}.$$

To justify the simplified propagators, consider a quark line with many gluons attached.



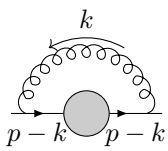
The colour algebra is already satisfied with the large-N limit, so we can drop all deltas. The numerator of the expression is $\gamma^+(\not{p}+m)\gamma^+\ldots\gamma^+$. This can be simplified using $\gamma^+\gamma^+ = \gamma^+\gamma^+\gamma^+ = 0$ and $\gamma^+\gamma^-\gamma^+ = 2\gamma^+$. Hence, only the p_- term in the numerator survives ($\gamma^+(\not{p}+m)\gamma^+ = 2p_-\gamma^+$), giving the simplified propagator. The factor of 2 goes to the vertex.

The exact quark propagator in $N \to \infty$ limit consists of only planar graphs, geometrically summing all planar one-particle-irreducible (1PI) self-energy insertions:

$S(p) =$ 

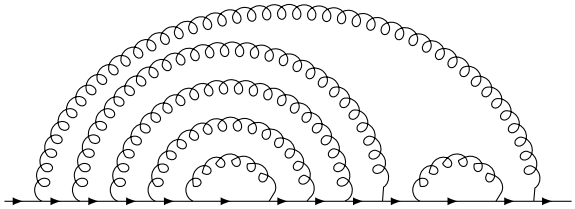
$= \dfrac{i}{\not{p} - m - \not{p}\Sigma - m\Sigma' + i\varepsilon}.$

Since $\Sigma(p)$ comes exclusively from 1PI planar graphs, it can only come from the amputated *rainbow diagrams*:

$-i\Sigma(p) =$ 

$= (-2ig/\sqrt{N})^2 N \int \dfrac{\mathrm{d}k_+\,\mathrm{d}k_-}{(2\pi)^d} S(p-k)\, i\dfrac{P}{k_-^2}.$

The factor of N comes from the colour loop. By iteration, the gluons can only attach to the fermion, and they cannot cross.



The definitions of $S(p)$ and $\Sigma(p)$ are recursive, which means they satisfy some integral equation. But we can do better.

- $-\not{p}\Sigma - m\Sigma'$ can be replaced by $-p_-\Sigma$ with the simplified Feynman rules.
- Shift $k_+ \to k_+ + p_+$ in the integral. Then the right-hand side does not depend on p_+. Hence, Σ does not depend on p_+.

The structure of the propagator is thus

$$S(p) = \frac{-ip_-}{2p_+p_- - m^2 - p_-\Sigma(p_-) + i\varepsilon}.$$

By Lorentz invariance, it has to be $\Sigma(p_-) = \text{const}/p_-$. Define $M^2 = m^2 + p_-\Sigma(p_-)$. We get

$$\Sigma = 4g^2 \int \frac{\mathrm{d}k_+\,\mathrm{d}k_-}{(2\pi)^2} \frac{i(p_- - k_-)}{2k_+(p_- - k_-) - M^2 + i\varepsilon} \frac{P}{k_-^2} = \frac{g^2}{2\pi} \int \mathrm{d}k_-\,\mathrm{sign}(p_- - k_-) \frac{P}{k_-^2} = -\frac{g^2}{\pi p_-}.$$

We can read off the physical mass of the quark

$$M^2 = m^2 - \frac{g^2}{\pi}, \quad S(p) = \frac{ip_-}{p^2 - M^2 + i\varepsilon}.$$

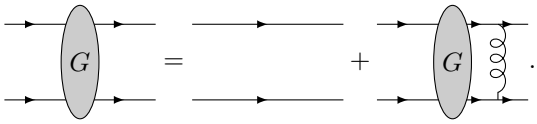
If $g^2 > \pi m^2$, the mass becomes tachyonic. However, quarks are not physical states in the theory, so this is not a problem. However, mesons (quark-antiquark bound states) are physical states, and we should check that their masses are physical. Remember the Green's four-point function for quark-antiquark scattering

$$G(x_1, x_2, y_1, y_2) = \langle\Omega|T\{q(x_1)\overline{q}(x_2)q(y_1)\overline{q}(y_2)\}|\Omega\rangle\,.$$

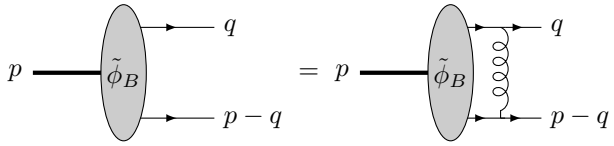
If there is a bound state of mass M_B and if $\phi_B(x_1, x_2) = \langle B(p)|T\{q(x_1)\overline{q}(x_2)\}|\Omega\rangle \neq 0$, then G will behave as

$$G = \frac{i}{p^2 - M_B^2 + i\varepsilon}\tilde{\phi}_B^\dagger(p, q_y)\tilde{\phi}_B(p, q_x) + \text{regular part}.$$

For large N the leading contribution to G comes from planar ladder diagrams:



The algebraic version of this equation is called the *Bethe–Salpeter equation*. Similarly, for $\tilde{\phi}_B$ we have (note that $\tilde{\phi}_B$ includes two external legs)



$$\tilde{\phi}_B(p,q) = \left(\frac{2iq}{\sqrt{N}}\right)^2 NS(q)S(-(p-q)) \int \frac{\mathrm{d}k_+\,\mathrm{d}k_-}{(2\pi)^2} i\frac{P}{(k_- - q_-)^2}\tilde{\phi}_B(p,k).$$

Denote

$$\Psi(p,q_-) = \int \mathrm{d}q_+\,\tilde{\phi}_B(p,q)$$

and then integrate both sides over q_+

$$\begin{aligned}\Psi(p,q_+) &= -\frac{ig^2}{\pi^2}\int \mathrm{d}q_+\, S(q)S(-(p-q)) \int \mathrm{d}k_-\,\frac{P}{(k_- - p_-)^2}\Psi(p,k_-) \\ &= -\frac{ig^2}{\pi^2}\frac{-i\pi}{2p_+ - M^2(1/p_- + 1/(p_- - q_-))}\int \mathrm{d}k_-\,\frac{P}{(k_- - q_-)^2}\Psi(p,k_-).\end{aligned}$$

The momentum q can lie anywhere on the interval $q \in [0,p]$, so we introduce the variable $x = q_-/p_-$, $x \in [0,1]$ (similarly, $k_- = p_- y$). The bound-state equation becomes

$$\left(2p_+ - \frac{M^2}{p_-}\left(\frac{1}{x} - \frac{1}{1-x}\right)\right)\Psi(x) = -\frac{g^2}{\pi}\int_0^1 \mathrm{d}y\,\frac{P}{(x-y)^2}\Psi(y).$$

We multiply both sides by p_- and take $M_B^2 = p^2 = 2p_+p_-$. The final form is called the *'t Hooft equation*:

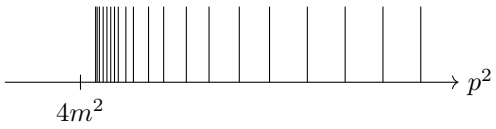
$$M_B^2\Psi(x) = \left(\frac{M^2}{x} + \frac{M^2}{1-x}\right)\Psi(x) - \frac{g^2}{\pi}\int_0^1 \mathrm{d}y\,\frac{P}{(x-y)^2}\Psi(y).$$

The eigenvalues M_B^2 are the squared masses of the meson bound states. The boundary conditions are $\Psi(0) = \Psi(1) = 0$, because the quark and antiquark cannot carry zero momentum fraction. The eigenvalues will be discrete, as only normalizable solutions are allowed. The spectrum can be solved numerically, the easiest way is to discretize the problem by expanding $\Psi(x)$ into a complete set of functions and truncate (Multhopp transform). It becomes a large matrix eigenvalue problem. The results are

- $M_B > 2m$: positive binding energy as in QCD,
- For $M_B \gg M$, we can drop M^2-depednent terms. Use ansatz $\Psi_n(x) = \sin(n\pi x)$ and

$$\int \mathrm{d}y\,P\frac{\cos(\pi n y)}{x-y} = -\pi\sin(\pi n x), \quad -\frac{\partial}{\partial y}\frac{1}{x-y} = \frac{1}{(x-y)^2}.$$

  After the integration by parts, one finds eigenvalues $M_B^2 = g^2\pi n$ for all n. Thus, at high excitations the spectrum is linear in n.



5.6 Operator product expansion and renormalons

5.6.1 Operator product expansion

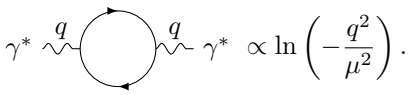
Consider two local operators $\mathcal{O}_1(x)$ and $\mathcal{O}_2(0)$. When $x \to 0$, we would expect singular behaviour in their product. However, we are interested in the case when x is just very close to 0:

$$\mathcal{O}_1(x)\mathcal{O}_2(0) \overset{x\to 0}{=} \sum_n c_n(x)\mathcal{O}_n, \quad c_n(x) \propto \frac{1}{(x^2)^a}\log x^2$$

where a is determined by the dimensions of the operators, $a = ([\mathcal{O}_1] + [\mathcal{O}_2] - [\mathcal{O}_n])/2$. The series is ordered by increasing dimension of the operators $\mathcal{O}_n$, so that the coefficients $c_n(x)$ become less and less singular as $x \to 0$. This is an operator equation, so the equality is meant on the operator level. Consider (only in this section) the following quantity

$$\Pi_{\mu\nu}(q) = (q_\mu q_\nu - q^2 g_{\mu\nu})\Pi(q^2) = i\int \mathrm{d}^4x\, e^{iqx}\,\langle\Omega|j_\mu^\dagger(x)j_\nu(0)|\Omega\rangle\,, \quad j_\mu = \overline{q}\gamma_\mu q.$$

The current j_μ is the electromagnetic quark current. The self-energy function $\Pi(q^2)$ is analytic on the cut complex plane. The cut starts at the physical threshold $q^2 = 4m_\pi^2$. In perturbation, the cut starts at $q^2 = 0$ (for massless quarks) as the main contribution comes from quark loops

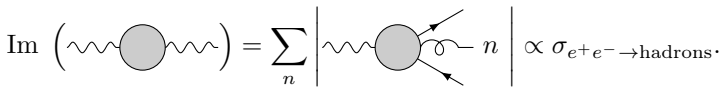


$$\propto \ln\left(-\frac{q^2}{\mu^2}\right).$$

Define the *Adler function* $D(Q^2)$ as the derivative of $\Pi(Q^2)$:

$$D(Q^2) = 4\pi^2 Q^2 \frac{\mathrm{d}\Pi(Q^2)}{\mathrm{d}Q^2}, \quad Q^2 = -q^2 > 0.$$

The optical theorem tells us that we can relate the imaginary part of the self energy to the cross-section of $e^+e^- \to$ hadrons:



$$\propto \sigma_{e^+e^-\to\text{hadrons}}.$$

Using the dispersion relation, we can write

$$\left.\frac{\mathrm{d}\Pi(q^2)}{\mathrm{d}q^2}\right|_{q^2=-Q^2} = \frac{1}{\pi}\int_0^\infty \frac{\mathrm{d}q^2}{(q^2+Q^2)^2}\mathrm{Im}\;\Pi(q^2+i\varepsilon).$$

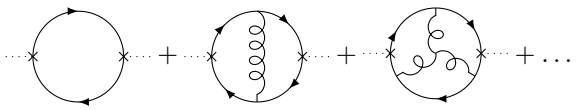
In general, the two-point functions of external electromagnetic or electroweak currents contain important information about QCD and Standard model parameters. Consider the operator product expansion of $\Pi_{\mu\nu}$, where $\mathcal{O}_1 = \mathcal{O}_2 = j_\mu$. What operators $\mathcal{O}_n$ can appear?

- They have to be gauge invariant.
- They have to conserve the same quantum numbers as $\mathcal{O}_1\mathcal{O}_2$ (parity, flavour, CP, etc.).
- We can drop $\mathcal{O}_n$ that vanish by the classical equations of motion when taking on-shell matrix elements (the Ward identity gives contact terms which vanish on-shell).

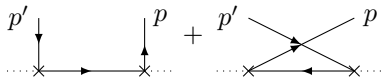
There are no gauge invariant 2-dimensional operators, so the first non-trivial operator appears at dimension four. There are two such operators:

$$i\int \mathrm{d}^4x\, e^{iqx}T\{j_\mu(x)j_\nu(0)\} = (q_\mu q_\nu - q^2 g_{\mu\nu})\left[c^{(1)}(q^2)\mathbb{1} + c^{(q\overline{q})}(q^2)m_q\overline{q}q + c^{(GG)}(q^2)G^{\mu\nu,A}G^A_{\mu\nu}\right]. \tag{$*$}$$

The first coefficient $c^{(1)}(q^2)$ can be calculated in perturbation theory. The identity has vacuum expectation value $\langle 0|\mathbb{1}|0\rangle = 1$, so when we take the vacuum expectation value, only the first term contributes at leading order. On the left-hand side we get exactly the perturbative result for $\Pi_{\mu\nu}(q)$ (the dotted lines represent amputated external photons):

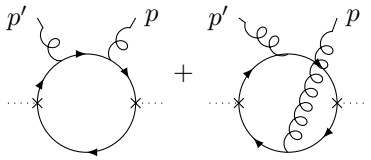


The second coefficient $c^{(q\overline{q})}(q^2)$ comes from $\langle q(p')|\ldots|q(p)\rangle$ and taking $p' \to p$ (both momenta have to be very small $p, p' \ll q$). On the right, we get $c^{(qq)}\overline{u}(p')u(p)$ (the contribution also comes from GG operator, but only at order g^2). When projected on this state, only the diagrams that have two external quark lines contribute to the left side. The leading contribution is



$$+\ldots, \quad c^{(q\overline{q})}(q^2) = \frac{2}{q^4} + \mathcal{O}(\alpha_\mathrm{s}).$$

For $c^{(GG)}(q^2)$ we have to project by $\langle A(p')|\ldots|A(p)\rangle$ and



$$+\ldots, \quad c^{(GG)}(q^2) = \frac{\alpha_\mathrm{s}}{12\pi q^4} + \mathcal{O}(\alpha_\mathrm{s}^2).$$

Now we sandwich (*) between the true, non-perturbative QCD vacuum states $\langle\Omega|\ldots|\Omega\rangle$. All $\langle\Omega|\mathcal{O}_n|\Omega\rangle$ are in general non-zero. The only scale in the problem is Λ_QCD, so by dimensional analysis

$$\langle\Omega|\mathcal{O}_n|\Omega\rangle = \mathcal{O}(\Lambda_\mathrm{QCD}^{[\mathcal{O}_n]}), \quad \langle q\overline{q}\rangle \sim \Lambda_\mathrm{QCD}^2, \quad \langle\alpha_\mathrm{s}G^2/\pi\rangle \sim \Lambda_\mathrm{QCD}^4.$$

These are quark and gluon condensates. Hence, the OPE is an expansion in powers of

$$\alpha_\mathrm{s} \sim \frac{1}{-\beta_0 \ln(q^2/\Lambda^2_\mathrm{QCD})} \quad \text{and} \quad \frac{\Lambda_\mathrm{QCD}}{q} \propto \exp\left(\frac{1}{-2\beta_0\alpha_\mathrm{c}(q)}\right).$$

This gives a transseries structure to the expansion of $D(Q^2)$. One further simplification is $m = 0$, which holds well for u and d quarks (and we are left with only one 4-dimensional operator):

$$(\alpha_\mathrm{s}G^2/\pi)_0 = Z_{G^2}(\alpha_\mathrm{s}G(\pi)), \quad \gamma_{G^2}(\alpha_\mathrm{s}) = \frac{1}{Z_{G^2}}\mu^2\frac{\mathrm{d}}{\mathrm{d}\mu^2}Z_{G^2}.$$

Since the bare operator is independent of the renormalization scale μ, it holds

$$\left(\mu^2\frac{\mathrm{d}}{\mathrm{d}\mu^2} + \gamma_{G^2}\right)\frac{\alpha_\mathrm{s}}{\pi}G^2 = 0.$$

Note that since G depends on mu^2 and $\alpha_s(\mu^2)$, the total derivative with respect to μ^2 can be rewritten as

$$\mu^2 \frac{\mathrm{d}}{\mathrm{d}\mu^2} = \mu^2 \frac{\partial}{\partial \mu^2} + \mu^2 \frac{\partial \alpha_s}{\partial \mu^2} \frac{\partial}{\partial \alpha_s} = \mu^2 \frac{\partial}{\partial \mu^2} + \beta(\alpha_s) \frac{\partial}{\partial \mu^2}.$$

Furthermore, $\mu^2 \mathrm{d}D(Q^2)/\mathrm{d}\mu^2 = 0$ so the renormalization group equation reads

$$\left(\mu^2 \frac{\partial}{\partial \mu^2} + \beta(\alpha_s) \frac{\partial}{\partial \alpha_s} - \gamma_{G^2}(\alpha_s) \right) c^{(GG)}(q^2, \mu^2, \alpha_s(\mu)) = 0$$

with the solution

$$c^{(GG)}(q^2, \mu^2, \alpha_s(\mu)) = \exp\left(- \int_{\alpha(\mu')}^{\alpha_s(\mu)} \mathrm{d}\alpha_s' \, \frac{\gamma_{G^2}(\alpha_s')}{\beta_0(\alpha_s')} \right) c^{(GG)}(q^2, \mu'^2, \alpha_s(\mu')).$$

Since the product of the coefficient and the condensate is renormalization group invariant, we also have

$$\left(\frac{\alpha_s}{\pi} G \right)(\mu) = \widehat{\left(\frac{\alpha_c}{\pi} G \right)}(\mu') \exp\left(- \int_0^{\alpha_s(\mu)} \mathrm{d}\alpha_s' \, \frac{\gamma_{G^2}(\alpha_s')}{\beta_0(\alpha_s')} \right).$$

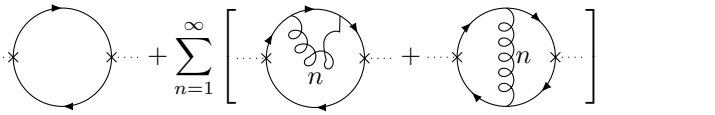
The lower limit of the integral is 0 because at $\alpha_s(\mu) = 0$ for $\mu \to \infty$. There is only one unknown parameter, the renormalization group invariant condensate $\widehat{(\alpha_s G^2/\pi)}$. Choosing $\mu' = \sqrt{Q^2}$, one can completely determine the Q-dependence of $c^{(QQ)}(q^2) G_{\mu\nu} G^{\mu\nu}$ in the OPE of the Adler function and express it as a series expansion in α_s. The expansion of G^2 operator in terms of α_s is

$$c^{(GG)}(q^2, \mu^2, \alpha_s(\mu)) \langle \alpha_s G^2/\pi \rangle (\mu) = \text{const} \times \exp\left(\frac{d}{2\beta_0 \alpha_s(Q)} \right) (-\beta_0 \alpha_s(Q))^{\frac{d\beta_1}{2\beta_0^2} - \frac{\gamma^0}{\beta_0}} (1 + \mathcal{O}(\alpha_s)),$$

where $d = [G^2] = 4$ is the dimension of the operator β, γ factors come from the expansion of beta and gamma functions, $\beta(\alpha_s) = \beta_0 \alpha_s^2 + \beta_1 \alpha_s^3 + \ldots$ and $\gamma(\alpha_s) = \gamma_0 \alpha_s + \ldots$. The constant in the front absorbs $\widehat{(\alpha_s G^2/\pi)}$ and $1/12$ from $c^{(G^2)}$. The exponential term shows the transseries nature of the OPE, and it equals $(\Lambda_{\mathrm{QCD}}/Q)^d$. Everything else is determined by anomalous dimension and short-distance coefficients, which can be calculated in perturbation theory due to asymptotic freedom.

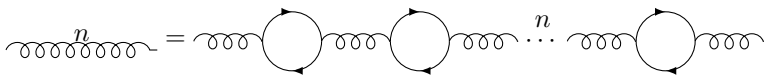
5.6.2 Bubble graphs and renormalons

We take $c^{(1)}(q^2)$ and the Adler function with $n_f \to -\infty$ (so that $\beta_0 = n_f/(12\pi)$). We use this weird limit in order to keep the theory asymptotically free. In this toy model, we expand the self energy function in terms of planar bubble diagrams

$D(Q^2) =$  $+ \sum_{n=1}^{\infty}$

$$= 1 + \sum_{n=0}^{\infty} \alpha_s(\mu) \int_0^{\infty} \frac{\mathrm{d}k^2}{k^2} F(k^2) \left(\beta_0 \alpha_s(\mu) \ln\left(-\frac{k^2}{\mu^2} e^{-5/3} \right) \right)^n$$

where $\mu = Q$. The calculation is very similar to the previous homework problem. The gluon propagator index n means that there are n fermion bubbles inserted into the gluon propagator:



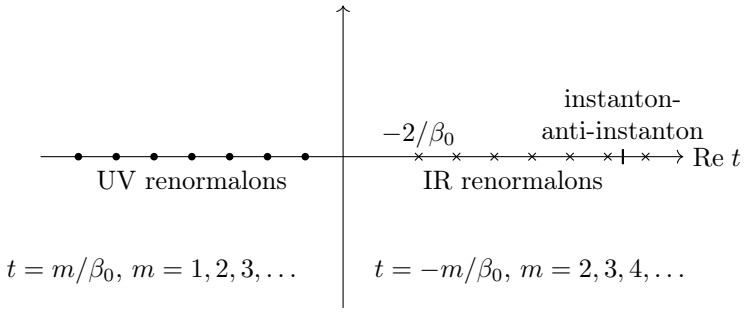
To get UV and IR renormalons, we only need to know the behaviour of $F(k^2)$ at small and large momenta (use $x = -k^2/Q^2$):

$$F(x) = \begin{cases} \frac{C_F}{3\pi}\frac{\ln(x)+5/6}{x} + \mathcal{O}(x^{-2}), & x \gg 1, \\ \frac{3C_F}{2\pi}x^2 + \mathcal{O}(x^3), & x \ll 1. \end{cases}$$

Performing the integral over k^2, we get

$$D(Q^2) = 1 + \frac{C_F}{\pi}\sum_{n=0}^{\infty}\alpha_s^{n+1}\left(\frac{1}{3}e^{-5/3}\beta_0^n n!(n+11/6) + \frac{3}{4}e^{10/3}(-\beta_0/2)^n n!\right).$$

The first term comes from the UV region and gives poles on the Borel plane (UV renormalon), the second term comes from the IR region and gives IR renormalon poles. The Borel plane for the exact Borel transform in the $n_f \to -\infty$ limit is shown below.



The leading UV renormalon at $t = -1/\beta_0$ gives formally the dominant behaviour as $n \to \infty$. In the $\overline{\text{MS}}$-scheme, the overall constant is a factor e^{-5} smaller than the leading IR renormalon.

5.6.3 OPE as a transseries

The IR renormalons come with constant sign in the expansion, so they imply non-Borel summability of the perturbative series. This is consistent with the fact that the perturbative vacuum in the Fock space of quark-gluon states is not the true QCD vacuum $|0\rangle \to |\Omega\rangle$. There is a lot going on in the true vacuum (not only instantons but also confinement, chiral symmetry breaking, condensates, etc.). When we interpret the Borel integral as a lateral Borel sum, the ambiguity that arises must be cancelled by non-perturbative contributions in the OPE. Schematically,

$$D(Q^2) = \sum_k c_k(\alpha_s(Q))\frac{1}{Q^{[\mathcal{O}_k]}}\langle\Omega|\mathcal{O}_k(0)|\Omega\rangle = \sum_k c_k(\alpha_s(Q))\left(\frac{\Lambda_{QCD}}{Q}\right)^{[\mathcal{O}_k]}.$$

This allows us to draw some general conclusions about the structure of $c_k(\alpha_s)$.

(1) The location of the IR renormalons can only be at $t = -m/\beta_0$, where $2m = 1, 2, 3, \ldots$, because the ambiguities must be cancelled by operators of integer dimension.

(2) For the Adler function, the series is in $1/Q^2$ not $1/Q$, hence only even m are allowed. The renormalon at $m = 1$ does not appear, because there is no gauge/Lorentz invariant operator of dimension 2.

(3) The overall constant has to be ambiguous. This follows from the fact local operators $[\mathcal{O}] = n$ have stronger UV divergence than logarithmic (power divergences with a UV cut-off).

(4) Since $\alpha_s(Q)$-dependence of the power corrections is known up to a constant, the large-order behaviour due to an IR renormalon will be fully determined (in terms of coefficients of beta and gamma functions) up to a constant. For the Adler function the leading IR renormalon at $-2/\beta_0$ is related to a single operator, $\alpha_s G^2/\pi$. The $\alpha_s(Q)$ dependence gives

$$D(Q^2) = 1 + \sum_n d_n \alpha_s^{n+1},$$

where the leading contributions to d_n come from IR renormalons are

$$d_n = K \left(-\frac{\beta_0}{2}\right)^n \Gamma(n + 1 - 2\beta_1/\beta_0)(1 + \mathcal{O}(n^{-1})).$$

The corrections in $1/n$ come from higher orders in beta and gamma functions (up to $1/n^4$ known).

One can also relate UV renormalons to local operators that appear if one considers QCD as an effective field theory below a cut-off Λ. UV renormalons at $t = m/\beta_0$ are then classified in terms of operators of mass dimension $4 + 2m$ in the non-renormalizable extension of QCD including all operators of any dimension compatible with Lorentz invariance and SU(3) gauge invariance.

Index

Symbols
θ-vacuum 88

A
Adler function 136
anomaly equation 106
asymptotic series 110
Atiyah–Singer index theorem 103

B
Bethe–Salpeter equation 134
bino 16
Bogomolny bound 84, 87, 98
Boltzmann equation 71, 72
Borel integral 111
Borel sum 111
Borel transformation 111
BPS monopole 87

C
central charges 6
Chern–Simons current 89
chiral multiplet 8
chiral symmetry 101
Christoffel symbols 25
Coleman–Mandula theorem 4
collective coordinates 95
collision term 71
constraint equation 72

D
de Donder gauge 27
Debye mass 57
Derrick's theorem 80
digamma function 114
dilute instanton gas approximation 77, 93, 119
Dirac quantization condition 87
domain wall 79

E
Einstein field equations 25
Einstein–Hilbert action 25
Euclidean action 40
Euler characteristic 130
extended supersymmetry 7

F
Faddeev–Popov method 27
Fayet–Iliopoulos term 12
Fierz identity 11
Fierz–Pauli Lagrangian 22
first homotopy group 82
fundamental group 82

G
gain term 71
gaugino 12
Gel'fand–Yaglom theorem 92
Georgi–Glashow model 85
Gevrey class 110
gluino 16
goldstinos 15
Goldstone boson equivalence theorem 35
gradient expansion 72
grand canonical ensemble 38
grand potential 38
Grassmann numbers 5
gravitino 15
graviton 27
graviton multiplet 8
Gross–Neveu model 123

H
hedgehog configuration 85

Higgsino 16
homotopy 82
homotopy theory 82

I
instanton 91
IR renormalon 125

K
Kadanoff–Baym equation 71
kink 74
Kubo–Martin–Schwinger relation 44

L
Landau damping 70
lateral Borel integrals 113
light-cone gauge 132
Linde problem 58
loss term 71

M
magnetic monopole 87
Matsubara frequencies 40
Mellin transform 121
metastable state 97
minimal coupling 29
minimally supersymmetric Standard Model 16
moduli space 98

N
n-th homotopy group 83
Nevalinna–Sokal theorem 111
Nielson–Olesen vortex 83

P
partition function 38
Passarino–Veltman reduction 56
Pauli–Lubanski vector 4, 8
Pauli–Villars regularization 99
Pontryagin index 101
Pöschl–Teller potential 117

Q
quasi-particle approximation 72

R
rainbow diagrams 133
resurgence theory 114
Ricci scalar 25
Ricci tensor 25
Riemann curvature tensor 25
ring diagram 50

S
Schwinger–Dyson equation 71
skyrmion 80
slepton 16
spectral function 63
sphaleron 107
squark 16
statistical function 63
Stokes lines 114
Striling formula 112
strong CP problem 102
Stuckelberg trick 34
superfield formalism 13
superhiggs mechanism 14
superpotential 11
supersymmetry 5

T
't Hooft coupling 124
't Hooft equation 135
't Hooft model 132
't Hooft vertex 105
't Hooft–Polyakov monopole 87
't Hooft symbols 98
thermal mass 49
transport equation 72
transseries 114

U
UV renormalon 125

V
Vainshtein mechanism 36
Vainshtein radius 37
vDVZ discontinuity 34
vector multiplet 8
vortex 83

W
Wess–Zumino model 9
Wigner space 72
wino 16