On quantum description of saturated systems

Gross-Neveu, axions, de-Sitter

Otari Sakhelashvili

München 2022

On quantum description of saturated systems

Gross-Neveu, axions, de-Sitter

Otari Sakhelashvili

Dissertation
der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Otari Sakhelashvili
aus Tbilisi, Georgia

München, den 15.07.2022

Supervisor: Prof. Gia Dvali
First Referee: Prof. Gia Dvali
Second Referee: Prof. Goran Senjanovic
Date of the exam: 31.08.2022

Zusammenfassung

In dieser Arbeit untersuchen wir bestimmte physikalisch motivierte Systeme, in denen "Large-N-Physik" und Eichredundanz eine wichtige Rolle spielen.

Kürzlich wurde eine neue Obergrenze für die Mikrozustandsentropie vorgeschlagen und es wurde gezeigt, dass die Objekte, die diese Grenze sättigen, sogenannte Saturonen, alle Schlüsseleigenschaften eines Schwarzen Lochs teilen. Dieser Befund weist darauf hin, dass die Eigenschaften von Schwarzen Löchern nicht spezifisch für die Gravitation sind und für alle Saturonen universell sind. Wir überprüfen diese Idee explizit, indem wir zeigen, dass die maximal entropiegebundenen Zustände im Gross-Neveu-Modell Saturonen sind und die Eigenschaften eines Schwarzen Lochs tragen. Die Kraft der Large-N-Physik und die asymptotische Freiheit der Theorie erlauben es uns, ein sehr klares Verständnis der Verbindung zu vermitteln. Wir decken somit die Grundlage der zugrunde liegenden Quantenphysik des Schwarzen Lochs in einem vollständig berechenbaren Modell auf.

Zweitens diskutieren wir eine eichredundante Formulierung von Axion und demonstrieren seine Unempfindlichkeit gegenüber UV-Physik.

Schließlich konstruieren wir explizit einige wichtige invariante BRST-Zustände in der Elektrodynamik und in großer N Gravitation, mit besonderer Implikation der Formulierung des de Sitter-Raums als eines kohärenten Zustands, der auf dem Minkowski-Vakuum aufgebaut ist. Diese Konstruktion ist wesentlich für das Verständnis der Lebensfähigkeit und Eigenschaften von de-Sitter-ähnlichen Zuständen innerhalb der S-Matrix formulierung der Quantengravitation, der derzeit einzigen bekannten Formulierung. Diese Formulierung impliziert, dass de-Sitter nicht als Vakuum angesehen werden kann. Statt dessen muss es als kohärenter Zustand dargestellt werden, der sich nicht trivial mit der Zeit entwickelt. Dies hat wichtige Auswirkungen auf die Physik der Dunklen Energie.

Abstract

In this thesis, we study certain physically motivated systems in which large-N physics and gauge redundancy play an important role.

Recently, a new upper bound on microstate entropy was proposed and was demonstrated that the objects saturating this bound, so-called "saturons", share all the key properties of a black hole. This finding indicates that black hole properties are not specific to gravity and are universal to all saturons. We give the explicit check to this idea by showing that the maximal entropy bound states in Gross-Neveu model are saturons and carry the black hole like properties. The power of large-N physics and the asymptotic freedom of the theory allows us to give a very clean understanding of the connection. We thus uncover the foundation of black hole's underlying quantum physics within a fully calculable model.

Secondly, we discuss a gauge redundant formulation of axion and demonstrate its insensitivity towards UV-physics.

Finally, we explicitly construct some important BRST invariant states in electrodynamics and in large-N gravity, with particular implication of formulating de Sitter space as a coherent state constructed on top of the Minkowski vacuum. This construction is essential for understanding the viability and properties of de-Sitter like states within the S-matrix formulation of quantum gravity, currently the only known formulation. This formulation implies that de Sitter cannot be regarded as a vacuum. Instead, it must be represented as a coherent state that evolves in time non-trivially. This has important implications for physics of dark energy.

Publications

The thesis is based on the following publications:

- O. Sakhelashvili, "Consistency of the dual formulation of axion solutions to the strong CP problem", Phys. Rev. D **105**, 085020 (2022), arXiv:2110.03386 [hep-th],
- G. Dvali and O. Sakhelashvili, "Black-hole-like saturons in Gross-Neveu", Phys. Rev. D **105**, 065014 (2022), arXiv:2111.03620 [hep-th],
- L. Berezhiani, G. Dvali, and O. Sakhelashvili, "de Sitter space as a BRST invariant coherent state of gravitons", Phys. Rev. D **105**, 025022 (2022), arXiv:2111.12022 [hep-th],

We also mention ongoing works:

- G. Dvali, O. Sakhelashvili, and A. Paolini, "In preparation",
- G. Dvali and O. Sakhelashvili, "In preparation",
- L. Berezhiani, G. Dvali, and O. Sakhelashvili, "In preparation",

Contents

Introduction

This thesis is devoted to understanding certain profound quantum field theoretic phenomena via using the powers of gauge redundancy and large-N physics.

Usually, the classical systems are well described by quantum coherent states of high occupation number of quanta. However, it has been observed that among such systems there exist states with a rather peculiar quantum behavior. In particular, such states store the maximal amounts of quantum information, with their microstate entropy being proportional to the area. At initial stage, they decay in thermal like way, but start the release of information only by the time of their half decay. Moreover, by then, the classical approximation breaks down fully.

Namely, it has been suggested in [7] that a black hole represents a bound state of N soft gravitons at the point of quantum criticality. This explains its key properties, such as, e.g. the maximal entropy with area-law, Hawking evaporation, long time-scale of information retrieval.

More recently[8–10], it has been argued that none of these properties are specific to either black holes or gravity. Rather, they are universally shared by the objects, called "saturons", which saturate the new upper bound on entropy imposed by unitarity. In particular, this maximal entropy is equal to the area of the object measured in units of the Goldstone coupling. The Goldstone in question, results from the symmetries that are spontaneously broken by the object. In particular, the Goldstone boson that is universally present is the one from spontaneous breaking of Poincare symmetry. It has been proposed and demonstrated on several examples that saturons have all the above-mentioned properties of a black hole [8–10].

The first original result of the thesis, based on publication [2], is the reliable demonstration of this connection in a fully calculable large-N model of Gross and Neveu [11]. This model is known to be asymptotically free and exactly solvable in large-N. It therefore represents an ideal laboratory for demonstrating the properties of saturation and its connection to black hole physics. In this model, we explicitly identify saturon states in the form of maximal entropy bound states and show that they indeed possess all the properties of a black hole. In particular, they obey the same area-law entropy as the black holes, they evaporate in a thermal like way very similar to Hawking, and the time-scale of information retrieval is given by the same formula as suggested for black holes by Page.

We thus discover that in full accordance with the proposed black hole/saturon correspondence, the saturated bound states of Gross-Neveu posses all the properties of a black hole. Needless to say, it is of extraordinary importance that we are able to reproduce the black hole features in a calculable asymptotically free theory. This proves that black hole features originate from the phenomenon of saturation rather than from the particularities of quantum gravity.

The next system of our interest is an axion. It is known that the gauge theories such as QCD possess toplogically non-trivial vacuum structure. Due to this, the theory splits into infinite set of superselection sectors that differ by the strength of CP-violation. Within the ordinary QCD, the amount of measured CP-breaking is un-observably small. This smallness goes under the name of "strong-CP problem". The axion is the field that supposedly solves the puzzle since it relaxes the vacua to a unique ground state. The original realization was offered by Peccei and Quinn [12], with the role of the dynamical axion pointed out in [13, 14]. However, the effect is highly sensitive to operators that break the axion shift symmetry explicitly. Such operators are expected to be generated by various non-perturbative effects in gravity. This issue was addressed in [15], where it was shown that the QCD axion admits an alternative (dual) formulation that does not require any global shift symmetry. Instead the theory is protected by the gauge redundancy of QCD. The axion, which is represented by an antisymmetric 2-form, shifts under the QCD gauge transformation. It has been suggested that due to gauge symmetry this formulation is insensitive to any UV-corrections. In the thesis, we explicitly check this claim (following our article [1]) in various regimes by coupling the theory to all possible heavy sources and by resolving these sources by an explicit UV-physics. Our results have implications to other proposed solutions to the naturalness puzzles in which the forms are present. Such is a so-called attractor mechanism [16] that was introduces for solving the Hierarchy Problem.

Finally, we move to construction of explicitly BRST invariant coherent states in gauge theories such as electrodynamics and gravity. We give certain important examples of coherent states that describe space-time metrics produced by classical sources such as the cosmological constant. This source is known to produce a so-called de-Sitter metric in gravity. However, it has been argued that such a space-time cannot play the role of the vacuum for quantum gravity [17]. A particularly strong argument [18] for this claim is that de-Sitter cannot support the S-matrix, which currently is the only existing formulation of quantum gravity. For this reason, it was proposed [17] that de Sitter must be viewed as a coherent state constructed on top of the Minkowski vacuum. In this thesis, based on our published work [3], we offer the first BRTS invariant construction of such a state. We also discuss a special double scaling large-N limit of gravity, called the "species limit" [19]. In this limit, the number of particle species N coupled to gravity is taken infinite, whereas the scale of quantum gravity (the so-called species scale [20]) is kept finite. As a result, the infinite set of corrections in which the quantum gravitational coupling is not accompanied by N vanishes and the analysis simplifies significantly. In this limit, our BRST invariant construction of de-Sitter as of coherent state of gravitons becomes exact. Within this limit, we study various effects, including the dynamics of Gibbons-Hawking radiation. Our analysis has important implications both for fundamental understanding of de Sitter and for the observational consequences of dark energy. In the view of [21], dark energy must necessarily come from a time-dependent source and not from the cosmological constant.

The thesis is organized as follows:

- The first chapter is a review of basic building blocks of the gauge theories, including the celebrated BRST quantization, anomaly cancellation and antisymmetric gauge forms.
- The second Chapter establishes the large-N limit and semi-classical approximation. Instead of general considerations, the large-N limit is demonstrated on a specific example. Then the chapter reviews classical topological solutions and WKB approximation. After this, the chapter establishes newly found objects, "saturons". These are the objects that saturate the recently proposed unitarity upper bound on entropy. It has been demonstrated that all saturons in many respects behave like black holes. Next, we discuss the original results, namely the saturons in the Gross-Neveu model. Our statements are exact, since the model is exactly solvable. The discussion is based on our work [2], we also mention the work in preparation [4, 5]. Finally, the chapter introduces the large-N gravity.
- The third Chapter is dedicated to the axion physics. We review the properties of the θ-vacua and some of its consequences. We establish a deep connection between structure of the θ-vacua and gauge redundancy and the gauge-redundant formulation of axion. We also discuss connection with large-N-physics. We then discuss our original result published in [1], which studies and confirms the insensitivity of gauge-redundant formulation of axion and of 3-forms towards UV-corrections. We discuss application of these results to the so-called attractor solution to the Hierarchy Problem, proposed earlier in [16].
- The fourth Chapter follows our article [3]. We build BRST-invariant coherent states in the QED and in the linear gravity. We then discuss the emergence of classical physics in this framework and a consistent treatment of the quantum corrections. We also explain in details the inclusion of classical sources in the above theories. As an important example of such a source, we discuss the cosmological constant. We formulate the metric produced by it, the de Sitter metric, as a coherent state of gravitons. We discuss the large-N limit of gravity and the saturation of the entropy bound by the corresponding coherent state. We further clarify some details of the framework [6], the clarifications are acknowledged during the discussion.

Chapter 1

Gauge Symmetries

In physics, symmetries play an important role. We usually build a physical model invariant under space-time translations. This assumption immediately conserves energy and momentum due to Nöther's theorem (see, e.g., [22, 23]). Physical systems can have so-called global internal symmetries in addition to space-time ones. They are conserved because of the validity of the Nöther's theorem for such symmetries. It appears that we can make global symmetries local by gauging them. But there is one caveat, the gauge symmetries are redundancies in reality. For instance, let us take a massless vector field. The analysis of Lorenz group leads us to only two helicity modes (see, e.g., [23]) of the vector field. On the other hand, we need a four-vector to keep an explicit Lorenz invariance. This means that we have some redundancy in the description. Indeed, local gauge transformations keep the system exactly redundant, resulting with the compatibility with the Lorenz group. On the one hand, this approach makes the system explicitly Lorenz invariant, but it is hard to deal with redundant systems, especially during the quantization procedure. There are different methods to quantize gauge redundant systems, but the most elegant and consistent way is a so-called BRST quantization (see e.g., [24]). In BRST framework, gauge transformations are parts of the bigger global symmetry. The BRST symmetry itself is not redundancy per se, for the full system it is only a global symmetry. So, it is possible to use usual Quantum field theoretic methods, including the Nöether's theorem.

In this chapter, we will talk about gauge redundant systems. We will review BRST symmetry, empathizing methodology of building physical states. We will also cover antisymmetric forms and their gauge structures.

1.1 Nöther's Theorem

In this section, we will review Nöther's theorem (see, e.g., [22]) for internal symmetries. Let us consider some fields Φ_n, where n is a label of a representation of some symmetry. Now we take into account an infinitesimal symmetry transformation, it takes the form

$$\delta\Phi_n^A = \epsilon_{nm}\Phi_m = T_{nm}^A\Phi_m\epsilon^A, \tag{1.1}$$

where ϵ_{nm} is a parametrization of the transformation. We should consider that a symmetry forms groups, meaning that this parametrization can be written as generators T and ϵ^A as an infinitesimal parametrization of the corresponding symmetry group, where A labels the generators. Let us consider an action $S[\Phi]$ which is invariant under such transformations. For infinitesimal ones we get

$$\begin{aligned} 0 = \delta S[\Phi] = \frac{\delta S}{\delta\Phi_n}\delta\Phi_n + \frac{\delta S}{\delta\partial_\mu\Phi_n}\delta\partial_\mu\Phi_n = \\ = \left(\frac{\delta S}{\delta\Phi_n} - \partial_\mu\frac{\delta S}{\delta\partial_\mu\Phi_n}\right)\delta\Phi_n + \int d^4x\partial_\mu j_\mu^A\epsilon^A, \end{aligned} \tag{1.2}$$

where j is the Nöether's current, and is conserved on-shell, i.e., the equations of motions are satisfied. It has the following form

$$j_\mu^A = \frac{\delta S}{\delta\partial_\mu\Phi_n}T_{nm}^A\Phi_m. \tag{1.3}$$

Now we consider this theory in the quantum regime; the transition amplitude in the path integral formalism is given by

$$W = \int D[\Phi]e^{iS[\Phi]}. \tag{1.4}$$

Let us assume that this symmetry transformation leaves the quantum measure invariant, i.e.,

$$D[\Phi'] = D[\Phi]. \tag{1.5}$$

Now we take a variation of the field with proper boundary conditions, such that the field does not vary at initial and final time moments and vanishes at infinity (Hamilton's Action principle). Hence, we get

$$-i\delta W = \int D[\Phi]e^{iS[\Phi]}\int d^4x\left([\frac{\delta S}{\delta\Phi_n} - \partial_\mu\frac{\delta S}{\delta\partial_\mu\Phi_n}]\delta\Phi_n\right). \tag{1.6}$$

Requiring an amplitude and an action (1.2) being invariant under the symmetry while considering the previous expression for the field on-shell, we find

$$0 = i\delta W = \int D[\Phi]e^{iS[\Phi]}\int d^4x\partial_\mu j_\mu^A\epsilon^A. \tag{1.7}$$

After taking derivatives with respect to ϵ's, we can derive the Ward identities in terms of Feynman diagrams.

We should mention that during the discussion, the quantum measure was assumed to be invariant under the symmetry transformations. This consideration is true only for the so-called "non-anomalous" symmetries. If a measure is not invariant, then a symmetry is called anomalous and this discussion is not valid anymore. We will cover such symmetries later.

1.2 Faddeev-Popov method and BRST symmetry

In this section, we will review Faddeev-Popov method and the BRST symmetry. Let us consider Yang-Mills Lagrangian with one Dirac fermion ψ,

$$\mathcal{L} = -\frac{1}{4}G^2 + \bar{\psi}(i\not{D} - m)\psi, \tag{1.8}$$

where G is a field strength tensor of a non-abelian field A, D corresponds to a covariant derivative, and we take the fermion ψ as a fundamental representation of the non-abelian group.

For definiteness, let us consider $SU(N)$ gauge symmetry. (Anti-)fundamental indices are contracted via matrix notation, and adjoint ones are denoted by capital Latin letters. In this notation infinitesimal gauge transformations have the following form

$$\begin{aligned} A_\mu^A \to A_\mu^A + f^{ABC}A_\mu^B\epsilon^C + \frac{1}{g}\partial_\mu\epsilon^A &= A_\mu^A + \frac{1}{g}\left(D_\mu\epsilon\right)^A \\ \psi &\to \psi + i\epsilon^A T^A\psi, \end{aligned} \tag{1.9}$$

where ϵs are N^2-1 different gauge transformations (dependent on space-time), and f is a totally antisymmetric structure constant of the group.

The gauge redundancy maintains correspondence between Lorenz group representations and fields describing physical phenomena. This description complicates quantization of systems, for instance we can not apply Legendre transformation to A_0s, since

$$\Pi_0^A = \frac{\partial\mathcal{L}}{\partial\partial_0 A_0^A} = -G_{00}^A = 0. \tag{1.10}$$

A problem arises when we apply a naive path integral quantization to the system. We will over-count field configurations due to gauge redundancies. Usual method for dealing with this problem is a gauge fixing. Following Faddeev-Popov procedure [25] (for a modern review, see e.g., [22, 24]) we should integrate in such ghost fields which maintain gauge fixing condition.

The Faddeev-Popov method factorizes the path integral, and we are left with the physical path integral multiplied by the volume factor of the gauge group. Let us consider for simplicity a pure glue. Then the naive path integral has the following form,

$$\int \mathcal{D}A e^{iS(A)}, \tag{1.11}$$

where $S(A)$ is a pure glue action. This path integral over-counts physical configurations, due to gauge redundancy. Now let us consider the following identity,

$$1 = \int \mathcal{D}\epsilon \delta(\epsilon) = \int \mathcal{D}\epsilon \delta(G(A)) \det\left(\frac{\delta G(A)}{\delta \epsilon}\right), \tag{1.12}$$

where ϵs are local gauge transformation parameters and $G(A)$ represents a gauge fixing condition. Now, we can rewrite (1.11) as

$$\int \mathcal{D}A e^{iS(A)} = \int \mathcal{D}\epsilon \delta(G(A)) \det\left(\frac{\delta G(A)}{\delta \epsilon}\right) \int \mathcal{D}A e^{iS(A)}. \tag{1.13}$$

Switching integrals gives us the following result

$$\int \mathcal{D}\epsilon \int \mathcal{D}A \delta(G(A)) \det\left(\frac{\delta G(A)}{\delta \epsilon}\right) e^{iS(A)}. \tag{1.14}$$

If the determinant does not depend on ϵ, then integration over gauge transformations can be factored out. For instance, let us choose $G(A)^A = \partial_\mu A^A_\mu - \omega^A$, where ω is an arbitrary function independent of A and ϵ. Then,

$$\delta G(A)^A = \frac{1}{g} \partial_\mu (D_\mu \epsilon)^A \tag{1.15}$$

and the final result will be

$$\int \mathcal{D}\epsilon \int \mathcal{D}A \delta(G(A)) \det\left(\frac{1}{g} \partial_\mu D_\mu\right) e^{iS(A)}. \tag{1.16}$$

The two integrals can be factored out, $\int \mathcal{D}\epsilon$ is just a volume of the gauge group, and we can work with the rest to compute matrix elements. This expression is still inconvenient, since it includes a delta function in a functional space. Let us now take a look at the integral (1.14), which does not depend on ω, irrespective of the form of $G(A)$. So, in principle, we can write

$$N^2 \int \mathcal{D}\omega F(\omega) \int \mathcal{D}\epsilon \int \mathcal{D}A \delta(G(A)) \det\left(\frac{\delta G(A)}{\delta \epsilon}\right) e^{iS(A)}, \tag{1.17}$$

where $F(\omega)$ is an arbitrary functional of ω, converges w.r.t the path integral and N^2 is an appropriate normalization factor. For definiteness, we choose

$$F(\omega) = e^{-i \int d^4x \frac{1}{2\xi} \omega^2}. \tag{1.18}$$

Switching the integrals and factoring out integration over the gauge sector leave us with the following expression,

$$N^2 \int \mathcal{D}\epsilon \int \mathcal{D}A \int \mathcal{D}\omega F(\omega) \delta(G(A)) \det\left(\frac{1}{g} \partial_\mu D_\mu\right) e^{iS(A)}. \tag{1.19}$$

We can get rid of the delta function via integration. So, the final the result is,

$$N^2 \int \mathcal{D}\epsilon \int \mathcal{D}A\, F(\partial_\mu A_\mu) \det\left(\frac{1}{g}\partial_\mu D_\mu\right) e^{iS(A)}. \tag{1.20}$$

Let us write down the determinant using fermionic variables,

$$\det\left(\frac{1}{g}\partial_\mu D_\mu\right) = \int \mathcal{D}c\mathcal{D}\bar{c}\, e^{iS_{gh}}, \tag{1.21}$$

where

$$S_{gh} = -\int d^4x \bar{c}^A \partial_\mu (D_\mu c)^A. \tag{1.22}$$

Here we absorbed g and i in the field normalization, cs and $\bar{c}$s are called Faddeev-Popov ghosts, they are scalars with fermion statistics.

All the above can be written effectively as a Lagrangian density, and we can consistently quantize the system in the path-integral formulation. Then the Lagrangian is represented as

$$\mathcal{L} = -\frac{1}{4}G^2 + \bar{\psi}(i\not{D} - m)\psi + \partial\bar{c}Dc - \frac{1}{2\xi}(\partial A)^2. \tag{1.23}$$

This Lagrangian keeps gauge redundancy at the level of amplitudes. Hence, all Ward identities are intact (see, e.g., [22]). Now, if we try again to compute canonical conjugate momenta of A_0 (1.10), we get

$$\Pi_0^A = -\frac{1}{2\xi}\partial A^A. \tag{1.24}$$

Here, usage of the Legendre transform is not straightforward. So, to address this issue and to answer the question regarding the cause of Ward identities being intact, we can integrate in new auxiliary fields B^A [26, 27] (see e.g., [24]). So, the Lagrangian will boil down to

$$\mathcal{L} = -\frac{1}{4}G^2 + \bar{\psi}(i\not{D} - m)\psi + \partial\bar{c}Dc + \frac{\xi}{2}B^2 - \partial BA \tag{1.25}$$

This expression, except the ghost part, looks almost gauge invariant, if we interpret ∂B as a current. This is precisely what happens in the case of QED [28]. Since B has the power of two in Lagrangian, this and previous formulation are equivalent at any loop level. We can also see that,

$$\Pi_0^A = B^A, \tag{1.26}$$

Therefore, we can build a Hamiltonian, corresponding to the Lagrangian. It also appears, that this system has a so-called BRST symmetry [29, 30] (for a review see, e.g., [22, 24]). The fields transform in the following way,

$$\begin{aligned} \delta_\theta A_\mu^A &= \theta\,(D_\mu c)^A \\ \delta_\theta \psi &= ig\theta c^A T^A \psi \\ \delta_\theta B^A &= 0 \\ \delta_\theta \bar{c}^A &= \theta B^A \\ \delta_\theta c^A &= -\theta g \frac{1}{2} f^{ABC} c^B c^C, \end{aligned} \tag{1.27}$$

where θ is a global fermionic variable. From this transformation, we see that "matter fields" (including the gauge fields) transform under the BRST, in the same way as under the gauge transformation, if we take θc^As as gauge parameters. So, basically, BRST symmetry is a generalization of a gauge symmetry, but a global one. Using the above Lagrangian, we have the possibility to build a Hamiltonian of this system. The BRST symmetry has a corresponding conserved charge, as a consequence of Nöther's theorem.

The symmetry also has the following important property:

- It is nilpotent, i.e., two BRST transformations δs give zero, $\delta_1\delta_2\phi = 0$, where ϕ is a field consists of any combination of the elementary fields.

Since the current of the symmetry is conserved, all quantum corrections should be BRST invariant, but there is an even more restriction on the corrections. The Lagrangian can be written in the following form,

$$\mathcal{L} = -\frac{1}{4}G^2 + \bar{\psi}(i\not{D} - m)\psi + \mathcal{L}_{BRST}, \tag{1.28}$$

where

$$\theta\mathcal{L}_{BRST} = \delta_\theta\left(\frac{\xi}{2}\bar{c}B - \partial\bar{c}A\right), \tag{1.29}$$

Nilpotency guarantees that the full Lagrangian is BRST invariant. Due to this fact, certain radiative corrections tend to be zeros. For instance, only transverse part of a 2-point function of a gauge field will be corrected. Hence, we can only get gauge invariant corrections.

We already covered the importance of the BRST symmetry to get the effective action. Now let us take a look at the building of states in the framework [28] (for a short review, see, e.g., [22]). Therefore, it is better to use a charge corresponding to BRST symmetry than symmetry transformations. In the Hamiltonian formulation of the theory, symmetries are generated by an action of the charge,

$$\delta_\theta\phi = i[\theta Q, \phi], \tag{1.30}$$

where Q is a BRST charge. It is clear, that we should commute charge with the bosonic fields and anti-commute with the fermionic ones. Nilpotency of the above transformations manifests as a peculiar property of the BRST charge,

$$Q^2 = 0, \tag{1.31}$$

telling us that the space of Hamiltonian's eigenstates are split into 3 different sub-spaces. Since the spanned space of eigenstates coincides with the full Hilbert space, we will not distinguish them during the discussion. So, we have the following types of vectors

- vectors for which $Q\left|\psi_1\right\rangle \neq 0$.
- vectors $\left|\psi_2\right\rangle = Q\left|\psi_1\right\rangle$, for which $Q^2\left|\psi_2\right\rangle = 0$, by definition.

- vectors $|\psi_0\rangle$, for which $Q|\psi_0\rangle = 0$, but $|\psi_0\rangle \neq Q|\psi_1\rangle$.

$|\psi_2\rangle$s have peculiar properties, they end up with zero scalar products with themselves and with $|\psi_0\rangle$. One can show that, $|\psi_0\rangle$ have positive norms, and hence physical states should live there. Moreover, as BRST charge Q is conserved, if we start from the state for which

$$Q|\psi\rangle = 0, \tag{1.32}$$

during evolution we end-up with the state satisfying the same condition.

Let us assume having the S matrix process from a state $|\psi\rangle$ to a state $|\phi\rangle$, where for both states $Q|\rangle = 0$. Then we have the following general expressions,

$$\begin{aligned} QS|\psi\rangle &= 0 \\ \langle\phi| S^+ Q &= 0. \end{aligned} \tag{1.33}$$

Since the BRST charge is conserved, $[Q, S] = 0$. Therefore, we can apply the following manipulations

$$\langle\phi|\psi\rangle = \langle\phi| S^+ S|\psi\rangle = \sum_\xi \langle\phi| S^+|\xi\rangle\langle\xi| S|\psi\rangle = \langle\phi|\left[S^+\right][S]|\psi\rangle, \tag{1.34}$$

where $Q|\xi\rangle = 0$, and [S] means the restriction to the $|\psi_0\rangle$ space. Here we utilized unitarity of the full S matrix, then we used conservation of the BRST charge, resulting with the identity $|\xi\rangle\langle\xi| = [1]$. Since ψ_2s have zero scalar products, they do not contribute to the division of the identity. So, we get

$$[S^+][S] = 1. \tag{1.35}$$

This means, that S matrix restricted to the physical space is unitary. We can also show that, it is gauge invariant. Let us take the Lagrangian (1.28) and change gauge fixing. So, we get a new $\tilde{\mathcal{L}}$ Lagrangian. Since we only change $\mathcal{L}_{BRST}$ part of the Lagrangian, the difference between these two will be

$$\theta(\mathcal{L} - \tilde{\mathcal{L}}) = \delta_\theta \Phi = i[\theta Q, \Phi], \tag{1.36}$$

where Φ is an arbitrary function, with ghost number -1. This difference will affect the full S matrix. Let us consider a transition process from a physical state $|\psi\rangle$ to a physical state $|\phi\rangle$, using $\mathcal{L}$ and using $\tilde{\mathcal{L}}$, at the lowest order we get,

$$\langle\phi| S|\psi\rangle - \langle\phi|\tilde{S}|\psi\rangle = i\langle\phi|\int d^4x(\mathcal{L} - \tilde{\mathcal{L}})|\psi\rangle. \tag{1.37}$$

Using (1.36) gives us

$$\langle\phi| S|\psi\rangle - \langle\phi|\tilde{S}|\psi\rangle = -\langle\phi|\{Q, \Phi\}|\psi\rangle = 0, \tag{1.38}$$

since both states are physical.

So, the physical states $|\psi\rangle$ are the ones which satisfy

$$\begin{aligned} Q\,|\psi\rangle &= 0 \\ |\psi\rangle &\neq Q\,|\rangle\,. \end{aligned} \tag{1.39}$$

Since the first line is $KerQ$ and the second is ImQ, usually the BRST cohomology is defined as

$$physical\,space = \ker Q/\ ImQ. \tag{1.40}$$

In practice, we can add any vector from the ImQ to the physical vectors without altering transition amplitudes. We should also note, that for each physical state the ghost number should be strictly zero to maintain the above procedure (see [28] for rigorous derivation).

1.3 Chiral Anomalies

In the section, we will review chiral anomalies. We present the derivation via Fujikawa method [31–33]. For this purpose, we take the Lagrangian of a massless Dirac fermion which is coupled to a YM-field. This system is described by

$$\mathcal{L} = i\bar{\psi}\not{D}\psi - \frac{1}{4}G^2, \tag{1.41}$$

where ψ is a fundamental Dirac fermion and G is a gauge field strength tensor (the discussion is also valid for an Abelian field), D is the corresponding covariant derivative. This Lagrangian is invariant under two symmetries,

$$\begin{aligned} \psi &\to e^{i\alpha}\psi, \\ \psi &\to e^{i\beta\gamma_5}\psi, \end{aligned} \tag{1.42}$$

where α and β are constants. The first transformation is called a vectorial transformation, while the second is axial.

The measure (1.5) is invariant under the vectorial transformation, while it is not for the axial one. To show this rigorously, let us write ψ and $\bar{\psi}$ in some basis

$$\begin{aligned} \psi &= \sum_n a_n\phi_n \\ \bar{\psi} &= \sum_n b_n\phi_n^+, \end{aligned} \tag{1.43}$$

where all information about the Fermi statistics is included in the as and bs, while ϕs are just functions. For our purposes we pick a basis where ϕs satisfy the following properties

$$\not{D}\phi_n = \lambda_n\phi_n \tag{1.44}$$

$$\langle\phi_n|\phi_m\rangle = \delta_{mn}. \tag{1.45}$$

They are mutually orthogonal, normalized to unity, and they are eigenstates of the covariant derivative. In such basis, the measure of the fermions can be written in the form

$$D\psi D\bar{\psi} \to \prod_n da_n db_n. \tag{1.46}$$

By symmetry transformations (1.42) we get a transformation rule for the measure,

$$a'_n = \sum_m \int d^4x \phi_n^+ e^{i\gamma_5\beta} \phi_m a_m = \sum_m C_{mn} a_m. \tag{1.47}$$

From the above expression the transformation rule for the measure (1.46) will be

$$\prod_n da'_n = (detC)^{-1} \prod_n da_n. \tag{1.48}$$

Considering the identity

$$\ln(detA) = tr(\ln A), \tag{1.49}$$

for the infinitesimal β, we get the Jacobian,

$$\ln\Big((detC)^{-1}\Big) = -i\beta \int d^4x \sum_k \phi_k^+ \gamma_5 \phi_k. \tag{1.50}$$

In general, the summation over k is ill-defined. Therefore, it must be regularized in a Lorentz covariant way. For that purpose, we introduce a scale M and multiply every term by the factor $e^{-\frac{\lambda_n^2}{M^2}}$. Then we get

$$\sum_k \phi_k^+ \gamma_5 \phi_k = \lim_{M\to\infty} \sum_k \phi_k^+ \gamma_5 \phi_k e^{-\frac{\lambda_n^2}{M^2}} = \lim_{M\to\infty} \sum_k \phi_k^+ \gamma_5 e^{-\frac{\not{D}^2}{M^2}} \phi_k. \tag{1.51}$$

The last expression can be derived by changing the eigenvalues with the appropriate operator.

The result can be rewritten in a basis-free form

$$\lim_{M\to\infty} \sum_k \phi_k^+ \gamma_5 e^{-\frac{\not{D}^2}{M^2}} \phi_k = \lim_{M\to\infty x\to y} Tr\gamma_5 e^{-\frac{\not{D}^2}{M^2}} \delta(x-y), \tag{1.52}$$

where we considered that ϕs form a complete set

$$\sum_k \phi_k^+(x)\phi_k(y) = \delta(x-y). \tag{1.53}$$

Now, we can rewrite the square of the covariant derivative in a more convenient form

$$\not{D}^2 = D_\mu D_\nu \left(\eta_{\mu\nu} + \frac{1}{2}[\gamma_\mu, \gamma_\nu]\right) = D^2 + \frac{g^2}{4}[\gamma_\mu, \gamma_\nu] G_{\mu\nu}, \tag{1.54}$$

where g is a coupling constant. Let us consider configurations near the vacuum, which means that the covariant derivative is so weak that it coincides with the normal one

$$\lim_{M\to\infty x\to y}\int\frac{d^4k}{(2\pi)^4}Tr\gamma_5 e^{(-D^2-\frac{g^2}{4}[\gamma_\mu,\gamma_\nu]G_{\mu\nu})/M^2}e^{ik(x-y)}=$$
$$\lim_{M\to\infty}\left(\frac{g^2}{16}Tr\gamma_5([\gamma_\nu,\gamma_\mu]G_{\mu\nu})^2\frac{1}{2!M^4}\int\frac{d^4k}{(2\pi)^4}e^{-k^2/M^2}+O(1/M)\right). \quad (1.55)$$

So, finally, the Jacobian corresponding to the transformation (for as) is

$$detC^{-1}=e^{i\beta\int d^4x\frac{g^2}{32\pi^2}G_{\mu\nu}\tilde{G}_{\mu\nu}}. \quad (1.56)$$

Here we assume that the term $G\tilde{G}$ is summed up over all indices, internal, and Lorentz. The transformation of the bs leads us to the same Jacobian, we do not need to repeat the calculation. Finally, the transformation rule is,

$$D\psi D\bar{\psi}\to D\psi D\bar{\psi}e^{i\int d^4x\beta\frac{g^2}{16\pi^2}G\tilde{G}}. \quad (1.57)$$

We see that this symmetry shifts the Lagrangian in the following way

$$\mathcal{L}\to\mathcal{L}+\beta\frac{g^2}{16\pi^2}G\tilde{G}. \quad (1.58)$$

These results, with combination of normal Ward identities (1.7), will correct the axial current divergence. It will have 2 terms, the first one is a result of Nöther's derivation while the second corresponds to the change of the measure. The divergence of the axial current j_5 has the form

$$\partial_\mu j_{5\mu}=\frac{g^2}{16\pi^2}G\tilde{G}, \quad (1.59)$$

The above result can also be derived from general considerations. Using a well-known index theorem [34], we can write

$$\Delta n_L-\Delta n_R=\int d^4xA(R)\wedge Ch(F), \quad (1.60)$$

where Δn_L and Δn_R are changes (in time) of numbers of left and right fermion modes respectively (see e.g., [35]), A is the characteristic of the representation ($\mathcal{A}$-genus, see e.g., [36]) and $Ch(F)$ is the Chern class proportional to $G\tilde{G}$ in this case. The above equation for $U(1)$ axial anomaly can be obviously written as (1.59).

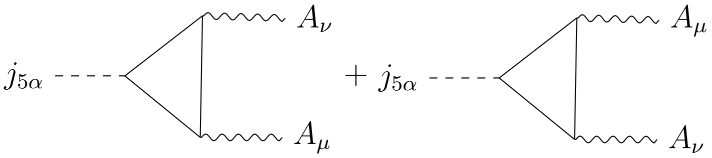
The equation (1.59) implies that two gauge bosons (γ) can be attached to that current with a non-trivial amplitude

$$\langle\gamma\gamma|\,\partial_\mu j_{5\mu}\,|0\rangle\neq 0. \quad (1.61)$$

For further discussions let us consider an amplitude, where we will keep the axial current to be general with an arbitrary index A. The index A combines a group index and an enumeration of different currents together. Nevertheless, we assume that it corresponds to some general axial symmetry, and hence, it is just a group index. Let us consider the amplitude from two gauge bosons to the vacuum of the axial current divergence,

$$\int d^4x e^{iqx} \langle p, k| \partial_\alpha j^A_{5\alpha} |0\rangle = (2\pi)^4 \delta(p+k-q) i q_\alpha \epsilon^{*B}_\mu(p) \epsilon^{*C}_\nu(k) \mathcal{M}^{ABC}_{\mu\nu\alpha}. \tag{1.62}$$

The amplitude can be computed diagrammatically. At the lowest order, it has the following form



. (1.63)

The first diagram will give us the result

$$(-1)(-ig)^2 \int \frac{d^4l}{(2\pi)^4} tr\left(\gamma_\alpha T^A \gamma_5 \frac{i(\not{l}-\not{k})}{(l-k)^2}\gamma_\nu T^C \frac{i\not{l}}{l^2}\gamma_\mu T^B \frac{i(\not{l}+\not{p})}{(l+p)^2}\right), \tag{1.64}$$

where we get a relative sign because of the closed fermionic loop. We get a vertex factor of $-ig$ for each gauge interaction and a corresponding generator T for each of them. The second diagram will give us the same contribution, but with interchanged gauge particles. Let us take T-generator terms and trace them

$$tr(T^A T^B T^C) = \frac{1}{2} f^{ABC} + \frac{1}{2} d^{ABC}, \tag{1.65}$$

where d and f are totally symmetric and antisymmetric group structure constants. The antisymmetric structure constant corresponds to the normal Ward identities having the same contribution as non-anomalous transformations. It is obvious that gauge bosons will vanish between them (bosons are symmetric). So, the expression vanishes for non-anomalous symmetries, while the anomalous transformations survive after symmetrization. The direct calculation shows that the normal parts of ward identities are still satisfied, for example,

$$p_\mu \mathcal{M}^{ABC}_{\mu\nu\alpha} = 0. \tag{1.66}$$

In the case of anomalous index, this identity is not satisfied, since acting on the amplitude with the q-momentum does not give us zero. From the first part of the diagram, we get (up to group structure constants):

$$g^2 \int \frac{d^4l}{(2\pi)^4} tr\left(\gamma_5 \frac{i(\not{l}-\not{k})}{(l-k)^2}\gamma_\nu \frac{\not{l}}{l^2}\gamma_\mu + \gamma_5 \frac{\not{l}}{l^2}\gamma_\mu \frac{i(\not{l}-\not{p})}{(l-p)^2}\gamma_\nu\right). \tag{1.67}$$

Through simple power-counting, we can see that this expression is divergent and needs to be regulated. The Final result should be proportional to the ϵ-tensor because it includes γ_5 with 4 gamma matrices. We should be careful with γ matrices during the dimensional regularization. The matrix γ_5 is well-defined only in $4D$, which means that this expression will not be shift-symmetric w.r.t l. Therefore, it can be non-zero in the sum with the second diagram (naively, it seems that shifting the second diagram and bringing in the form of the first one the two cancel each other). We can write down that the contribution should be proportional to a general expression which has the following form,

$$q_\alpha \mathcal{M}^{ABC}_{\mu\nu\alpha} \propto d^{ABC} g^2 \epsilon_{\mu\nu\alpha\rho} p_\alpha k_\rho. \tag{1.68}$$

It gives us the ϵ-tensor which was expected from the trace. It is clear that normal Ward identities are satisfied due to the properties of the totally antisymmetric tensor. Divergence of the amplitude coincides with the path integral derivation. It seems that (1.68) equals to full re-summation of all orders of the perturbation theory. This can be easily explained. This diagram is not the lowest order in perturbation series, but it is an exact result. On the other hand, we can check that all high order diagrams, which give contributions to this 3-point function, vanish. As they do not need to be regularized, meaning that they admit a shift symmetry, they will be canceled pair-wisely.

This result simplifies our calculations a lot. We can skip detailed derivations of anomalies and just calculate group index structures for some representations. The rest is the same for any anomaly. By checking of the group index structures, we conclude the anomaly property of the symmetry. Non-anomalous ones have trivial structures. An important remark is that if some system admits anomaly at some scale, the system should admit it at any scale, since the anomalies are an exact result of path integral derivation.

Gravity admits a global chiral anomaly too (see, e.g., [33]). For simplicity, we can take the measure transformation (1.55) and extend it to the presence of gravity. Instead of ϵ, we should use the Riemannian version of it, and we should consider that the commutators of the gravitational covariant derivatives give us the Riemann tensor. Hence, taking all these facts together, the final result can be written in a simple form

$$\delta\mathcal{L} \propto \beta \epsilon_{\mu\nu\alpha\beta} R^{\mu\nu}_{\rho\sigma} R^{\alpha\beta\rho\sigma}. \tag{1.69}$$

So far, we discussed global anomalies, their presence in the theories is consistent. They just bring new physical phenomena, i.e., axial current can decay into two gauge bosons, etc. Now let us discuss what happens in case of anomalous gauge symmetry.

From the equation(1.27) it is clear that fermionic current is a part of the BRST charge

$$J^{BRST}_\mu \supset g\bar{\psi}\gamma_\mu T^A \psi c^A. \tag{1.70}$$

Now, let us imagine, that ψ is a chiral fermion ($d^{ABC} \neq 0$), for instance $\psi = \frac{1+\gamma_5}{2}\psi$, then from (1.59)

$$\partial_\mu \bar{\psi}\gamma_\mu T^A \psi = g^2 \frac{d^{ABC}}{32\pi^2} G^B \tilde{G}^C. \tag{1.71}$$

Hence, for the BRST current we get

$$\partial_\mu J^{BRST}_\mu = g^3 \frac{d^{ABC}}{32\pi^2} G^B \tilde{G}^C c^A. \tag{1.72}$$

We can see clearly that it is not conserved, during the evolution we can end up with outside physical states. Conservation of the BRST charge was essential (1.33) for the S-matrix unitarity too. So, to maintain theory self-consistent, d^{ABC} for the matter content must vanish, or in other words gauge anomalies must cancel to maintain BRST symmetry exact.

Now we have a complete picture of self-consistency. The gauge systems should be quantized via BRST symmetry, their Physical states live in the kernel of the BRST charge and the BRST symmetry should not be anomalous to maintain self-consistency.

1.4 Exterior forms and gauge redundancy

In this section, we will describe gauge fields in the formalism of exterior forms (this introduction follows [37], for a review see e.g., [24]). Let us first consider a simple case, when the theory has only one Maxwell field A. In the formalism of the forms it can be written as (one form)

$$A = A_\mu dx^\mu. \tag{1.73}$$

After taking the exterior derivative of it, we get

$$F = dA = \partial_\nu A_\mu dx^\nu \wedge dx^\mu. \tag{1.74}$$

The expression is a usual field strength tensor $F_{\mu\nu}$ in index notation. It is easy to understand Bianchi identities in this language. As the exterior derivative (d) of the derivative is zero ($dd = 0$) and the wedge product is antisymmetric, we immediately get

$$\partial_{[\alpha} F_{\mu\nu]} = 0 \rightarrow \partial_\mu \tilde{F}_{\mu\nu} = 0, \tag{1.75}$$

where square brackets denote anti-symmetrization over the enclosed indices. Using the properties of the derivative, we can consequently see gauge redundancy of the field A. We can do the following transformations,

$$A \rightarrow A + da, \tag{1.76}$$

where a is a scalar (zero form). We observe that the field strength does not change under such transformations. To write the equations of motion, it is useful to use a Hodge star operator. This is defined as

$$*(dx^{\mu_1} \wedge ... \wedge dx^{\mu_n}) = \frac{1}{(4-N)!} \epsilon^{\mu_1 ... \mu_N \nu_1 ... \nu_{4-N}} dx^{\nu_1} \wedge ... \wedge dx^{\nu_{4-N}}. \tag{1.77}$$

We can introduce a vector current

$$^*J = J_\mu dx^\mu. \tag{1.78}$$

Now we can easily write Maxwell equations in this formalism

$$d^*F = J. \tag{1.79}$$

It is clear from this expression that the current conservation is required (just acting on the equation with d).

It is essential to count propagating and non-propagating degrees of freedom. From (1.76) we see that we can perform an arbitrary shift by scalar a, but let us fix a gauge condition. For instance,

$$ {}^*d^*A = 0, \tag{1.80} $$

which corresponds to the gauge $\partial A = 0$ in index notation. It is clear, that for an arbitrary a, we cannot do gauge transformations anymore. It looks like that the gauge is completely fixed, and we used one scalar condition for that. This scalar itself is a fully legitimate propagating degree of freedom. Three propagating degrees of freedom are left out of four. Let us take a look at the following identity,

$$ {}^*d^*d = \Box\,. \tag{1.81} $$

If we try to do gauge transformation with some field a keeping invariant (1.80), we get,

$$ \Box a = 0\,. \tag{1.82} $$

Since this equation has non-trivial solutions, it means that a previous gauge fixing condition was incomplete. The gauge fixing condition still left one degree of freedom in gauge transformations. However, this degree of freedom does not have any source. This means that it is an only propagating degree of freedom. Hence, only two propagating degrees of freedom are physical out of four and another one is a non-propagating degree of freedom, which corresponds to the static (classical) interaction (Coulomb force).

We can count the degrees of freedom in other ways too. The field A_0 does not have a second time derivative, so it is a pure constraint. We can integrate it out, which will generate terms, like $\rho\frac{1}{\nabla^2}\rho$ in the energy (Coulomb force). We call this a non-propagating degree of freedom. After integrating out A_0, the system still has gauge redundancy $A_j \to A_j + \partial_j\alpha$, leaving us with 2 propagating degrees of freedom. Here we see that previous discussion is consistent with explicit counting. The treatment just give us a shortcut to count degrees of freedom. It is very useful for the forms with higher rank, which we see shortly.

The next important aspect is a duality transformation. A duality transformation of the Abelian vector field is trivial, but it shows us how the machinery works. Let us consider the pure Maxwell Lagrangian

$$ \mathcal{L} = -\frac{1}{4}F_{\mu\nu}F_{\mu\nu}. \tag{1.83} $$

Let us take F as an independent variable, neglecting its dependence on A. We can see that Bianchi identities (1.75) are not necessarily satisfied anymore. Hence, we have to impose them as constraints. As we need to take the antisymmetric derivative of field strength, we should use a derivative traced with epsilon. Then the Lagrangian will take the form

$$ \mathcal{L} = -\frac{1}{4}F_{\mu\nu}F_{\mu\nu} + B_\alpha\epsilon_{\beta\alpha\mu\nu}\partial_\beta F_{\mu\nu}, \tag{1.84} $$

where B_μ is a Lagrange multiplier. Taking a variation w.r.t. F will give us the following equation of motion

$$-\frac{1}{2}F_{\mu\nu} - \epsilon_{\mu\nu\alpha\beta}\partial_\alpha B_\beta = 0. \tag{1.85}$$

Solving F and putting back the solution into the Lagrangian, we will get (during this discussion we will use the notation $dB = F(B)$)

$$\mathcal{L} = -F(B)_{\mu\nu}F(B)_{\mu\nu}. \tag{1.86}$$

We see that a duality transformation gives us a Maxwell Lagrangian for B. A magnetic field of B-field corresponds to an electric one of A-field and vice versa (1.89). We do not get an entirely new description by duality transformation in this case. Furthermore, we will see that duality transformations give us an important tool for more complicated cases.

The next important aspect to discuss is giving a mass to a field. We can write down directly the so-called Proca Lagrangian for the vector field, $\tilde{A}$

$$\mathcal{L} = -\frac{1}{4}\tilde{F}^2 + \frac{1}{2}m^2\tilde{A}^2, \tag{1.87}$$

where we assume the summation over all appropriate indices, and we do not explicitly write them down. It looks like that the gauge redundancy is lost in this theory. This theory is not invariant under $\tilde{A} \to \tilde{A} + da$. This means that this theory should exhibit pathologies because an analysis of a Lorentz group shows us the necessity of the gauge redundancy. It is clear from the Lorentz group that a massive vector should have 3 polarizations. To see the gauge symmetry manifestly, we should decompose $\tilde{A}$ in the following way

$$\tilde{A} = A + \frac{1}{m}db, \tag{1.88}$$

with the requirement that

$$\begin{aligned} A &\to A + da \\ b &\to b - ma. \end{aligned} \tag{1.89}$$

We see that the gauge redundancy is not lost, it was just hidden. In terms of the fields A and b, the Lagrangian will take the form

$$\mathcal{L} = -\frac{1}{4}F^2 + \frac{1}{2}m^2(A + \frac{1}{m}db)^2. \tag{1.90}$$

It is manifestly clear that the above Lagrangian is gauge invariant and describes 3 propagating degrees of freedom, as it was expected. According to this formalism, Stückelberg field becomes physical in the massive case.

We can also double-check the description, taking the limit when $m \to 0$,

$$\mathcal{L} = \left(-\frac{1}{4}F^2 + \frac{1}{2}m^2A^2 + \frac{1}{2}(\partial b)^2 + mA_\mu\partial_\mu b\right)|_{m\to 0} = -\frac{1}{4}F^2 + \frac{1}{2}(\partial b)^2. \tag{1.91}$$

From the above Lagrangian, we can see the decoupling limit. In the limit, we still get a massless Maxwell field and a massless goldstone scalar. Under the gauge transformation (1.89) in the limit, only the vector field transforms and the scalar is invariant under the constant shifts.

We see that the exterior forms give us a beautiful description for gauge vector fields, including the mass term. Now we try to generalize it for more sophisticated examples.

First, let us consider 2-forms. They can be written in the above formalism,

$$\omega = \omega_{\mu\nu} dx^\mu \wedge dx^\nu. \tag{1.92}$$

Their field strength is

$$F = d\omega\,. \tag{1.93}$$

The presence of gauge redundancy is clear, since the field strength is invariant under

$$\omega \to \omega + dB, \tag{1.94}$$

where B is a one-form. ω-field has 6 components. Obviously, not all of them will be physical degrees of freedom due to the gauge redundancy. Let us fix the gauge by imposing a Lorenz constraint (as it was done in the case of the vector field)

$$\partial_\mu \omega_{\mu\nu} = 0. \tag{1.95}$$

Let us consider ω', gauge transformed field, which does not satisfy this constraint. We can do a gauge transformation with one-form B to bring this field back to ω and to satisfy the constraint

$$0 = \partial_\mu \omega_{\mu\nu} = \partial_\mu \omega'_{\mu\nu} + \partial_\mu F(B)_{\mu\nu}, \tag{1.96}$$

where the quantity $F(B)$ is a field strength for B. Since the above equation is a Maxwell equation for B with the source $\partial_\mu \omega'_{\mu\nu}$, from vector field analysis we know that B carries two propagating and one non-propagating degrees of freedom. After fixing the gauge, we can see that 3 degrees of freedom are not physical out of 6. Similarly to the vector field, we still have ambiguity of gauge transformations. It is possible to do gauge transformations of ω with a source-less B' vector field, i.e., it has a divergence-less field strength. Such transformations do not change the gauge fixing constraint. As we have seen, a sourceless vector field has only two propagating degrees of freedom. For a 2-form, it means that 2 degrees of freedom are not propagating out of 3. So, the 2-form has two non-propagating and one propagating degrees of freedom.

Now we can derive the equations of motion for a 2-form. Foremost, we should build the Lagrangian. We start from a massless case with some external source J, which, in fact, is 2-form too. The Lagrangian has the following form,

$$\mathcal{L} = \frac{1}{12} F^2_{\mu\nu\rho} + \omega_{\mu\nu} J_{\mu\nu}. \tag{1.97}$$

The current should respect the gauge redundancy, which means that it must be conserved (at least) on-shell. This 2-form will carry one propagating degree of freedom and will create non-propagating static potentials between charges. The equation of motion has the following form

$$\partial_\mu F_{\mu\alpha\beta} = J_{\alpha\beta}. \tag{1.98}$$

It is obvious that the field strength tensor satisfies Bianchi identities, which follow from (1.93), in the index notation,

$$\partial_{[\mu} F_{\alpha\beta\gamma]} = 0\,. \tag{1.99}$$

Equations of motion of the 2-form with Bianchi identities give us a very useful description of one propagating degree of freedom including 2 non-propagating degrees of freedom with a gauge redundant construction. Intuition from the previous duality transformation says that propagating degrees of freedom of the 2-form should be dual to a scalar field. To perform the duality transformation, like in the case of the vector field, we should take a field strength as an independent variable and the Bianchi identity as a constraint. Integrating the 2-form field out gives us the dual description. A corresponding Lagrangian has the following form,

$$\mathcal{L} = \frac{1}{12}F^2 + a\epsilon_{\mu\alpha\beta\gamma}\partial_\mu F_{\alpha\beta\gamma}, \tag{1.100}$$

where a is a Lagrange multiplier. The field a is invariant under the shift symmetry,

$$a \to a + c, \tag{1.101}$$

where c is an arbitrary constant. The transformation adds a total derivative to the Lagrangian. Corresponding equations of motions are

$$\frac{1}{6}F_{\alpha\beta\gamma} - \partial_\mu a \epsilon_{\mu\alpha\beta\gamma} = 0. \tag{1.102}$$

putting the solution of the 2-form field strength back in the Lagrangian, we get

$$\mathcal{L} = 3(\partial a)^2. \tag{1.103}$$

This Lagrangian is clearly shift invariant (1.101). As a note, we could start with the shift invariant scalar and make the duality transformation to the 2-form. The procedure is straightforward. The gauge redundancy imposes strong constraints on model building. This duality is a remarkable tool, since the 2-form description has a gauge redundancy, we have more control under the dynamical system. We will talk about this in details, during the next chapter about the axion.

Similarly to the vector fields, 2-forms may be massive. The construction should be done in a gauge invariant way too. In this case, Stückelberg field is a one-form B. Such Lagrangian with a mass m has the form

$$\mathcal{L} = \frac{1}{12}F^2 - \frac{1}{2}m^2(\omega + \frac{1}{\sqrt{2}m}dB)^2, \tag{1.104}$$

which is invariant under

$$\begin{aligned} \omega &\to \omega + dA \\ B &\to B - \sqrt{2}mA, \end{aligned} \tag{1.105}$$

since we combined 2 objects, which have 3 propagating degrees of freedom, the field $\omega' = \omega + \frac{1}{\sqrt{2}m}dB$ carries 3 propagating degrees of freedom. A dual picture of this configuration is clear; the 2-form becomes Stückelberg of a vector, which is dual to the one form B. Therefore, this Lagrangian is dual to the massive vector description. Again, similarly to the vector field, a decoupling limit holds and counting of degrees of freedom is not altered. However, this picture has an advantage, all the components are gauge redundant quantities, which provides us a powerful tool to deal with the dynamical systems.

To complete our discussion, let us consider a 3-form C

$$C = C_{\alpha\beta\gamma} dx^\alpha \wedge dx^\beta \wedge dx^\gamma, \tag{1.106}$$

The corresponding field strength is

$$F = dC. \tag{1.107}$$

We can shift this 3-form by the derivative of a 2-form Ω without changing the field strength. Due to its antisymmetric nature, the field C has four components, but not all of them are independent degrees of freedom, because of the gauge redundancy. As it was done in the case of the 2-forms, we can choose a Lorenz gauge for the C-field and do a gauge transformation. The result will be C'

$$0 = \partial_\mu C_{\mu\alpha\beta} = \partial_\mu C'_{\mu\alpha\beta} + \partial_\mu F(\Omega)_{\mu\alpha\beta}, \tag{1.108}$$

here $F(\omega) = d\Omega$ is a field strength tensor for Ω-field.

We see that this is an equation of motion for Ω with the source $\partial_\mu C'_{\mu\alpha\beta}$. The 2-form carries 1+2 degrees of freedom, which means that 3 degrees of freedom out of 4 are not physical. We are left with one degree of freedom, but there is a possibility to do gauge transformations with a sourceless Ω' 2-form, under which the field strength is invariant. We can conclude that the 3-form does not carry any propagating degrees of freedom, but it has one non-propagating degree of freedom.

Now we can build a Lagrangian for the massless field C. It has the following form,

$$\mathcal{L} = -\frac{1}{48}F^2 - J_{\alpha\beta\gamma}C_{\alpha\beta\gamma}, \tag{1.109}$$

where J is a conserved current. The equation of the motion has the form,

$$\partial_\mu F_{\mu\alpha\beta\gamma} = J_{\alpha\beta\gamma}\,. \tag{1.110}$$

An interesting thing is that we do not have Bianchi identities in this case. Maximum dimension of antisymmetric form is 4 in four dimensions. The field strength of a 3-form already has that dimension. Bianchi identities require anti-symmetrization with the derivative. Therefore, it must be 5-dimensional in this case, and it is impossible to write down any Bianchi identities in case of 3-forms. This means that we cannot apply duality transformations in the same manner as we did for lower-rank forms. However, since totally antisymmetric 4-form is Hodge dual to a scalar, the field strength tensor is always dual to some scalar.

$$^*F = \frac{1}{24}\epsilon_{\mu\alpha\beta\gamma}F_{\mu\alpha\beta\gamma} = -E, \tag{1.111}$$

where E is called an “electric” field. We can see that all information about the 3-form is encoded in one function, which was expected by counting degrees of freedom. The field does not have any propagating degrees of freedom, it is just a constraint. This fact makes massless 3-forms very useful for description of boundary conditions. On the one hand, an electric field is a boundary term, while on the other hand, it is a field strength of a massless 3-form.

Similarly to the lower rank forms, it is possible to give mass to 3-forms, without violation of the gauge redundancy. So, we should make Stückelberg field physical. Let us consider a 3-form C and a two-form B, then a massive Lagrangian with mass m is

$$\mathcal{L} = -\frac{1}{48}F^2 + \frac{1}{2}m^2(C + \frac{1}{\sqrt{6}m}dB)^2, \tag{1.112}$$

and it has the following gauge invariance

$$\begin{aligned} C &\to C + d\Omega \\ B &\to B - \sqrt{6}m\Omega, \end{aligned} \tag{1.113}$$

The combination $C' = C + \frac{1}{\sqrt{6}m}dB$ carries one propagating degree of freedom. It is dual to a massive scalar field.

In this chapter we saw, that high rank forms are dual to scalar or to constraints. One can ask why we should even consider them, since the dual picture is much less difficult to treat. In the literature, we see scalar with a mass, not the massive three forms. The answer is obvious, in this case we have the powerful BRST symmetry, since we are dealing with gauge redundant objects. The symmetry controls all possible options to build up Lagrangian and all possible options of the types of quantum corrections the system can get. The dual description in that sense is more powerful than usual textbook descriptions. We will see, during the next chapters, usage of this language, for instance with the strong-CP problem.

Chapter 2

Large-N / Semi-classical Physics

We usually think that adding degrees of freedom increases the complexity of the system. At least two contradictions to this exist, the super-symmetry (see e.g., [38–40]) and large N physics (see e.g., [41, 42]). The first one adds extra symmetry, and all simplifications arise due to new conserved quantities. We won't discuss it in this chapter. The second one increases *species* of the fields, taking pairwise interaction negligible, but keeping a collective coupling fixed. It turns out, that in this limit many simplifications happen and in some cases Semi-classical physics is exact (see, e.g., [43]). We can have similar to large-N limit, if we populate a large number of quanta, in an ordinary non-large-N theories. During this chapter we will use a term "large-N" for both cases, we will not specify which number meant to be large, unless it is not clear from the context.

In this chapter we will discuss large-N systems, their simplifications, a semi-classical approximation, importance of classical topological configurations and new-found universal phenomena of saturation, including the Saturons [8–10]. We discuss the above phenomena in details in the context of Gross-Neveu model following our article [2].

2.1 The Gross-Neveu Model

In a well-known work by Gross and Neveu [11], a 2-dimensional quantum field theoretic model was introduced (for the review, see, [41, 42]). The model is integrable in a large-N limit, the chiral symmetry is spontaneously broken, and the bound states are known [44]. In this section, we will mostly review [11] and the other aspects will be presented later.

The Gross-Neveu model consists of N Dirac ψ fermions in 2 dimensions.

$$\mathcal{L} = i\bar{\psi}\partial\!\!\!/\psi + \frac{\alpha}{2}\left(\bar{\psi}\psi\right)^2 , \tag{2.1}$$

where α is a dimensionless coupling constant of the theory. The contractions of flavour and space-time indices are obvious and not shown explicitly. The theory has a discrete chiral symmetry, which acts as $\psi \to \gamma_5\psi$ in Dirac notations. The above theory has an interaction term with a negative sign in the Hamiltonian. To address this issue, we can integrate an auxiliary field σ in and write down a fully equivalent theory,

$$\mathcal{L} = i\bar{\psi}\partial\!\!\!/\psi - \frac{1}{2\alpha}\sigma^2 + \sigma\bar{\psi}\psi \,. \tag{2.2}$$

The field σ is a singlet under the flavour group and transforms under the discrete chiral symmetry as $\sigma \to -\sigma$. Obviously, we can interpret the new field as a very heavy one, and hence it does not have derivative terms. Here, we adopted a normalization for the σ field (2.2) used in [42], hence, it is not canonically normalized. This normalization is a matter of choice, since the physical quantities, such as masses of particles and the vacuum expectation value of the $\bar{\psi}\psi$ operator, are independent of this normalization.

Usually, the model is considered in a large-N limit. In the limit, a 't Hooft-like coupling [45] is defined as

$$\lambda = \alpha N \,, \tag{2.3}$$

and $N \to \infty$, where λ is kept fixed. In this case, the only correction to σ propagator is exact in one loop (2.1). We can see from the picture, that the extra loops give sub-leading contribution.

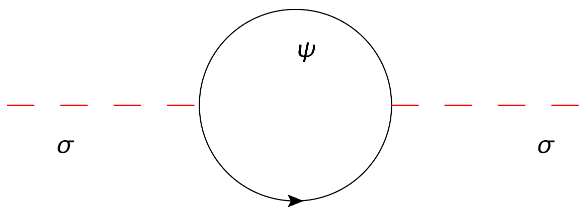


Figure 2.1: One loop correction to σ

We can compute one loop contribution using the path integral. We should integrate over fermionic fields, since the integration over the σ-field just restores the original form of the Lagrangian (2.1). We also want to investigate a vacuum structure of the theory. So, we should integrate over fermions to compute effective potential [46] for the σ,

$$\int \mathcal{D}\sigma\mathcal{D}\psi\mathcal{D}\bar{\psi}e^{i\int d^2x\mathcal{L}} = \int \mathcal{D}\sigma\,\det\Big(i\partial\!\!\!/ + \sigma\Big)e^{-\frac{i}{2\alpha}\int d^2x\sigma^2}. \tag{2.4}$$

The result tells us that we have the following effective action

$$S = \int d^2x \frac{-1}{2\alpha}\sigma^2 - i\,\mathrm{Tr}\log\Big(i\partial\!\!\!/\mathbb{I} + \sigma\mathbb{I}\Big). \tag{2.5}$$

We are currently interested in only computation of effective potential, and hence, we can take σ to be constant. We can also take trace over the flavour structure to get

$$S = -VT\frac{1}{2\alpha}\sigma^2 - \frac{iN}{2}\,\mathrm{Tr}\log\Big(\partial^2 + \sigma^2\Big). \tag{2.6}$$

Taking trace in the spinor space, rewriting the expression in the momentum space and using dimensional regularization $d = 2 - 2\epsilon$ result in

$$V_{eff} = -\frac{1}{2\alpha}\sigma^2 - iN\tilde{\mu}^{2\epsilon}\int\frac{d^dp}{(2\pi)^d}\log\Big(-p^2 + \sigma^2\Big), \tag{2.7}$$

where effective scale $\tilde{\mu} = \mu e^{1-\gamma}\sqrt{4\pi}$, γ is an Euler-Mascheroni constant. Taking integration explicitly gives

$$V_{eff} = -\frac{1}{2\alpha}\sigma^2 - \frac{N\sigma^2}{4\pi}\left(1 - \frac{1}{\epsilon} + \gamma - 1 + \log\left(\frac{\sigma^2}{\tilde{\mu}^2}4\pi\right)\right),\,. \tag{2.8}$$

After removing divergence by renormalization and writing everything in the terms of μ, we get the final well-known expression [11]

$$V = \frac{N}{2\lambda}\sigma^2 + \frac{N}{4\pi}\sigma^2\left(\log\left(\frac{\sigma^2}{\mu^2}\right) - 1\right), \tag{2.9}$$

where coupling $\lambda(\mu)$ depends on a renormalization scale. The above potential has two minima at $\sigma_0 = \pm\mu e^{-\frac{2\pi}{\lambda}}$. This implies a spontaneous breaking of the chiral symmetry, and the result above is exact for an infinite N.

Since the minima of the potential should not depend on the renormalization scale, we can derive the beta function.

$$\mu\frac{d}{d\mu}\sigma_0 = 0, \tag{2.10}$$

giving us

$$\mu\frac{d}{d\mu}\lambda = -\frac{\lambda^2}{2\pi}, \tag{2.11}$$

which implies a negative beta function,

$$\beta(\lambda) = -\frac{\lambda}{2\pi}. \tag{2.12}$$

So, the theory is asymptotically free, and the chiral symmetry is broken. From the Lagrangian (2.2), we can see that fermion gets a mass,

$$m_f = \sigma_0 \,. \tag{2.13}$$

In the spectrum [44] (will be discussed in the following sections), the lightest particle is the fermion. This implies freezing of beta function at $\mu \sim \sigma_0$ scale, which corresponds to $\lambda \sim 1$. This scale will be an important during the next discussions.

2.2 Classical solutions in the Quantum field theory

In the previous section, we saw a close relation between large-N physics and validity of the semi-classical approximation. The semi-classical approximation gives us the exact effective potential. It appears that semi-classical approximation is a good tool to find, for instance, bound states in certain Quantum field theories (see e.g., [43]). This procedure usually requires of finding classical solutions in quantum field theories. The most attractive solutions for this purpose are topological solitons/instantons and breathers. We will review some topological solitons/instantons in this chapter.

2.2.1 Kink in 2 dimensions

Let us consider one real scalar field in $1+1$ dimensions with a potential giving vev to the field at the minima [47] (for a review see e.g., [35, 39, 43])

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - \frac{\lambda}{4}(\phi^2 - v^2)^2. \tag{2.14}$$

The corresponding Hamiltonian for the above system reads

$$\mathcal{H} = \frac{1}{2}\dot{\phi}^2 + \frac{1}{2}(\phi')^2 + \frac{\lambda}{4}(\phi^2 - v^2)^2, \tag{2.15}$$

where $\dot{}$ and $'$ are time and space derivatives respectively. It is clear from the Hamiltonian that the theory has a vacuum when $\phi^2 = v^2$ and hence, two different choices, $\phi = \pm v$, exist. The Lagrangian and, obviously, the corresponding Hamiltonian have the symmetry $\phi \to -\phi$. If we do such transformation twice, the field returns into itself. This transformation forms the group $Z_2 = \{0, 1\}$. Despite the Z_2-symmetry invariance of Lagrangian, the vacuum does not fulfill the same property. If we transform the vacuum v, we will get the vacuum $-v$ and vice versa. This phenomenon is called a spontaneous symmetry breaking (SSB). The vacuum solutions are topologically trivial ones. We can look for non-trivial solutions, namely, the solutions which start in one vacuum and end in another. For this purpose, let us consider time-independent configurations, their equation of motion has the following form

$$-\phi'' = -\lambda(\phi^2 - v^2)\phi \,. \tag{2.16}$$

Applying the proper boundary conditions (starting from a $-v$ vacuum and ending in a v vacuum, or vice-versa) we get the following result

$$\phi' = \pm\sqrt{\frac{\lambda}{2}}(\phi^2 - v^2). \tag{2.17}$$

Foremost let us consider the + sign; it corresponds to the boundary conditions where the field equals to v at + spatial infinity, and it equals to $-v$ at $-$ infinity. The solution to this equation with the mentioned boundary conditions has the following form,

$$\phi = v \tanh\left(\frac{m}{\sqrt{2}}(z - z_0)\right), \tag{2.18}$$

where $m^2 = \lambda v^2$. If we have chosen opposite boundary conditions, this solution would be the same, but with an opposite sign. This solution is called kink and the one with the opposite sign - anti-kink. Let us consider the current,

$$J_\mu = \frac{1}{2v}\epsilon_{\mu\nu}\partial_\nu\phi, \tag{2.19}$$

where ϵ is a two-dimensional totally antisymmetric tensor, which obeys time and space-z indices. This current is trivially conserved by definition,

$$\partial_\nu J_\nu = \partial_\nu\partial_\mu\epsilon_{\mu\nu}\frac{\phi}{2v} = 0. \tag{2.20}$$

The charge corresponding to this current with appropriate boundary conditions gives us

$$Q = \int dz J_0 = \frac{1}{2v}\int dz\partial_z\phi = \frac{1}{2v}(\phi(\infty) - \phi(-\infty)) = \pm 1. \tag{2.21}$$

This means that kink and anti-kink (we will call both of them kinks, if context does not require any distinction) solutions are stable. They are protected by a conserved charge because this charge is zero for the vacuum $\partial_z v = 0$, while it is not for the above solution. The charge has a topological origin; it characterizes the vacuum manifold, and it is not a Nöther's charge.

An origin of this charge is topology. Due to spontaneous symmetry breaking, vacuum manifold is not trivial, it has the following form

$$\mathcal{M} = Z_2/\{1\} = Z_2. \tag{2.22}$$

So, it has disjoint regions, which cannot be mapped to each other by continuous deformations. Hence, we cannot destroy kink configuration with a continuous process. Disjoint regions of a manifold are counted by the zeroth homotopy group π_0 (see, e.g., [48]). In this case,

$$\pi_0(Z_N) = Z_N, \tag{2.23}$$

which means that kinks are stable configurations (can be checked explicitly, via stability of perturbations, see e.g., [35]).

We considered the kink, since we want to show that Semi-classical approximation gives amazingly good results. Semi-classical quantization of ϕ^4 kink was done in [49]. We will not review the results, since we want to demonstrate above methods for different models, which has a better overlap with our general discussion. Therefore, to avoid the discussion of the same topic twice, we skip it here.

2.2.2 Zero modes

Zero modes play a crucial role in the analysis of topological structures. They are similar to goldstone bosons, since each zero mode corresponds to some broken generator under which the action is still invariant while the solution is not. If we compare it to spontaneously broken symmetry, an action is invariant under the symmetry and the vacuum is not. The existence of zero modes gives us significant physical results. For instance, they increase the degeneracy of the object, which is essential for our future discussions.

Let us take a look at the kink solution (2.18). The solution has an integration constant z_0 which breaks translational invariance. Despite this fact, the action is invariant under translations $z \to z + a$, where a is a constant. Hence, if the action is invariant under the shift symmetry while a solution is not, we should expect zero modes as fluctuations around the solution. This is a very similar (or almost the same) phenomenon, as having goldstone bosons in spontaneous symmetry breaking. In the manuscript, we may use the term "zero modes" and "goldstones" interchangeable.

Let us consider fluctuations around the kink,

$$\phi = \phi_{kink} + \eta. \tag{2.24}$$

If we constrain ourselves to quadratic terms of η we get

$$\mathcal{L} = \frac{1}{2}(\partial\eta)^2 - \frac{1}{2}V''(\phi_{kink})\eta^2. \tag{2.25}$$

Here V is the potential of the kink Lagrangian (2.14). To simplify equations for fluctuations, we can make the following ansatz for η

$$\eta = \sum_j f_j e^{i\omega_j t}. \tag{2.26}$$

The equation of motion becomes Schrödinger-like,

$$\left(-\frac{d^2}{dz^2} + V''(\phi_{kink})\right) f_j = \omega_j^2 f_j. \tag{2.27}$$

For a rigorous proof for the existence of the zero-mode solution for this equation see, e.g., [35]. We will not prove this statement, but instead we will check it. Let us take the kink solution and differentiate it once,

$$\frac{d}{dz}\phi_{kink} = -\frac{d}{dz_0}\phi_{kink}. \tag{2.28}$$

It satisfies the equation of motion with $\omega_0 = 0$. Now let make this object time-dependent, giving a motion possibility to the kink. To study only the zero mode, we can integrate all other modes out. For this purpose, let us consider only its time-dependence. Now we insert the kink solution back into the Lagrangian with the dynamical zero mode

$$L = \int dz \left(\frac{1}{2}(\partial\phi)^2 - V(\phi)\right) = \int dz \left(\frac{1}{2}\dot{z_0}^2\phi'^2 - \frac{1}{2}\phi'^2 - V(\phi)\right), \tag{2.29}$$

where dot $\dot{}$ is a time derivative. We can integrate ϕ field over z, and get

$$L = \frac{1}{2}M\dot{z_0}^2, \tag{2.30}$$

where M is called kink mass. The Lagrangian shows that the kink's free motion is described by a collective time-dependent coordinate z_0. This leads us to the conclusion that a kink is a part of the spectrum as a free particle (at the lowest order of the perturbation theory).

A more interesting phenomenon is connected to fermions and their zero modes [50] (for the review, see, e.g., [35]). Let us consider a Dirac fermionic Lagrangian in the background of the kink,

$$\mathcal{L} = i\bar{\psi}\partial\!\!\!/\psi - G\bar{\psi}\phi\psi, \tag{2.31}$$

where G is a Yukawa coupling constant. In the vacuum, where $\phi = v$, the last term is just the fermion mass. In the kink background, this is not anymore an exact statement. To solve fermions in the Kink background, we should consider equations of motion from the Lagrangian and take an ansatz

$$\psi = \sum_j g_j e^{i\omega_j t}. \tag{2.32}$$

Hence, we get the equation,

$$\left(-i\gamma_0\gamma_1\partial_z + G\gamma_0\phi_{kink}\right) g_j = \omega_j g_j. \tag{2.33}$$

Let us concentrate on zero modes $\omega_0 = 0$ only. First of all, let us remind that the Dirac equation describes particles and anti-particles together, they carry opposite charges. In case of 2 dimensions, the charge conjugation operator has the form

$$g_j^c = -i\gamma_1 g_j. \tag{2.34}$$

Particle is a solution with a positive ω_j frequency, while anti-particle has a negative $-\omega_j$ one. A particle and an antiparticle are not linearly independent in the case of zero modes because they share the same zero frequency. We can see that Dirac equation is simplified for zero modes

$$0 = \gamma_0\left(-i\gamma_0\gamma_1\partial_z + G\gamma_0\phi_{kink}\right) g_j = \left(-i\gamma_1\partial_z + G\phi_{kink}\right) g_j. \tag{2.35}$$

So, the solutions can be chosen as eigenvalues of the γ_1 matrix

$$g_0 = f s_\pm, \tag{2.36}$$

where f is the z-dependent part of the equation and $s_{\pm}$ is an eigenspinor of the γ_1 matrix. Taking these facts into account, the solution can be written in the form

$$g_0 = e^{\mp G \int_0^z dz' \ \phi(z')_{kink}} s_{\pm}. \tag{2.37}$$

Different signs of s spinors correspond to solutions in kink and anti-kink background, respectively. One of them is normalizable in the kink background, while the other one in the anti-kink background. As it is known, zero modes must be normalizable solutions. Since there are 2 solutions for each matrix, zero modes are degenerated, namely each s has 2 solutions. If we consider N fermions, then the degeneracy of the zero modes will be 2^N. This degeneracy plays an important role, we will see it during the next chapters. In addition, these fermions carry fractional charge [50], but we will not discuss this beautiful phenomenon here, since it is beyond of the scope of the manuscript.

2.2.3 Cosmic Strings

We will not use the cosmic strings [51] in our discussion. They are still a nice prototype to discuss more sophisticated classical solutions than kinks and prepare our discussion for instantons. Let us consider a $U(1)$ charged scalar field in $1+2$ dimensions,

$$\mathcal{L} = |\partial\phi|^2 - \frac{\lambda}{2}\left(|\phi|^2 - v^2\right). \tag{2.38}$$

The above model has a non-trivial vacuum, so SSB happens. The vacuum has the following form,

$$\phi = ve^{i\alpha}, \tag{2.39}$$

where α is an arbitrary constant angle, and therefore we have a continuum number of vacua labeled by it. The vacuum manifold in this case is more sophisticated, than in the kink case,

$$\mathcal{M} = U(1)/\{1\} = U(1) = S^1. \tag{2.40}$$

With the above information we can seek for topologically stable solutions. In the kink case, we had 2 options and two infinities (infinities at one dimensional axis). In the case of the cosmic string, we have the continuous number of infinities, in 2D plane. Naively, we can continuously rotate between these infinities, but here we characterize vacuum manifold with one angle α. To see what kind of topological configurations we can have, let us consider the following quantity,

$$N[C] = \frac{1}{2\pi}\int d\vec{l}\vec{\nabla} arg\phi = \frac{1}{2\pi}\int d(arg\phi). \tag{2.41}$$

The integral counts full rotations for an arbitrary phase of ϕ. The field must be single-valued, namely, it should not change after 2π rotation. This means that the phase of the field should be periodic, with a period of 2π. The above integral of a periodic function over the period can give us only integers, winding numbers. Continuous deformations cannot change them, and hence they represent topological properties. Mathematically, this property is written in the following form [48]

$$\pi_1(U(1)) = \pi_1(S^1) = Z, \tag{2.42}$$

where π_1 is called the first homotopy group.

Unfortunately, topologically non-trivial solutions in the above model have infinite energy. The divergent part of the Hamiltonian only has an angle distribution. Schematically, the expression is

$$E \propto \int r dr \frac{1}{r^2} \left(\frac{d\alpha}{d\theta} \right)^2 \propto \int \frac{dr}{r} N^2 \propto \log r. \tag{2.43}$$

The energy for an angle distribution diverges logarithmically. Apparently, the reason appears to be "solvable", the theory lacks some degrees of freedom. By gauging the global $U(1)$, the problem is not present anymore. The corresponding Lagrangian has the following form

$$\mathcal{L} = |D\phi|^2 - \frac{\lambda}{2} \left(|\phi|^2 - v^2 \right)^2 - \frac{1}{4} F^2, \tag{2.44}$$

where D is an Abelian covariant derivative and F is the field strength for an Abelian gauge field (A). Similarly to the previous case, let us discuss time-independent configurations. Then only the spatial part of the covariant derivative survives in the temporal gauge ($A_0 = 0$), which is

$$\vec{D}\phi = e^{i\alpha}(\nabla|\phi| - i|\phi|(\nabla\alpha - e\vec{A})), \tag{2.45}$$

where e is a coupling constant. If we choose $\vec{A} = \frac{1}{e}\nabla\alpha$, the second term of the covariant derivative will vanish. This choice cancels divergence of the energy, which we saw in the previous case. Gauging the symmetry does not change the fact that the phase of the field is quantized (2.41); In this case we get,

$$N[C] = \frac{1}{2\pi} \int d\vec{l}\nabla\alpha = \frac{e}{2\pi} \int d\vec{l}\vec{A} = \frac{e}{2\pi} \int d\vec{s} \left(\nabla \times \vec{A} \right) = \frac{e}{2\pi} \int d\vec{s}\vec{B}. \tag{2.46}$$

The first equality is true for topological configurations, while the second one holds for configurations which include gauge fields because we chose it as a pure gauge. Stokes theorem was used in the third equality, where the line integral transforms into a surface one. The definition of the magnetic field was employed in the last equality. The derived quantity $\Phi = \frac{e}{2\pi} \int d\vec{s}\vec{B}$ is magnetic flux. The topological effects lead us to the physically measurable effect,

$$\Phi = \frac{2\pi N}{e}. \tag{2.47}$$

Namely, the magnetic flux of such configurations is quantized by the winding number. It is important to mention that such configurations are stable, as kinks, since they admit conserved topological charge.

Cosmic strings give us an interesting phenomenon, magnetic monopole [47, 52] type solutions in lower dimension.

2.2.4 Instantons in Yang-Mills

Pure Yang-Mills theories have the following Lagrangian

$$\mathcal{L} = -\frac{1}{4}GG, \tag{2.48}$$

where G is a field strength, corresponding to the gauge field A. Let us focus on gauge $SU(N)$ groups and for definiteness fix the space-time dimensions to four. In the conventional QCD we do not see naked colourful charges. Kugo and Ojima [28] argued, even further, that such charges could not be observed. The confinement paradigm is widely accepted. Hence, we can safely consider the full colour group $SU(N)$ as a vacuum manifold of the theory. Even, if we forget about the above paradigm, solutions considered in this section are still valid and rigorous. So, let us discuss topological properties of $SU(N)$. Following [48], we see that

$$\pi_3(SU(N_c)) = Z, \tag{2.49}$$

where π_3 is a third homotopy group. It counts winding numbers of the maps from S^3 to S^3, as π_1 does for a usual circle S^1.

Similarly to the previous cases, we should map the internal group with the space-time group at the infinity. Unfortunately, in this case we cannot do this procedure, since the Lorenz group $SO(1,3)$ is a non-compact one. It has only locally $SU(2) \times SU(2)$ algebra. Nevertheless, we can perform Wick rotation and consider only space coordinates. The symmetry group $SO(4)$ is doubly covered by $SU(2) \times SU(2)$, and we can have topologically non-trivial solutions in 4-dimensional Euclidean YM theories. The solutions will be the same for 5-dimensional YM theories, since their spatial part is a 4-dimensional euclidean YM-theory. Since we want to give a physical interpretation to the solutions, we use Wick rotation during the discussion. We will switch from 4-dimensional Euclidean to Minkowski space, back and forth, without mentioning it explicitly.

The Euclidean version of the action has the following form

$$S_E = \int d^4x \frac{1}{4}GG, \tag{2.50}$$

which can be written in the equivalent form

$$S_E = \pm \int d^4x \frac{1}{4}(G\tilde{G}) + \int d^4x \frac{1}{8}(G \mp \tilde{G})^2, \tag{2.51}$$

where the first term is topological ((1.59), (1.60), see e.g., [35, 39]). It can be rewritten as

$$\frac{g^2}{32\pi^2} \int d^4x G\tilde{G} = \int d^4x \partial_\mu K_\mu, \tag{2.52}$$

where K is called the Chern current (see e.g., [35]) and has the following form

$$K =^* \frac{g^2}{32\pi^2} tr\left(A \wedge dA + \frac{2g}{3} A \wedge A \wedge A\right), \tag{2.53}$$

where $*$ is the Hodge dual (1.77). In $1+3$ dimensions, the above expression can be decomposed in the following way

$$\frac{g^2}{32\pi^2}\int d^4x G\tilde{G} = \int d^4x \partial_\mu K_\mu = \int dt \int d^3x \left(\partial_0 K_0 + \partial_j K_j\right). \tag{2.54}$$

Here the last term vanishes at infinity and $Q_C = \int d^3x K_0$ is the Chern charge. So, we get

$$\frac{g^2}{32\pi^2}\int d^4x G\tilde{G} = Q_C(t=\infty) - Q_C(t=-\infty), \tag{2.55}$$

meaning that the topological term is simply change of the Chern charge during the evolution. Meantime, the quantity

$$\int \mathrm{tr}(dU \wedge dU \wedge dU), \tag{2.56}$$

counts π_3 winding of $U \in SU(N)$ up to a coefficient. If a gauge-field is almost a pure-gauge configuration (like an Abelian gauge field in cosmic strings (2.45)), then the Chern charge just counts the winding number of the gauge configuration due to (2.53). Hence, (2.55) shows change of the winding during the evolution. We can also interpret it in 5-dimensions, showing winding carried by a 4-dimensional monopole. As we have seen, the above solution differs from the previous ones. This solution only uses a part of the group, still giving a full topological number. This happens due to the factorization, since the rotations $O(4) \sim SU(2) \times SU(2)/Z_2$, and we map only one of the $SU(2)$ to the vacuum manifold.

So far, we only discussed the possibility of topologically non-trivial solutions. We also established their behaviors at the spatial infinities. To find them, we should minimize (2.51). Since the first part of the expression is a topological number, we can minimize only the second part, leaving us with

$$G = \pm\tilde{G}. \tag{2.57}$$

Solutions of the above equation, with a unit winding are celebrated BPST instantons [53] (for a review see e.g., [35, 39, 54]). The solution of the above equation with the plus sign is

$$A^a_\mu = \frac{2}{g}\eta_{a\mu\nu}\frac{(x-x_0)_\nu}{(x-x_0)^2+\rho^2}, \tag{2.58}$$

where η is called 't Hooft symbol [55] having the form

$$\eta_{a\mu\nu} = \begin{cases} \epsilon_{a\mu\nu} & \mu,\nu = 1,2,3 \\ -\delta_{a\nu} & \mu = 4 \\ \delta_{a\mu} & \nu = 4 \\ 0 & \mu=\nu=4 \end{cases}. \tag{2.59}$$

The solution at the temporal infinity is pure gauge, as it was intended. It depends on 5 parameters (x_0 and ρ) and carries the positive winding. To solve the equation (2.57) with the negative sign, we just need to switch the 't Hooft symbol with an anti-'t Hooft one (see e.g., [35])

$$\bar{\eta}_{a\mu\nu} = \begin{cases} \epsilon_{a\mu\nu} & \mu,\nu = 1,2,3 \\ \delta_{a\nu} & \mu = 4 \\ -\delta_{a\mu} & \nu = 4 \\ 0 & \mu = \nu = 4 \end{cases} . \tag{2.60}$$

The configuration carries a negative winding. Therefore, for both configurations, the action is

$$S_E = \frac{8\pi^2}{g^2}. \tag{2.61}$$

In general, we can have (and we do) multi-instanton configurations (see e.g., [35]). The action for these cases has again the simple form,

$$S_E = \frac{8\pi^2}{g^2}|k|, \tag{2.62}$$

where k is a π_3 winding number.

Similarly to the case of the Kink, zero modes also play an important role in the case of instanton. The unit winding solution in $SU(2)$ has in total 8 zero modes. Five of them are easy to read since objects correspond to translation and dilatation moduli, and three of them correspond to combined rotation in colour and Lorenz spaces, leaving 't Hooft symbols unaltered. In larger groups, with arbitrary winding k, counting of zero modes is different. Following [35, 56], we can count the number of independent parameters n needed for writing an instanton with a fixed winding number. The general formula has the form

$$n = \begin{cases} 4Nk - (N^2-1), & \frac{N}{2} \le k \\ 4k^2+1, & \frac{N}{2} > k \end{cases} . \tag{2.63}$$

In large-N, obviously, they coincide to the number of zero modes.

If we couple massless fermions with the gauge field, they will add extra zero modes. Adding one fundamental Dirac fermion changes the YM Lagrangian in the familiar form

$$\mathcal{L} = i\bar{\psi}\not{D}\psi - \frac{1}{4}G^2, \tag{2.64}$$

If we consider instantonic background, we can deduce that the fermion should have a zero mode. Using the index theorem [34] gives us (1.60), and since the term $G\tilde{G}$ is not zero,

$$\Delta(n_L - n_R) = 2\Delta Q_C. \tag{2.65}$$

So, during the evolution, chirality should flip (see e.g., [39]). However, we do not discuss this phenomenon here. We see from the equation above the existence of the fermionic zero modes.

To find an exact form, let us employ the following notation,

$$\sigma_\nu^\pm(\vec{\sigma}, \mp i), \tag{2.66}$$

where σ_νs are Euclidean Pauli matrices and $\vec{\sigma}$s are usual Pauli matrices. Then we can write down the Dirac equation for eigenvalues and eigenfunctions,

$$-i\gamma_\mu D_\mu u_n(x) = \lambda_n u_n(x), \tag{2.67}$$

where λs are eigenvalues and us are eigenfunctions. The function u_0 corresponds to the zero mode. As it is known, a Dirac spinor can be decomposed into two chiral ones. This decomposition in Euclidean space-time has the form

$$u_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \xi_L + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \xi_R, \tag{2.68}$$

where ξs are chiral left and right-handed spinors, respectively. Using them and Pauli matrices σs of the Lorenz group, the equations of motions can be written in the following way:

$$\begin{aligned} \sigma_\mu^+ D_\mu \xi_L &= 0 \\ \sigma_\mu^- D_\mu \xi_R &= 0. \end{aligned} \tag{2.69}$$

It is proved rigorously that the solutions of the equations above are the same as for the equations(see e.g., [39] and the references therein)

$$\begin{aligned} D^2\xi_L &= 0, \\ D^2\xi_R &= 4\vec{\sigma}\vec{\tau}\frac{\rho^2}{[(x-x_0)^2+\rho^2]^2}\xi_R, \end{aligned} \tag{2.70}$$

where τs are Pauli matrices in $SU(2)$ while σs are the ones in the Lorenz Group. The quantities x_0 and ρ are instanton collective coordinates. We see that left moving fermions in the instantonic background do not have any zero modes. A solution of the equation is trivial because $-D^2$ is a positive operator. However, there can be a zero mode for a right-handed spinor. By direct checking, we can see that the above equation is satisfied only if

$$\xi_R^{\alpha j} \propto \epsilon^{\alpha j}, \tag{2.71}$$

where α and j are Pauli matrices' indices of colour and space-time groups, respectively. The object ϵ is just 2-dimensional. The complete solution has the form

$$u_0^{\alpha j} = \frac{1}{\sqrt{2}\pi}\frac{\rho}{(x^2+\rho^2)^{\frac{3}{2}}}\begin{pmatrix} 0 \\ 1 \end{pmatrix}\epsilon^{\alpha j}, \tag{2.72}$$

normalized to unity. We see that each right-handed colour-fundamental fermion has exactly one zero mode in the instantonic background. The story is the same with the left-handed fermions in case of anti-instanton background.

Instantons and their zero modes play an important role in QCD. During the next chapters we will have the discussion about it.

2.3 Semi-classical/WKB approximation

Semi-Classical methods tend to be very useful, as we saw in Gross-Neveu example, where the full computation was only one loop. Even though we cannot do fully re-summation, small quantum corrections to the fully classical dynamics appear to be useful (e.g., one loop effective potential [41]). If we combine this method with large-N, our results are almost always exact. In this section, we will consider WKB methods in two different scenarios. We will compute tunneling amplitudes in different models, and we also compute particle spectra in certain QFT.

2.3.1 Instantons in Quantum mechanics

In this section, we will review, how euclidean solutions can describe tunneling processes. We mostly follow the book [35], same computations are also present in the books [57, 58].

Let us consider quantum mechanical one degree of freedom, or in other words, $1+0$ dimensional QFT. The Hamiltonian has the following form,

$$H = \frac{1}{2}p^2 + V(q), \tag{2.73}$$

where p and q are canonical variables. The potential has the following form,

$$V(q) = \frac{\lambda}{4}\left(q^2 - v^2\right)^2, \tag{2.74}$$

symmetry is clearly broken and $q = \pm v$. Naively, we would say that one can work in any of the two vacua, but in Quantum mechanics we have a unique ground state [59]. We can still start our pertubative approach and see if the tunneling process can restore the true vacuum. The tunneling process can be computed via the path integral,

$$\langle v| e^{-H\tau} |\pm v\rangle = \int \mathcal{D}q e^{-S_E}, \tag{2.75}$$

where we employ euclidean time, since our interest concerns on the tunneling process. There S_E is a Euclidean action with proper boundary conditions, and $\pm v$s are vev values for the field q. Moreover, we can write the above amplitude as,

$$\langle v| e^{-H\tau} |\pm v\rangle = \sum_n \langle v|n\rangle \langle n|\pm v\rangle e^{-E_n\tau}, \tag{2.76}$$

where ns are energy eigenvalues. In infinite time we get the following expression

$$\int \mathcal{D}q(t) e^{-S_E} = \langle v|n\rangle \langle n|\pm v\rangle e^{-E\tau}, \tag{2.77}$$

where $E_\pm$ are the lowest possible energy levels (correspond to symmetric and antisymmetric choice of the states. Here, energy and vacua $\pm v$ signs are not related). From here we can restore the usual result from quantum mechanical textbooks. Let us take a look at the expression

$$\int \mathcal{D}q(t)e^{-(S_E-E_{\pm}\tau)} = \langle v|\pm\rangle \langle\pm|\pm v\rangle . \tag{2.78}$$

On the left side we see the euclidean version of the following identity

$$L + H = p\dot{q}. \tag{2.79}$$

So, the transition is

$$\int \mathcal{D}q(t)e^{-\int pdq} = \langle v|\pm\rangle \langle\pm|\pm v\rangle , \tag{2.80}$$

If we assume extreme trajectories, then we get

$$\langle v|\pm\rangle \langle\pm|\pm v\rangle = e^{-\int pdq}, \tag{2.81}$$

evaluated on-shell euclidean trajectory. The expression usually appears in the quantum mechanical textbooks, but is derived through a different method.

Now, we want to compute the transition probability. To carry it out we should compute the action corresponding to the Hamiltonian (2.73)

$$S_E = \int d\tau \left(\frac{1}{2}\dot{q}^2 + \frac{\lambda}{4}\left(q^2-v^2\right)^2\right), \tag{2.82}$$

It is not hard to recognize (2.15), together with the (2.18). Therefore, the action is

$$S_E = \sqrt{\frac{8\lambda v^6}{9}}, \tag{2.83}$$

and we get simple transition amplitudes. Now we can compute leading corrections, which will multiply amplitude by the following quantity

$$\det\left(\frac{d^2}{d\tau^2} - V''(q)\right)^{1/2}. \tag{2.84}$$

The quantity should be evaluated on top of the solution (2.18), but here the problem arises. The solution has a zero mode/collective coordinate. We redefine the determinant,

$$\det{}'(\frac{d^2}{d\tau^2} - V''(q))^{1/2}, \tag{2.85}$$

where $'$ means all other modes except the zero mode. Despite removing the zero mode, we still have it as a part of the path integral integration. So, the integral over the zero mode has the form

$$\frac{1}{\sqrt{2\pi}}dc_0. \tag{2.86}$$

Since ψ_0 is the zero mode, it is a derivative of q and it should be normalized to unity $\psi_0 = \frac{1}{\sqrt{S_E}}\frac{dq}{d\tau}$. A small change in coordinate is given by

$$\psi_0 dc_0 = \frac{1}{\sqrt{2\pi}}\frac{dq}{d\tau}d\tau, \tag{2.87}$$

From here we get

$$dc_0 = \sqrt{\frac{S_E}{2\pi}}d\tau, \tag{2.88}$$

giving us the result

$$amp \propto \sqrt{S_E}e^{-S_E} \times \tau(\overset{'}{\det})^{1/2}. \tag{2.89}$$

Here $\det'$ has a dimension of energy square. We computed the contribution of only one instanton, but it is important to understand what happens in case of the diluted gas [55]. Now we should compute instanton-anti-instanton separated in time. The action in the exponent will just have a factor n. The only contribution which should be treated carefully is an integration over moduli. Since one instanton should end before the other starts, we get the following cascade of integrals (in the case of 2 instantons)

$$\int_{\tau 1}^{\tau} d\tau_1 \int_0^{\tau_1} d\tau_2. \tag{2.90}$$

Such integrals simply give us $\frac{1}{n!}\tau^n$. We should take the sum of all contributions and finally, our amplitude is

$$amp(v, -v) = (\overset{'}{\det})^{1/2}\sinh\left(e^{-S_E}\sqrt{\frac{S_E}{2\pi}}(\overset{'}{\det})^{1/2}\tau\right). \tag{2.91}$$

Similar computation gives us a very similar expression for $\langle v| e^{-H\tau} |v\rangle$ transition amplitude [35], where we should just substitute the function sinh by the function cosh. Afterwards, the level split between symmetric and antisymmetric states can be computed

$$E = E_- - E_+ = e^{-S_E}\sqrt{\frac{S_E}{2\pi}}(\overset{'}{\det})^{1/2}, \tag{2.92}$$

The exponential factor of the result is a standard WKB computation, the corrections fix exactly only the prefactor. This machinery for quantum mechanics looks really like an overkill. Nevertheless, it shows two advantages, giving us a recipe for computation of transition amplitudes in QFT and showing high dimensional topological objects kinks/monopols serving the role of the instantons in lower dimensional theories.

2.3.2 YM Instantons and YM vacua

In the previous chapter we saw how instantons describe the tunneling process in a simple theory, they connect different “vacua”. Hence, they carry the topological charges, since the “vacua” are in different regions of π_0. The same is true in the case of YM instantons, they carry π_3 winding number, and they have non-zero action (2.61). To find what types of tunnelings they describe, we should take a look at the Chern current (2.53). The object is clearly not gauge invariant, hence, the instantons must describe physics of non-gauge invariant states. Naively, one might immediately stop thinking about such processes, since we know that physical objects must be gauge (or even more BRST) invariant (see Section 1.2). However, we have not talked about the vacuum of YM yet. So, the instantons should be the key ingredients to build one. It appears that, they indeed are, they give a nice framework to study the vacuum structure of YM. The framework was introduced, and many aspects were discussed in [60, 61]. Apparently, YM has a rich vacuum structure. We will talk about it in this section (for a review, see, e.g., [35, 39]).

Let us again consider a pure YM theory, which is described by the Lagrangian,

$$\mathcal{L} = -\frac{1}{4}G^2, \tag{2.93}$$

where G is a $SU(N_c)$ non-Abelian field strength tensor. In perturbative computations, usually we assume that $A = 0$ in the vaccum. However, this assumption is too constraining, since we have π_3 winding and instantons. For instance, (2.58) gives us a pure gauge configuration at temporal infinities. So, instead of assuming that $A = 0$, we should consider the pure gauge configurations. They give the same zero field strength, but they can carry non-trivial winding. For simplicity, let us assume that $A_0 = 0$ (temporal gauge) for all As. Therefore, the pure gauge configurations have the following form

$$A_j = \frac{i}{g}U^{-1}\partial_j U, \tag{2.94}$$

with the boundary condition $U(\infty) = 1$. For such configurations obviously the field strength is zero. As we saw, this configuration can carry a non-trivial winding (2.56). So, the Chern charge for such configurations is not zero,

$$\int d^3x K^0 = N[C]. \tag{2.95}$$

Let us pick a state with a definite winding as our “vacuum”. The question immediately arises regarding the guarantee of the BRST invariance of the vacuum. One can argue, that the Chern charge is not invariant if we consider singular gauge transformations only, which change a topological number. So maybe we can take regular gauge transformations, and any “vacua” with arbitrary winding will do a job. Apparently it is not doable, since we have instantons and as soon as we fix the winding, the physical system will jump to another winding. Hence, we do not have a gauge invariant vacuum yet. Meantime, transition from one to the other “vacua” is measured by gauge invariant quantity (2.55). So, it gives us a clue regarding the construction of gauge invariant vacua. Following the discussion [35], we consider an operator T which increases winding by unity.

$$|n\rangle \to |n+1\rangle \ . \tag{2.96}$$

Assuming that the operator commutes with the Hamiltonian (for instance, divergence of Chern current) of the theory, i.e. $[T, H] = 0$. Hence, it is clear that the Hamiltonian and the operator share the eigenvectors. We can construct a real vacuum from superposition of the "vacua"

$$|\theta\rangle = \sum_{n\in Z} e^{-i\theta n} |n\rangle . \tag{2.97}$$

The $|\theta\rangle$ vacuum is an eigenstate of an operator T with an eigenvalue $e^{i\theta}$. Considering its definition, it is clear that the parameter θ is defined up to modulus 2π. This object has possibility to be a vacuum, since it is a superposition of energy zero states (at least classically). The question arises regarding the possibility of building true vacua satisfying super-selection rules.

Despite the fact, that we built a theta vacuum as an eigenstate of the Hamiltonian, let us compute its evolution and let us study the following correlator

$$\langle\theta'| e^{-iHt} |\theta\rangle = \sum_{m,n} e^{im\theta'-in\theta} \langle m| e^{-iHt} |n\rangle = \sum_{m,k} e^{im(\theta'-\theta)} e^{ik\theta} \langle k| e^{-iHt} |0\rangle . \tag{2.98}$$

In the first equality we used a definition of the θ vacuum, while the gauge invariance was applied to the second one. This means that a physical quantity can depend entirely on the difference of windings (since the difference is a gauge invariant quantity (2.55)). So, the final result is

$$\langle\theta'| e^{-iHt} |\theta\rangle = 2\pi\delta(\theta-\theta') \sum_{k} e^{ik\theta} \langle k| e^{-iHt} |0\rangle , \tag{2.99}$$

where we used a Fourier decomposition of a delta function with a compact argument. The above result tells us, that the θ-vacua have a super-selection property. Once we pick a θ, it cannot be changed by a physical process in the pure YM.

Now let us compare the above result with the path integral quantization. The amplitude is given by

$$\sum_{k} e^{ik\theta} \langle k| e^{-iHt} |0\rangle \propto \sum_{k} e^{ik\theta} \int DA_\mu^k e^{i\int d^4x\mathcal{L}}, \tag{2.100}$$

where we assumed that the quantum measure sums only configurations with a topological number k. If we look at the connection between the change in the winding number and gauge fields (2.55), we can bring an exponential factor in the Lagrangian by adding the following term

$$\mathcal{L}_\theta = \frac{\theta g^2}{32\pi^2} G\tilde{G}. \tag{2.101}$$

So, to manifest the θ-vacuum, we should add this extra term in the YM Lagrangian

$$\sum_{k} e^{ik\theta} \langle k| e^{-iHt} |0\rangle \propto \int DA_\mu e^{i\int d^4x(\mathcal{L}+\mathcal{L}_\theta)}. \tag{2.102}$$

The integration can be done over all configurations. As this extra term is a total derivative, it does not play any role in a perturbative expansion. Despite this, the θ-term is essential, it gives us observable physical phenomena, P and CP violations due to the presence of the ϵ tensor.

Nowadays, experiments do not observe above phenomena in the QCD sector. They give us a bound for the parameter theta [62]

$$\theta < 10^{-9}, \tag{2.103}$$

Usually, community talks about smallness of the angle. But, the θ-vacua is a super-selection puzzle, which may be connected to the consistency issue [63]. Usually, those issues are referred as "the Strong-CP problem". We will discuss it in the following chapter.

2.3.3 Bound states, a naive WKB approximation

Applying WKB approximation in a certain form (The "Old quantum mechanics") to find bound states in quantum systems was known before the birth of the actual quantum mechanics (see, e.g., [59]). In certain systems, the approximation can give the exact spectrum too. Such systems in quantum mechanics are well known, the hydrogen atom and harmonic oscillator (see, e.g., [59]). Using a path-integral formulation, the WKB analysis can be done in QFT [49, 64] (see, e.g., [43, 57]). In several examples, the spectrum can be extracted explicitly [44, 65].

Let us consider the following identity before discussing WKB approximation,

$$G(E) = tr\left(\frac{1}{H-E}\right) = \sum_k \left(\frac{1}{E_k - E}\right) = itr\int_0^\infty dTe^{i(E-H)T}dT, \tag{2.104}$$

where E_k are bound state poles. The Green function G has poles, when the energy is equal to the energy of the bound state. The last expression gives us the possibility to write the Green function ($i\epsilon$ prescription should be done in the Hamiltonian) in terms of the path integral. So, finally, we get the following equation

$$G(E) = \int \mathcal{D}\phi e^{i(S+\int EdT)} \tag{2.105}$$

In ϕ we mean all quantum fields. The above expression looks very similar to the one from the previous chapter, but in this case, we have a complex pre-factor. This gives us periodicity, and we can make transition amplitudes infinite (giving poles in the propagator), if exponential is a unity. This means, that the expression in the exponent should be an integer time 2π, giving us

$$S + ET = 2\pi N\,. \tag{2.106}$$

One can recognize naive WKB quantization ("the old quantization") rule here:

$$\int dT \int d^n x \sum \dot{\phi}\pi = 2\pi N, \tag{2.107}$$

where π's are canonical momenta, corresponding to ϕs. The above expression is valid, since

$$\sum \dot{\phi}\pi = \mathcal{L} + \mathcal{H}. \tag{2.108}$$

It appears, that this type of quantization requires solutions which are compact and periodic [49, 64].

From (2.106), differentiating with respect to T, taking into account that $E = -\frac{\partial S}{\partial T}$, considering N and E as function of T, we get

$$\frac{dN}{dE} = \frac{T}{2\pi} = \frac{1}{\omega}, \tag{2.109}$$

where ω is a characteristic frequency of the classical equations of motions. Since we consider periodic configurations, we can always find a characteristic frequency in terms of the energy and N. The solution of the above differential equation gives energy levels in terms of a quantum number N. We should normalize the ground state energy, for instance, $N = 0$ corresponds to E_0.

2.3.4 The spectrum of the Sine-Gordon

The spectrum of sine-Gordon was found in [65], which turns out to be exact (see, e.g., [43]). In this section, we will investigate it.

The sine-Gordon is defined by the following action

$$S = \int d^2x \left(\frac{1}{2}(\partial\phi)^2 + \frac{m^2}{\beta^2}(\cos(\beta\phi) - 1) \right), \tag{2.110}$$

where m, β are parameters of the theory and ϕ is a one field.

In the spectrum, the theory has topological Kinks (similar ones were considered in the section 2.2.1)

$$\phi = \frac{4}{\beta}\,\text{atan}\Big(e^{\pm(z-z_0)}\Big). \tag{2.111}$$

The Kink has the following mass $M_s = \frac{8m}{\beta^2}$. Its existence we can predict from the vacuum manifold

$$\mathcal{M} = Z/\{0\} = Z. \tag{2.112}$$

It is possible, in this model, to do quantization on top of the Kink and get an exact quantum mass of it [65] (for a review see e.g., [35, 43]). However, in this chapter we are interested in derivation of the spectrum of the theory. So, we skip this discussion here.

The model also exhibits so-called “breater” solutions. They have the form

$$\phi = \frac{4}{\beta}\,\text{atan}\left(\frac{\eta\sin(\omega t)}{\cosh(\eta\omega z)}\right), \tag{2.113}$$

where $\eta = \frac{\sqrt{m^2-\omega^2}}{\omega}$. At the negative infinity the solution has the following behavior

$$\phi \approx \frac{4}{\beta}\,\text{atan}\Big(e^{mx+\log(2\eta\sin(\omega t))}\Big), \tag{2.114}$$

which looks like a soliton (Kink), with moduli depend on time. We get a similar result if we consider positive infinity. Therefore, this solution is a well-separated Kink-anti-Kink system. They are "breating" during periodic motion. This solution is discussed in details in [35, 43]. The energy of this solution is given by the equation

$$E = \frac{16m}{\beta^2}\sqrt{1-\frac{\omega^2}{m^2}} = 2M_s\sqrt{1-\frac{\omega^2}{m^2}}, \tag{2.115}$$

For real frequencies, we should interpret the breather as a soliton-anti-soliton bound state, since the energy of the breater is less than the energy of the Kinks. We can also solve frequency as a function of the energy,

$$\omega = m\sqrt{1-\left(\frac{\beta^2 E}{16m^2}\right)^2}. \tag{2.116}$$

This equation is ready-made for WKB quantization. Using (2.109), we get

$$\frac{dE}{dN} = m\sqrt{1-\left(\frac{\beta^2 E}{16m}\right)^2}, \tag{2.117}$$

which gives the solution (assuming $N = 0$, $E = 0$)

$$E = 2M_s \sin\left(\frac{\beta^2}{16}N\right), \tag{2.118}$$

with $N < \frac{8\pi}{\beta^2}$. Such a simple treatment can give us masses of the bound states. Surprisingly, the full quantum spectrum has also a similar form

$$E = 2M_s \sin\left(\frac{\xi}{2}N\right), \tag{2.119}$$

where $\xi = \frac{\pi\beta^2}{8\pi-\beta^2}$ and $M_s = \frac{m}{\xi}$. The result for small β coincides with the naive WKB treatment.

2.3.5 The proper WKB

In previous chapters we considered naive WKB approximation, which is a good approximation for the systems like hydrogen atom, or Sine-Gordon model. In those systems, we have one family of classical solutions, and we can quantize on top of that. This is not true in general, and we should be a bit more careful. We did not derive the WKB method rigorously. In this section, we will give general ideas from the articles [49, 64], how the treatment should be modified.

In general, we should consider all saddle points and sum over them using a path integral. The particle can move through the same trajectory back and forth. This is similar to the multi-instanton solution considered in the section 2.3.1. We will not derive the result, since the technical details are beyond of the scope of manuscript. We will just give the results from [64].

Let us define the short action $W(E) = S(T(E)) + ET$ (the quantity which enters in (2.107)). In terms of the short action, we get the following Greens function (2.104) for the quantum mechanical system

$$G(E) \approx -iT(E)\frac{e^{iW}}{1+e^{iW}} \, . \tag{2.120}$$

From the above equation we can see,

$$W(E_n) = 2\pi \left(n + \frac{1}{2}\right), \tag{2.121}$$

This expression is more accurate, since it counts a vacuum energy in the harmonic oscillator. There are also some technical caveats not covered here.

2.3.6 Bound states in the Gross-Neveu model

As we presented in the chapter 2.1 the Gross-Neveu model is exact in large-N limit. It turns out that it is also possible to solve exactly (using WKB) its spectrum [44].

In this chapter we will not go through technical details, we will only cover important aspects. Let us first consider the Gross-Neveu model in the form (2.2). During the previous discussion, we only considered a part of its symmetry group, $U(N)$. Apparently, it has a larger symmetry, namely $SO(2N)$. Let us decompose a 2-dimensional Dirac spinor

$$\psi_j = \psi_j^1 + i\psi_j^2, \tag{2.122}$$

where, both ψ^1 and ψ^2 are real spinors and j numerates $SU(N)$ index. Then, the Lagrangian (2.2) can be written in the following equivalent form

$$\begin{aligned} \mathcal{L} = -\frac{1}{2\alpha}\sigma^2 + \\ + \sum_j \left(i\psi_j^1 \partial_t \psi_j^1 + i\psi_j^2 \partial_t \psi_j^2 + i\psi_j^1 \partial_z \sigma_3 \psi_j^1 + i\psi_j^2 \partial_z \sigma_3 \psi_j^2 + \sigma\psi_j^1 \sigma_2 \psi_j^1 + \sigma\psi_j^2 \sigma_2 \psi_j^2 \right), \end{aligned} \tag{2.123}$$

with the choice of $\gamma_0 = \sigma_2$, $\gamma_1 = i\sigma_1$. In the above Lagrangian, all the mixed terms are dropped out. Usually, this basis is called “Majorana basis”. The Lagrangian, obviously, has a larger $SO(2N)$ symmetry, the spinors transform under the vector representation of the $SO(2N)$ group. The σ field is a scalar under the symmetry. This has a big impact on the spectrum of the theory.

From the equation (2.5) we expect kink solutions, due to the vacuum degeneracy (2.9). So, as in the case of sine-Gordon, we expect to have a kink-anti-kink system, which gives us bound states. Since the scalar field σ does not carry $SO(2N)$ quantum numbers, we expect group representations. Since elementary building blocks are fermions, the representations can be totally-anti-symmetric objects only. Such representations indeed exist in the group and its size can be calculated as

$$n_{st} = \frac{2N!}{n!(2N-n)!}, \tag{2.124}$$

where n labels representation (number of the indices).

It turns out, that the model indeed has bound states, and they form totally antisymmetric multiplets of the group $SO(2N)$. Their degeneracy is given by (2.124), as it was supposed. Their masses depend on the representation, n

$$M_n \; = m_f \frac{2N}{\pi} \sin\left(\frac{n}{N}\frac{\pi}{2}\right), \tag{2.125}$$

which for $n = 1$ reproduces the mass of the elementary fermion.

We will not reproduce these results here. The derivation [44] is very technical and not too relevant to our discussion.

2.4 Saturons

In the series of papers [8–10] (for the summary, see [66]) universal unitarity bound was established on quantum field theoretic objects. Unitarity is the most important concept in a quantum field theory. Self-consistent QFTs always respect unitatarity. In the literature, there was a misconception between *unitarity* and *pertubative unitarity.* The first one cannot be violated in self-consistent QFTs, while the violation of the latter leads to the change of the regime. For instance, in QCD, the second one is violated. It means that instead of quarks and gluons, mesons and baryons represent the S matrix states. During the last century, we saw a so-called Wilsonian UV-completion of old physics many times, which preserved perturbative unitarity at high energies. For instance, the famous $V - A$ theory is UV-completed by the standard model, integrating in new degrees of freedom at high energy. That is not always the case or the necessity. For instance, it is conjectured in [7, 67–69] that the Gravity maintains unitarity not by integrating new degrees of freedom in (which can maintain perturbative unitarity at high energies) but via the classicalization, where composite objects (Black Holes) keep theory unitary at high energies. This looks like a low-energy QCD, where unitarity is restored by baryons. One of the aim of the above papers [8–10] was to check the existence of similar effects in UV-complete theories. One may naively expect pertubative unitarity being always present in no regime changing and renormalizable QFTs. This is not true, since the pertubative series are asymptotic ones. One can always find a large number of quanta, for which pertubative unitarity is broken. The opposite is argued in the above papers and emergence of bound states is demonstrated, as in QCD or gravity, which unitarizes the theory. These objects emerge at the optimal truncation point, and they have maximal amplitudes. Nevertheless, it is hard to manufacture such states, since they have a tiny window in the phase space. Above-mentioned objects imply that the states of the S-matrix are constrained. Their degeneracies are restricted by unitarity. The papers conclude, that unitarity is preserved at the optimal truncation point, despite the signal regarding its violation. The theory simply creates a bound state.

2.4.1 Entropy bound

Setting the bound on states' degeneracy by unitarity was discussed in [8]. In other words, in a self-consistent QFT, S matrix states can not be degenerated more than a certain threshold. This bound can be computed from the parameters of the theory. To see the importance of the consideration, we will talk about highly degenerated states. Let us imagine that we have a scattering process, namely $2 \to n$. The states with the same energy-momentum of the n quanta can be different, i.e., those quanta can have different quantum numbers. Hence, the outcome of this process is highly degenerated. Such scenario is possible in QFT models with the property of having different species of similar quanta. So, this type of consideration is almost impossible, for instance, in usual QED. A typical process is $2 \to 2$ scattering, e.g., positron-electron annihilation in 2 photons (two polarization is far too small to talk about the degeneracy). Nevertheless, talking about degenerate states is important already in the Standard model (e.g., QCD sector). In QCD, degeneracy is present due to the quark flavour. For a moment, let us consider exactly the massless limit of light quarks, which means that stable Baryons have the same masses. Looking at lepton-anti-lepton annihilation process we can deduce, that light baryons are born with the same rate (let us ignore the difference in electric charge among the quarks), we get an extra degeneracy for a cross-section due to their multiplicity. One can argue that QCD itself at low energy is hard enough, and these considerations might not be rigorous. It would be an important point if similar behaviors were not observed in non-confining theories. Numerous models are discussed in [8–10], which do not confine, and the amplitudes are still degenerate as the QCD ones.

To demonstrate the importance of degeneracy, we consider the following model, a real scalar field coupled with N-flavoured Dirac fermions via Yukawa coupling. Let us assume that fermions are lighter than the scalar and there is no SSB in the scalar sector. In this setup, a decay of scalar in two quanta is enhanced, comparatively to the setup where just one fermion is provided. The enhancement comes from the degeneracy of fermions, due to many flavours. In other words, scalar can decay in N different ways. The decay product is identical, up to the flavour content. This example cannot be used for next discussions, despite the exact demonstration of amplitude enhancement. The process is still very quantum, since we can determine flavours of the final products with a small effort.

Following [8], we should quantify the degeneracy of the final state. Let us imagine that we have N bosonic sectors, which are exactly degenerated e.g., flavours, and we consider $2 \to n$ process. Then the final product can be distributed among the sectors in the following way,

$$n = \sum_{j=1}^{N} n_j, \tag{2.126}$$

where n_j is a number of quanta in the j'th sector. From the above equation we can immediately see degeneracy n_{st} of such microstates

$$n_{st} = \binom{n+N}{N}. \tag{2.127}$$

Assuming that for each fixed flavour content cross-sections are the same, we get the following enhancement for the scattering process

$$\sigma = \sum_{microstate} \sigma_{2\to n} = \sigma_{2\to n} n_{st}, \tag{2.128}$$

where σ and $\sigma_{2\to n}$ are full and with a fixed flavour content cross-sections, respectively.

One may think, naively, that the above cross-section diverges rapidly for a large number of quanta. However, this in not the case, since the $\sigma_{2\to n}$ itself depends on n. In various sources [70–75], which we also confirmed for several models in our formalism [5], the cross-section scales in the following way

$$\sigma_{2\to n} \propto c_n n! \alpha^n, \tag{2.129}$$

where α is $2 \to 2$ processes effective coupling, and c_n dependence on n is weak and not relevant. The above equation should be used carefully. It does not always hold, since perturbation series are asymptotic ones and the above consideration is valid up to optimal truncation. During the rest of the chapter we assume that n is exactly at the optimal truncation,

$$\lambda_c \equiv \alpha n \sim 1. \tag{2.130}$$

Additionally, we assume that $\alpha \to 0$, hence n is large. The Stirling approximation gives the following scaling of the cross-section (at the truncation point)

$$\sigma_{2\to n} \propto e^{-n} = e^{-\frac{1}{\alpha}}. \tag{2.131}$$

This makes transparent that the expectation of full cross-section (2.128) being divergent with large number of quanta is simply wrong. Now we should remember, that not just scaling of cross-section, but also the degeneracy of final states n_{st} is important. Since the object is degenerated, it is natural to define entropy based on it

$$S = \log n_{st}. \tag{2.132}$$

It gives us possibility to quantify the degeneracy of the states in an easier way. The full cross-section has the following form

$$\sigma \propto e^{-\frac{1}{\alpha}+S}. \tag{2.133}$$

To preserve unitarity of the theory, entropy should satisfy the inequality

$$S \leq \frac{1}{\alpha}. \tag{2.134}$$

It shows us, that maximal degeneracy of states is limited by unitarity. The cross-section should not grow infinitely, which is a reason of the inequality. Above restriction should not be understood as an upper bound for a maximal particle number, there is nothing wrong to have almost infinite number of electrons in QED (where usually only a pair of particles scatter). It contradicts considerations, where all particles scatter simultaneously. So, basically, this restriction sets the maximal possible degeneracy of asymptotic states (for instance, number of quarks in QCD states). This limit also puts restriction on the number of quanta taking place in the S matrix processes. An exact realization of the above restriction in a real example will be demonstrated in next chapters.

At the optimal truncation, the entropy saturates unitarity bound, which can be seen from its definition (2.132) and the collective coupling λ_c (2.130). Existence of composite objects, with the same entropy, was conjectured in [8–10]. Therefore, many quanta scattering at the optimal truncation point gives Saturons. This explains the reason for not being able to go beyond the truncation point. After that point, the theory changes degrees of freedoms and elementary quanta are no more good descriptions.

Despite the Saturons being maximally degenerated states, we can still scatter them, for instance, two Saturons. In such scattering the S matrix will exchange at most half of the "elementary" quanta between the Saturons*. This implies that Saturons hardly scatter†.

In $3+1$ dimensions, entropy of the Saturon can be written in an Area-Law form [8–10]

$$S = f^2 Area, \tag{2.135}$$

where f serves as a "Goldstone decay constant". Here, we call all types of gapless modes "Goldstones". Large-N analysis tells us how the Goldstone decay constant scales with N [76]. If the system has size R, then

$$f \sim \frac{\sqrt{n}}{R}, \tag{2.136}$$

considering the definition of area-Law [77, 78] entropy, we get

$$S = (Rf)^2 = Area/G = \frac{1}{\alpha}\,. \tag{2.137}$$

at the optimal truncation point. The entropy also can be written, in the form of 2-dimensional area, where G is a "Newtonian" constant. The above equations are usually considered in the context of gravity. In this chapter, we have not talked specifically about the gravity. So, the black holes are just particular cases of the Saturons.

Now let us consider all quanta inside the Saturon with characteristic momenta $q \sim \frac{1}{R}$, then the energy of the Saturon is

$$E \sim \frac{n}{R}, \tag{2.138}$$

*This is not entirely true if Saturons carry the same degeneracy, e.g., creation of Saturon-anti-Saturon pair

†Highly degenerate objects hardly scatter in general, e.g., mesons in large-N QCD [76]

So, the entropy of this system is given by

$$S = ER = \frac{1}{\alpha}. \tag{2.139}$$

Above expression coincides with the Bekenstein entropy [79].

2.4.2 Page's time

It was argued [8] that Saturons and black holes share many key features. We already discussed some of them in the previous section. Surprisingly, there is one more important property, a minimal time, after which we can decode information from the object. The time is called Page's time, and it is given by the expression

$$t_{min} = \frac{R}{\alpha} = SR\,. \tag{2.140}$$

Information cannot be obtained at any shorter timescale. For $\alpha = \alpha_{gr}$ this expression reproduces the Page's time [80] for a black hole.

The above-mentioned similarities between black holes and Saturons were demonstrated for huge number of examples [8–10, 81]. All the examples include high amount of occupation of quanta. So, this picture supports the idea, that black hole consists of highly occupied gravitons [7]. It was also argued [82] that YM has a candidate [83] with the above properties. We should stress that, correspondence between black holes and other Saturons does not mean the same between gravity and other theories. From here the existence of universal phenomena can be deduced, shared among them. For instance, in the article [84] a phenomenon was generalized from the [81] Saturon to the black holes. This universality is important in two ways, to extract some universal phenomena from the simpler models and to have guidance from gravity and QCD on what type of universal proprieties we should seek in simple models.

We see that universal phenomena, saturation, gives us certain similarities between different theories. Those objects carry maximal entropy permitted by unitarity. They are realized as bound states, simultaneously saturate Area-Law entropy bound and pop up at the Optimal truncation point. It is very hard to read information from them. Working with the Saturons in general is not an easy task. They usually involve high complexity, in most of the models we can only find them at the qualitative level. Nevertheless, there is an exception. During the next sections we will demonstrate one example of the Saturon, we will do an exact computation and show all the properties of the Saturon on an integrable model [2].

2.5 A perfect Saturon in the Gross-Neveu model

The Gross-Neveu Model is an ideal laboratory to study Saturons. We found that it has a perfect Saturon in the spectrum. The model has nothing to do with the 4-dimensional general relativity. So, it also demonstrates the point, that saturation is a universal phenomenon. In this section, we will go through our paper [2] empathizing analogies between the Gross-Neveu Saturon and black holes.

2.5.1 The entropy of bound states

Bound states in the Gross-Neveu model are degenerate (2.124). The quanta with different flavour structures have equal masses, we call them multiplets. In is natural to define microstate entropy (2.132) for each multiplet similarly (2.127) to the previous chapter,

$$S = \ln(n_{st})\,, \tag{2.141}$$

The entropy reaches its maximum for $n = N$. Due to large N, we can use Stirling approximation, which gives,

$$S = 2N\ln(2)\,. \tag{2.142}$$

This expression is valid up to relative corrections of order $\ln(N)/N$. Since we work in a large-N limit, they are negligible. The entropy (2.142) is the limiting entropy of the bound state in Gross-Neveu model. From (2.125), the mass of this bound state (a candidate of saturon, which we simply call during the discussion saturon) is given by,

$$M_N\ =\ \frac{2N}{\pi}m_f\,. \tag{2.143}$$

Thus, both the mass and the entropy of the bound state scale as N in units of the fermion mass m_f. We also note that the size of the bound state is set by the Compton wavelength of the fermion,

$$R \sim \frac{1}{m_f}\,. \tag{2.144}$$

From the definition of the 't Hooft coupling (2.3) we see that

$$S \sim \frac{1}{\alpha}, \tag{2.145}$$

where α is a $2 \to 2$ coupling. We see similarities here, if we change α to the α_{gr}, the analogy is exact [8–10].

We can define an area law entropy in the Gross-Neveu model. To define an area in a low dimensional theory, let us first look at the expression (2.137). The expression is well-known in the case of the gravity, where we have Newton constant G_{gr}. In the case of black-holes we get

$$S_{BH} \sim \frac{\mathcal{A}rea}{G_{gr}} \sim \frac{1}{\alpha_{gr}}\,. \tag{2.146}$$

Now, we should find an analog of the G_{gr} for a Gross-Neveu Saturon. In terms (2.134) of G_{Gold}, the above entropy is given too (2.137). In each case, the coupling G_{Gold} is unambiguously defined. The reason is that any object, carrying entropy, breaks spontaneously only a part of the Poincare symmetry.

In particular, for a generic saturated state of N quanta, of the size R, the coupling G_{Gold} is given by (a general form of the equation (2.137)),

$$G_{Gold} \sim \frac{R^{d-2}}{N} \sim \frac{R^{d-2}}{S} \,. \tag{2.147}$$

Correspondingly, the form of the entropy is

$$S \sim \frac{\mathcal{A}rea}{G_{Gold}} \sim \frac{1}{\alpha} \,, \tag{2.148}$$

This is a general statement, which is also valid in the Gross-Neveu. There is only one peculiarity, since the model is defined in two dimensions the coupling is dimensionless $G_{Gold} = 1/\alpha = N$. The above equation is still valid, since we can define area in this case. For four dimensional QFTs area is a 2-dimensional object, which traps 3-spatial-dimensional compact objects. In the case of $d = 2$, such objects are just two dots. So, the "area" in this case is simply a dimensionless number 2, but we will forget about order one factors and just write $\mathcal{A}rea \sim 1$.

We see that all the equations which we usually use in gravity can be generalized and used in non-gravitational theories too. We note here, that (2.148) and (2.145) coincide with each-other. It appears that the connection is deeper, the Gross-Neveu Saturon also saturates the Bekenstein entropy bound. The Bekenstein bound says that the entropy of an object of mass M_N and radius R is bounded by (2.139) (including order one factors [79]),

$$S = 2\pi M_N R \,, \tag{2.149}$$

considering the size (2.144) and the mass (2.143), this expression takes the form

$$S \sim N \,, \tag{2.150}$$

which coincides with (2.142). So, we get 3 different entropies (Micro, Area Law, Bekenstein) which are simultaneously saturated and coincident with each-other.

We may ask the reasons for these bounds being maintained. Let us compare (2.142) and (2.145). It is clear from here that violation of the entropy bounds require growth of the coupling faster, than the number of the quanta,

$$\alpha >> \frac{1}{N}, \tag{2.151}$$

which means that the 't Hooft coupling should be more than one at the scale of bound states. This is impossible, since at the bound states scale the 't Hooft coupling is frozen, like we discussed in the chapter (2.1). $\lambda \sim 1$ also means the change of the regime in the theory. For instance, in QCD, we go from weakly interacting gluons and quarks to baryons and glueballs. The theory eventually changes the degrees of freedom at the edge of the strong coupling. So above results are protected by QFTs themselves, we cannot violate the bound on the entropies. The only surprise given by the Gross-Neveu model is a multiplet of states which saturate the bound. So, it has a Saturon, an analog object to the black holes in the gravity.

2.5.2 Saturation of scattering amplitudes

Following the general criteria of [8], we shall now establish the correlation between the entropy bound and unitarity in Gross-Neveu model.

The connection goes through the 't Hooft coupling λ. As (2.145) indicates, the saturation of the entropy bound (2.134) coincides with the regime where λ becomes of order one. Of course, λ has to be understood as the running coupling evaluated at the scale m_f. This scale sets the size of the bound state according to (2.144).

Similarly to the discussion in the previous chapter, we cannot increase λ, without jeopardizing the properties of the theory. We cannot push the entropy (2.142) above (2.134) by increasing λ, since the running will change the scale of the fermion, and we will end up with the same above bound.

The point $\lambda \sim 1$ is a critical point, where the bound states play an important role. For instance, let us look at the one-loop diagram (2.1). The loop is controlled by λ, and when it hits the order one, loop expansion breaks down. The series should be re-summed, which means the change of the regime in the new degrees of freedom.

We can also see a similar breakdown from the corrections to the fermion self-energy (using the Lagrangian (2.1)). For example, the two-loop contribution of the following type (fig 2.2), which is the order of λ^2. Each extra bubble brings an additional power of the λ. The expansion breaks down around $\lambda \sim 1$.

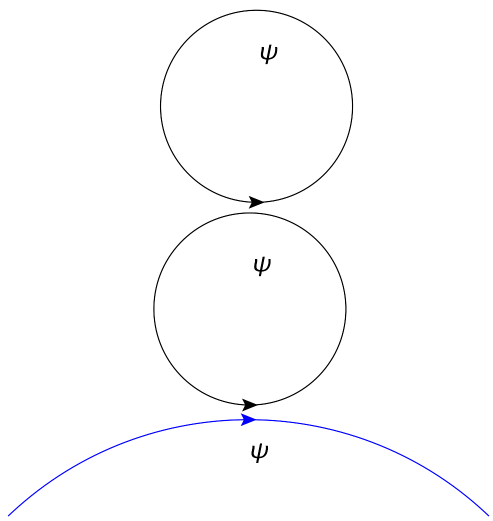


Figure 2.2: Corrections of fermion propagator

We can also see the regime-change effects from the simple tree level amplitudes. For some states at $\lambda \sim 1$, the S-matrix is of order one. Let us consider one of the such processes. Let us look for a transition from an $SO(2N)$-invariant two-particle state with the opposite momenta to the state with the two particles. We will compute the transition element from this state into a similar state. The initial and final states have the form,

$$|in\rangle = |f\rangle = \frac{1}{\sqrt{2N}} \sum_i |i,p\rangle\, |i,-p\rangle \ , \tag{2.152}$$

where $|i,p\rangle$ describes a one-particle state of a Majorana fermion of flavour i and momentum p. The transition amplitude at large N takes the following form,

$$\mathrm{Amp}_{in\to f} = \lambda\,. \tag{2.153}$$

For simplicity, a kinematic factor is skipped. The exact form of it can be reconstructed from [85] (the full $2 \to 2$ matrix is known in a large-N). For our purposes, it suffices to know that in a large-N limit, it is independent of N. The S-matrix element is of order one, so it tends to saturate unitarity for $\lambda \sim 1$. Hence, λ cannot be made arbitrarily large and $\lambda \sim 1$ sets the dimensional transmutation m_f scale. So, there is no way to push the entropy of the bound states beyond the (2.134).

The above statement is really similar to the one happens in the 4-dimensional QCD [8, 10]. The entropy of the baryon, like the entropy of the Gross-Neveu Saturon saturates bound (2.134), when the number of flavours $\sim N$, it cannot be increased more without jeopardizing the asymptotic freedom.

Now let us consider multi-particle amplitudes. In this case, we have a more interesting picture, the saturation is observed when the number of quanta is comparable with the number of fermions the Saturons consist of. This behavior was also observed with various 4-dimensional models [8]. Saturation of such processes can be understood as a regime change. The Saturon unitarizes the processes by creating the highly degenerate almost classical states, so-called Saturons. A nice thing about the Gross-Neveu model is the exactly known spectrum (using the semi-classical methods). We can directly identify the Saturons in the model. The object is in the multiplet with the maximal entropy, in the previous chapters we already labeled that object as a Saturon. It consists of N fermions, which is apparent from the (2.142). So, we, as a consistency check, look at $2N$-fermion scattering process.

We shall work in Dirac basis. Let us work with the Lagrangian (2.1). We prepare a state in which we have the equal number $n = \bar{n}$ of fermions (ψ) and anti-fermions ($\bar{\psi}$). We denote such a state by

$$|n, \bar{n}\rangle \ . \tag{2.154}$$

The fact that the fermions can be distributed in different flavours allows us to give them the same characteristic momenta.

Let us consider a S-matrix process of a transition from the vacuum $|\Omega\rangle$ to the state of n particle-anti-particle pairs,

$$\langle n, \bar{n}|\, \hat{S}\, |\Omega\rangle \ . \tag{2.155}$$

Obviously, the process in not an on-shell one, still the amplitude scaling can be extracted with respect to particle number n. If we forget about the kinematic factors, the amplitude will share similarities with the amplitude which is on-shell, for instance with the process

$1+\bar{1} \to n+\bar{n}$. In this process, particle-anti-particle pair creates a saturon-anti-saturon one. So, since these two processes share the same scaling regarding n, we will not distinguish them.

We compute the amplitude at the lowest order in the perturbation series in α,

$$\text{Amp} \propto \langle n\bar{n}| \frac{1}{n!}\frac{1}{2^n}\alpha^n T\left(:\bar{\psi}\psi\bar{\psi}\psi:\right)^n |0\rangle \;. \tag{2.156}$$

Although in a 't Hooft limit $\alpha \to 0$, the transition can be meaningfully analyzed due to the high degeneracy of the final state. As it was already said, our main interest is to extract the scaling with respect to n. We, therefore, will not derive the Green functions exactly.

Following the usual routine and collecting all the factors, the resulting amplitude of the process has the following scaling with respect to n

$$Amp \propto \frac{1}{n!}\frac{1}{2^n}\alpha^n \times (n!)^2\, 2^n \;=\; \alpha^n n!\,. \tag{2.157}$$

The corresponding cross-section scales as,

$$\sigma \,\propto\, (\alpha^n n!)^2\,, \tag{2.158}$$

notice that for any given value of α, this cross-section blows up for large n, due to the factorial growth. This growth is typical for n-particle processes[70–75]. The growth signals a breakdown of the Feynman series expansion for large n. Beyond the point, the series must be re-summed.

Let us work at the point of the optimal truncation for which the above expression gives a reliable approximation for the process. This point is given by

$$n \,=\, \frac{1}{\alpha}\,. \tag{2.159}$$

The error $\sim 1/n$. At the point of optimal truncation, the cross-section scales in the following way,

$$\sigma \,\propto\, \left(n!n^{-n}\right)^2 \sim e^{-2n} \,=\, e^{-\frac{2}{\alpha}} \tag{2.160}$$

where the Stirling approximation holds for large n (unimportant factors are dropped). The above expression represents a particular example of the general scaling (2.131).

We will relate the above computation to the production of the Gross-Neveu bound-state of the maximal entropy. For that purpose, we take $n = N$. This already gives us the first glimpse of the connection between the maximal entropy and the saturation of unitarity. Using $n = N$, the point of optimal truncation (2.159) gives $\lambda = 1$. As we already discussed, this value of the 't Hooft coupling is critical, beyond it the theory changes regime.

The total cross-section of the creation of the maximal entropy state is obtained by the summation over all degenerate microstates. This is equivalent to multiplication of (2.160) by the degeneracy factor, e^S. We obtain,

$$\sigma_{total} \,=\, \sum_{microstate} \sigma \propto e^{-2N+S} = e^{-\frac{2}{\alpha}+S}\,. \tag{2.161}$$

Let us have a closer look at the above expression.

The expression reproduces the bound (2.134) in the sense of the Gross-Neveu theory. The maximal entropy, permitted by unitarity, assumed by the state of N quanta, is $S \sim 1/\alpha \sim N$. It also indicates that the entropy of the bound-state (2.142) is very close to the saturation. It scales properly with N.

It should be noted that the expression (2.142) carries an extra factor of ln(2) compared to the critical value $S = 2N$ required for a complete saturation of (2.161). This mismatch, however, does not imply the saturation of the cross-section by pair-creation of maximal entropy bound-states being incomplete. The reason is that, at the level of the present analysis, there is an uncertainty, which can be interpreted in two ways.

First one is that the expression (2.161) is reliable when S approaches the saturation point from below, that is, from the region of weak λ. Very close to the saturation point, the higher corrections in λ must be re-summed. This cannot change the fact that the saturation takes place somewhere in the $\lambda \sim 1$ domain. Correspondingly, the bound (2.134) is robust. However, it leaves an uncertainty in exact saturation of the process by $n = N$ bound states.

Another reasoning is that the discrepancy follows from our approximation of describing the bound state as the state of N free fermions. This approximation also gives us error. The magnitude can be estimated in the following way. The fermions in the bound state are off-shell compared to their free counterparts. This is due to the interaction energy. Since we work in 't Hooft limit, α is vanishingly small. Correspondingly, interactions between the individual fermion pairs are infinitely weak. However, the collective effect is non-zero. The binding potential experienced by each fermion from the rest of N fermions is $\sim N\alpha m_f \sim \lambda m_f$.

For a weak 't Hooft coupling, $\lambda \ll 1$, this is a negligible correction to the fermion self-energy. Correspondingly, in this case, the fermion can be regarded as free. However, we have seen that the saturation of the entropy bound happens when $\lambda \sim 1$. In this case, the collective interaction cannot be ignored. The resulting potential puts the fermions off-shell.

Due to the discussion above, there are good news and some bad news. The bad news is that the approximation of a free fermion is no longer exact. That is, when approximating the production of bound states by the state of $2N$ free fermions, we are committing an error. The good news, however, is that the off-shellness is only of order one. In particular, the wave function overlap between a fermion in the bound state and its free version, is order one, rather than vanishingly small. Correspondingly, the cross-section of producing a true bound-state is expected to differ from the one of $2N$ free fermions by an exponential factor $e^{c_\lambda 2N}$. Here c_λ is an unknown number that depends on the collective coupling and order one for the critical value $\lambda = 1$. Taking into account that the maximal entropy bound state of Gross-Neveu theory is a Saturon, we estimate $c_\lambda = 1 - \ln(2) \simeq 0.3$.

Regardless of the exact saturation, the tendency is rather transparent. We see that the theory resists to an unlimited growth of the entropy of any n-particle state by hitting the unitarity bound. This is clear both from the breakdown of perturbation theory of loop expansion, and, from the saturation of unitarity by the scattering amplitudes.

Those are the signals how the theory protects itself from violating unitarity, thereby, restricting the degeneracy of the states. This restriction is manifested as the bound (2.134). In the present case, this translates as the bound on N

$$N \lesssim \frac{1}{\alpha}\,. \tag{2.162}$$

We must note that the evident saturation of the scattering cross-section by the maximal entropy bound states in Gross-Neveu model is strikingly similar to saturation of $2 \to N$ graviton scattering by black holes [69, 86]. In that analysis too, the cross-section of the individual N-particle states is suppressed by the factor $\mathrm{e}^{-\frac{1}{\alpha_{gr}}}$, which is compensated by the entropy of the black hole.

2.5.3 Time-scale of information retrieval

We have seen that the bound state with a maximal degeneracy is a Saturon. According to the conjecture [8], such objects should have Page's time, similar to black holes Page's time [80]. So, information retrieval time should be bounded from below by (2.140). In this section, we will show that, this indeed applies to Gross-Neveu saturon.

The Gross-Neveu saturon, carries flavour information, the multiplet in which the saturons live have N different indices. So, in the bound state, each of them are almost independent of the other. Hence, the Saturon practically carries a huge amount of the information. To read the information, we should determine the flavour content of the Saturon. The decoding can be done in two ways

- When Saturon has a decay channel, i.e., it is unstable, we can simply analyze the decay products. This option will be considered in the next section. The process is similar to the Hawking evaporation of the Black holes.

- If the Saturon is stable, we will not have decay products, but we can perform a scattering experiment. We can prepare an external probe, scatter it on the Saturon and look at the final states. In this section, this possibility will be discussed.

Let us prepare one Saturon and one free fermion. The Saturon is an almost classical object, so we can treat it semi-classically, while the fermion is a pure quantum object. So, the probe, in this case the fermion, should decode information from the Saturon. The free fermion forms a vector representation of $SO(2N)$ the symmetry group. The interaction rate between the fermion and the Saturon is suppressed by the coupling α^2, and enhanced by the number of fermions in the Saturon, which is N. The interaction rate between those objects is

$$\Gamma \sim m_f \alpha^2 N \sim \frac{m_f}{N}\,, \tag{2.163}$$

where the factor m_f comes from the characteristic energy of the fermion. In the final expression, we took saturation condition $N\alpha \sim 1$, which is true for the Saturon. The above expression can be translated to the information retrieval time-scale,

$$t_{min} \sim \frac{1}{\Gamma} \sim \frac{N}{m_f} \sim SR\,. \tag{2.164}$$

This time-scale represents minimum, after which we can decode at least one flavour carried by the Saturon. The time coincides with the general expression (2.140) for the minimal information retrieval time from any Saturon, including a black hole [80]. In this case, the black hole radius is replaced by the size of the Saturon in the Gross-Neveu $R \sim 1/m_f$. As a note, the real time, when we decode all the information, is bigger than the Page's time. During this time, we can only decode one flavour.

2.5.4 Hawking Evaporation in Gross-Neveu

In this section, we will give Saturon a possibility to interact softly with massless quanta. This will destabilize it and give a possibility to decay. The flavour quantum numbers should be given to the quanta, due to flavour conservation. This give us possibility to observe information retrieval via evaporation. At early stages of decay, the system should behave like the Hawking's thermal evaporation [8].

One may ask a question regarding the stability of the Saturon in the minimal Gross-Neveu model. Apparently, this effect is characteristic to all saturons, and it is called the memory burden effect discussed in [87, 88]. The stability of bound states can be understood in terms of the conserved quantum information. In the case of Gross-Neveu and generally in kink-anti-kink systems, we checked the above statement exactly [4]. Similar discussion in a different model is presented in [81].

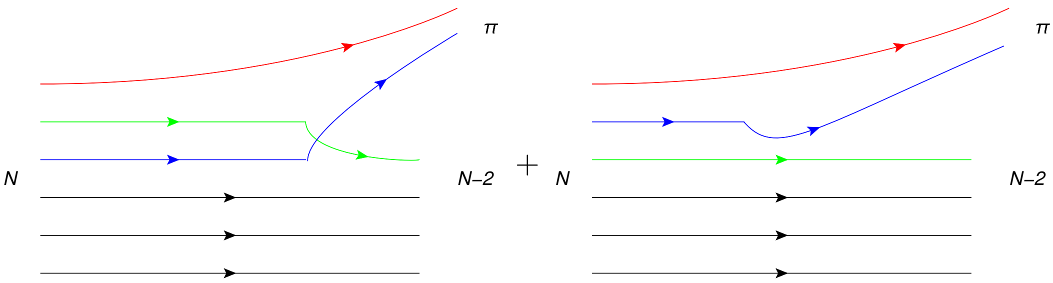
So, since we destabilize the saturon, we should observe emitted soft quanta. The emission appears to be thermal. The microscopic mechanism behind it was discussed in the case of gravity [7]. To use the same mechanism here, we should add the following field to the system, a massless pseudoscalar multiplet π_{ij}, transforming under the adjoint representation of the $SO(2N)$. Here $i, j = 1, 2, ..., 2N$ are $SO(2N)$ indices and π_{ij} is an antisymmetric tensor. We will call it a “pion”. This modification was discussed in the original work of Gross and Neveu in a different context.

In this section, we shall work in Majorana basis, so ψ_j spinor is a vectorial representation of $SO(2N)$ group. The coupling between fermions and the pion has the following form,

$$\frac{m_f}{f_\pi} \pi_{ij} \bar{\psi}_i \gamma_5 \psi_j \, , \tag{2.165}$$

where f_π is dimensionless pion decay constant. The unitarity demands that it scales as $f_\pi \sim \sqrt{N}$. The same scaling is required by the consistency of 't Hooft's planar limit [76].

The existence of the massless pion makes the Saturon unstable. Let us estimate the decay rate. The leading order contribution comes from the re-scattering of the constituent fermions, resulting in a single pion emission. The typical diagrams contributing to such a process are present in (2.166).



(2.166)

The coloured lines indicate the flow of $SO(2N)$ flavour among the constituent quanta participating in the process. The black lines denote the background fermions, their flavours at this moment are not relevant. The emitted pion carries away the flavour quantum numbers of two fermions.

To merge them in a single pion, the off-shellness of the bound state fermions must be considered. Since the theory in principle is still weekly coupled, we can consider Saturon as an off-shell fermionic bound state. We can take the interaction into account, with one rescattering. The rate of the process can be estimated, and the leading order in $1/N$ is given by,

$$\Gamma_\pi \sim m_f \alpha^2 \frac{1}{f_\pi^2} N^3 \sim m_f \sim \frac{1}{R}\,. \tag{2.167}$$

The factor $\alpha^2 \sim 1/N^2$ comes from the contact interaction among the fermions and the factor $\frac{1}{f_\pi^2} \sim 1/N$ comes from the fermion-pion interaction. Finally, the factor N^3 is the combinatorial one.

The above rate is similar to the Hawking evaporation rate of a black hole of temperature

$$T \sim 1/R \sim m_f\,. \tag{2.168}$$

We can look at $2 \to 2$ *saturons* scattering process, the cross-section has the form (2.160), the rate of the above process scales similarly. So, we get

$$\Gamma \propto e^{-\frac{2}{\alpha}}, \tag{2.169}$$

By using the fact that Bekenstein entropy coincides with the microstate one we get

$$\Gamma \propto e^{-\frac{M_N}{T}}\,. \tag{2.170}$$

Hence, in this case we get a thermal process for the early decays, as expected, since the Saturon carries a lot of flavour content, and they are emitted democratically. Here, as in case of the black holes, more energetic quanta are exponentially suppressed due to Boltzmann suppression. To emit a very massive object, we need to re-scatter a large amount of fermions, giving us the extra exponential factor e^{-E/m_f}. So, emitted quanta will be thermal. Still, it will have a small derivation from the thermality, which is proportional to $\frac{1}{N}$, or $\frac{1}{S}$ (a negligible quantity), in the case of the Saturon. This is also valid for black holes [89, 90]. The above picture has an exactly similar mechanism of hawking radiation in a large-N limit, which was conjectured for black holes [7].

Now, let us discuss what happens, when the Saturon looses the flavour content. In one process, only $\frac{1}{N}$ of the information is emitted. This information does not give us possibility to decode quantum numbers of the Saturon, or read the information which is carried by it. Let us still imagine that even this fraction of information matters. We should measure the outgoing pion. It couples with the fermions with a small coupling $\sim 1/\sqrt{N}$. So, the minimal time, in which we can decode quantum information carried by the pion is

$$t_\pi \sim NR\,. \tag{2.171}$$

Therefore, to start decoding of any significant fraction of quantum information of the Saturon, an observer has to collect at least $\sim N$ pion quanta. The time required for emitting this amount of quanta, is the same as (2.171), and this time-scale coincides with (2.164).

So, the Saturon in Gross-Neveu model has all the properties of the generic Saturon. This model does not leave any possibility to "mystify" these properties, since all of them are under control and calculable. This shows that the saturation is a universal phenomenon, not an exclusive one owned by gravity. So, we should rather study the phenomenon on simpler models and use our knowledge to extract more information about the gravity.

2.6 Large-N gravity

As in the QCD, we can add many species in the gravity. The species lower the gravitational species has the effect of lowering the gravitational cutoff $\Lambda_{\rm gr}$ [20, 91] to,

$$\Lambda_{\rm gr} = \frac{M_{\rm pl}}{\sqrt{N}}\,. \tag{2.172}$$

This gives us an idea how to define the 't Hooft-like limit [45] in this case. Following [19], instead of a single (massless) scalar, we can introduce N copies of them ϕ_j where $j = 1, 2, ..., N$ is the species index. For definiteness, we assume no self-interactions for scalars. Introduction of many scalars lowers the universal cutoff. To keep it non-zero in a large-N limit, we should take the following limit

$$M_{\rm pl} \to \infty\,, \;\; N \to \infty\,, \;\; \Lambda_{\rm gr} = \text{finite}\,. \tag{2.173}$$

As in the case of the 't Hooft limit the quantum gravitational coupling, which at any finite energy scale q is defined as,

$$\alpha_{\rm gr} \equiv \frac{q^2}{{M_{\rm pl}}^2}\,, \tag{2.174}$$

vanishes in the limit (2.173) as $1/N$. So, all processes where $\alpha_{\rm gr}$ is not accompanied by N vanish. Since the graviton does not carry the species index, simplification is much stronger than in the QCD [45]. For instance, all loop contributions vanish except the renormalization of the graviton kinetic term, which can be fully re-summed. So, the theory is simple, it is a linear gravity, with only first order interaction with the scalars.

$$\mathcal{L} = \mathcal{L}_\phi + \mathcal{L}_h - \frac{1}{2M_{\rm pl}} h_{\mu\nu} \sum_{j=1}^{N} T^j_{\mu\nu} \,. \tag{2.175}$$

The above Lagrangian, which is the limit (2.173), is exact in a large-N limit. It is the sum of the free Lagrangian of the gravity and scalars and the interaction term. We note that the number of the interacting quanta must be strictly smaller than N, or the limit is invalid. This is true, since the coupling of graviton to each of particular species is zero, but the collective effect is non-vanishing.

This example shows us how powerful the large-N techniques are and how the non-related models simplify in the same manner. We will get back to this model during the last chapter.

Chapter 3

Axion Physics and gauge forms

In the section 2.3.2 we argued, that the Yang-Mills theories do not have trivial vacua, instead they have a valley of the so called θ-vacua. Naively, this fact does not change treatment of the theory, but [92] showed that these vacua are not democratic at all, the vacuum energy is minimum for $\theta = 0$. The fact together with the observation that θ is small (2.103) motivates us to make it dynamical. An external object to QCD, makes the vacuum angle dynamical, is called an axion [13, 14]. Even without the axion, the QCD has its object, η' which is highly impacted by the QCD vacua. It possesses the famous $U(1)$ problem (for a review, see, e.g., [42]). In this chapter, we will discuss implications of QCD vacua and consistency of the axion solution. For the discussion, we highly use duality between scalars and forms using the objects introduced in the chapter 1.4. We will also check consistency to use the above objects to address other "naturalness" problems in the physics. These sections will follow our article [1].

3.1 The θ-vacua

In the section 2.3.2 we reviewed the famous θ-vacua. It appears that for different θs the vacua have different energies. The dilute-gas instanton approximation [60] (see also [58]) shows that the vacuum energy is

$$E \propto \cos(\theta), \tag{3.1}$$

valid for relatively small θs. Hence, in the vacuum valley, different vacua have different energies. So, if the superselection is lifted, the system will choose $\theta = 0$. There is a rigorous proof of $\theta = 0$ being indeed the minimum [92]. In the pure QCD, if quarks are massless, the θ-vacua, (2.101) becomes dynamical by the axial anomaly (1.59). So, if in the QCD we have a massless quark, the super-selection is disappeared.

Since the point $\theta = 0$ is special, and we experimentally observe a very strict bound on it (2.103), it is natural to assume the existence of dynamics which relaxes it towards zero. We should stress here that the problem is not a naturalness problem per se. If we treat it as a parameter, by the technical naturalness [93], due to its topological nature, it will not be corrected drastically. This perspective has a problem. θ is not a parameter, it is a solution. So, we have basically two regimes in the YM:

- We have a superselection among different θ-vacua. Then we can pick an arbitrary θ as a solution. If we do not observe it, then we can suspect, that it may be dynamically zero, and we are not in this regime.
- The theory has extra objects, for instance a massless quark, which makes parameter θ dynamical, and hence, the superselection is lifted. In this case, the theory sits at the minimum $\theta = 0$. This scenario is preferable, since the experiments leave less and less room between $\theta = 0$ and the upper bound.

So, "the strong CP problem" is more about vacuum superselection problem than a "naturalness" problem. The observation motivates us to consider the regime, where superselection is lifted. There is also a very solid theoretical argument against the regime, where we have the superselection. The quantum gravity requires absence of the superselection sector due to the consistency [63].

In this section, we will discuss further implications of the θ-vacua and during next sections we will consider the consistency of the axion physics.

3.1.1 The $U(1)$ problem

Let us consider QCD, a YM theory with N_f fundametal Dirac fermions. Let us concentrate on the regime, where the quarks are light/almost massless

$$\mathcal{L} = -\frac{1}{4}G^2 + i\bar{\psi}\partial\!\!\!/\psi. \tag{3.2}$$

The theory has the following global symmetry

$$U(N_f)_L \times U(N_f)_R, \tag{3.3}$$

where in L, R we mean rotations of left and right fermions, respectively. The above symmetry can be decomposed in the following way,

$$SU(N_f)_V \times U(1)_B \times SU(N_f)_A \times U(1)_A, \tag{3.4}$$

where $SU(N_f)$, $U(1)_B$ are vectorial and $SU(N_f)_A$, $U(1)_A$ axial symmetries, respectively. The $SU(N_f)$ is spontaneously broken and as a consequence goldstone bosons (we call all of them pions) are observed. Meantime, all the vectorial symmetries are unbroken. The problem was the $U(1)_A$, which was neither observed as an unbroken symmetry, nor the corresponding goldstone boson (η') (for a review see e.g., [42]). The 't Hooft [93] anomaly matching conditions suggest that all the axial symmetries should be broken spontaneously. So, missing the goldstone corresponds to $U(1)_A$ being a puzzle solved in two different ways by 't Hooft [55, 94, 95] using instantons and by Witten and Veneziano [96, 97] using θ-vacua properties in large-N QCD (see [42] for a review).

A long story short, there is still a Goldstone boson η', but due to anomaly the axial current is not conserved, hence η' gets mass from the anomaly.

Let us discuss the 't Hooft's treatment. In previous chapters, we discussed that change of fermion chirality is connected to the change of winding via anomaly (2.65). So chiral rotations β basically shift θ.

$$\theta \to \theta - 2N_f\beta. \tag{3.5}$$

which means that instantons explicitly break $U(1)_A$. Meantime, they do not break the rest of the symmetry and due to the periodicity of θ-vacua ($\theta \to \theta + 2\pi$, (2.97)) a discrete subgroup Z_{2N_f} of $U(1)$ survives. So, the instanton should generate an effective vertex out of fermions, respecting the above symmetries. The effective vertex, called the 't Hooft vertex, has the following form

$$\mathcal{L}_{t'Hooft} = C det\bar{\psi}\psi = Ce^{i\frac{\eta'}{f_\eta}}|det\bar{\psi}\psi|, \tag{3.6}$$

where C is some dimensional constant and determinant is taken in the flavour space. The phase of this determinant represents the η' goldstone, which means that it gets its mass from the fermion condensate. Obviously, the group Z_{2N_f} is a subgroup of $U(1)$ and the rest of the flavour symmetry remains unbroken.

To show the above more rigorously [55], we consider vacuum to vacuum transition amplitudes in presence of one instanton. Foremost, let us take an instantonic solution with a unit winding number, then consider all quantum fluctuations around that solution,

$$A = A^I + a, \tag{3.7}$$

where A^I is the instanton solution and a is a fluctuation on top of it. The Lagrangian up to the second order will be

$$\mathcal{L} = \mathcal{L}_{\mathcal{I}} - aM_1a - \bar{\psi}M_2\psi - \bar{c}M_3c, \tag{3.8}$$

where $\mathcal{L}_I$ is a pure instanton Lagrangian and Ms are instanton-dependent operators, and $c, \bar{c}$s denote Fadeev-Popov ghosts. A transition between two vacua is given by

$$\frac{\langle 0|0\rangle}{\langle 0_0|0_0\rangle} \propto finite, \tag{3.9}$$

where 0_0 denotes a vacuum without an instantonic background. The vacuum to vacuum amplitude (using the WKB (2.3.1)) of the configuration (3.8) is given by

$$\langle 0|0\rangle \propto (detM_1)^{-\frac{1}{2}} detM_2 detM_3 e^{-S_I}. \tag{3.10}$$

As we will see M_2 has zero modes (2.2.4), which means that this transition will be zero. We should take care of zero modes. We should change the winding number to have a non-trivial vacuum transition. So, let us add some external source, i.e., a new term to the Lagrangian (3.8)

$$\mathcal{L} = \bar{\psi}^s J_{st} \psi^t, \tag{3.11}$$

where indices enumerate flavours. It means that the source is colour-blind, but it should have some Lorentz structure. For example, it can include γ_5. There is exactly one zero mode for each flavour of fermions, since a fundamental Dirac fermion has exactly one zero mode in an instantonic background (2.2.4). Let us study the M_2 operator at the lowest order. The corresponding equation of eigenvalues is,

$$M_2 \psi_s + J_{st} \psi_t = E_s \psi_s, \tag{3.12}$$

where Es are the eigenvalues in the presence of the external source. Considering the fact of these functions being zero-mode wave functions, the above expression is simplified in the following way

$$J_{st} \psi_t = E_s \psi_s. \tag{3.13}$$

The zero-mode wave functions are square integrable and orthonormal. Therefore, we get

$$\int d^4x \psi_i^{*s} J_{st} \psi_f^t = E_i \int d^4x \psi_i^{*s} \psi_f^s = E_i \delta_{if}, \tag{3.14}$$

from which it follows

$$det(M + J) = detM' \prod_i E_i \propto det \int d^4x \psi_0 J_{st} \psi_0, \tag{3.15}$$

where M' is the part of the determinant which does not have zero modes. In fact, it is the same as the fermionic determinant in the vacuum. This integral tells us that the vertex (3.6) is present in an effective action. So, we can see the important role played by zero modes, they shape the low-energy physics. η' gets mass from the structure of the θ-vacua.

The same conclusion can be made by Witten and Veneziano treatment [96, 97]. Let us consider the θ term Lagrangian (2.101), but in a large-N, then we get

$$\mathcal{L}_\theta = \frac{\theta \lambda}{32 N \pi^2} G\tilde{G}, \tag{3.16}$$

where λ is the 't Hooft coupling [45]. Then vacuum energy (3.1) in terms of N powers (we will see it explicitly) has the following form,

$$E \propto N^2 F\left(\frac{\theta}{N}\right), \tag{3.17}$$

where F is an arbitrary function. This computation does not match with the 't Hooft computation, since one instanton is exponentially suppressed in this case. Nevertheless, there is a loophole, dilute-gas approximation fails in this case. So, the above two computations are not comparable. Now combining (3.1) and the path integral quantization including the θ term (3.16), we get

$$\frac{d^2E}{d\theta^2}|_{\theta=0} = \left(\frac{\lambda}{32N\pi^2}\right)^2 \int d^4x \, \langle | \, G\tilde{G}(x)G\tilde{G}(0) \, | \rangle \, , \tag{3.18}$$

During the discussion we always take a derivative at the point $\theta = 0$, so we do not write it explicitly anymore. One should note that here we have equal time correlator*. We can define the above correlator in a momentum space

$$U(k) = \int d^4x e^{ikx} \, \langle | \, G\tilde{G}(x)G\tilde{G}(0) \, | \rangle \, , \tag{3.19}$$

So, the equation has the following form

$$\frac{d^2E}{d\theta^2} = \left(\frac{\lambda}{32N\pi^2}\right)^2 U(0), \tag{3.20}$$

where the correlator $U(k)$ can be written in terms of the physical particles [76] (in a large-N limit),

$$U(k) = \sum_{glueball} \frac{N^2 a_n^2}{k^2 - m_n^2} + \sum_{mesons} \frac{N b_n^2}{k^2 - M_n^2}, \tag{3.21}$$

where m and M are meson and glueball masses respectively, and $a_n N$ and $b_N\sqrt{N}$ are their creation amplitudes from $\tilde{G}G$ respectively. In a pure glue, we have only the first term and hence, the second derivative of vacuum energy,

$$\frac{d^2E}{d\theta^2} \sim 1. \tag{3.22}$$

The (3.17) is evident from here. In a theory where we also have massless quarks, vacuum energy must be independent of θ. It is believed that there is one meson, with mass $\sim \frac{1}{\sqrt{N}}$. The above particle exactly cancels the dependence (the cancellation happens due to the equal time correlator [76]). We can assume that, it is η'. So, we get

$$U(0)_{pure\,YM} = N\frac{b_{\eta'}^2}{M_{\eta'}^2}, \tag{3.23}$$

*For the details see appendix of [96]

the above expression can be written in a simpler form, remembering that

$$\sqrt{N} b_{\eta'} = \langle 0| G\tilde{G} |\eta'\rangle \,, \tag{3.24}$$

using the anomaly (1.59,3.5) it simplifies as

$$\sqrt{N} b_{\eta'} = \frac{16\pi^2 N}{N_f \lambda} \langle 0| \partial_\mu j^A_\mu |\eta'\rangle = \frac{16\pi^2 N}{N_f \lambda} f_{\eta'} M^2_{\eta'}, \tag{3.25}$$

where j^A is an $U(1)_A$ axial current and $f_{\eta'}$ a η' decay constant. In the second equation we used the connection between a goldstone boson and a broken current,

$$j^A_\mu \Longleftrightarrow f_{\eta'} \partial_\mu \eta'. \tag{3.26}$$

The final form for η' mass is (The Witten-Veneziano formula)

$$M^2_{\eta'} = \frac{4N_f^2}{f^2_{\eta'}} \frac{d^2 E}{d\theta^2}|_{pure\, YM}. \tag{3.27}$$

In a large-N limit $f_{\eta'} = \sqrt{N_f} f_\pi$, where f_π is a decay constant of the goldstones corresponding to the $U(N_f)_A \to 0$ breakdown. So, the final formula is

$$M^2_{\eta'} = \frac{4N_f}{f^2_{\pi}} \frac{d^2 E}{d\theta^2}|_{pure\, YM}. \tag{3.28}$$

This discussion fills the gaps in 't Hooft's discussion. For instance, it is technically impossible to fix coefficient C in (3.6), since it depends on IR physics. Witten-Venziano formula tells us that in large-N leading terms already give a non-zero mass to η'. Most importantly, these discussion shows us $U(k)|_{k\to 0} \neq 0$ (topological susceptibility of θ-vacua). This fact will have a considerable impact on our next discussion.

3.1.2 Effective description of θ-vacua

In the previous chapter we discussed an important quantity $U(0)$, which shows how vacua responds to change of θ, it is called a topological susceptibility. The above quantity is a correlator of two $\tilde{G}G$s. In large-N we saw it should not be zero in pure YM, or η' will be massless. In the nature, we observe that η' has non-zero mass in the massless quark limit.

In addition, it was argued that the above correlator encodes all information of θ-vacua [98]. So, if we have a superselection sector, then

$$\langle 0| G\tilde{G}(0), G\tilde{G}(q) |0\rangle_{q\to 0} = const \neq 0. \tag{3.29}$$

We consider correlators in the momentum space, the Fourier transform is omitted when it is clear from the context. Considering $G\tilde{G} = dK$, (2.53) where K is Chern current, which is hodge dual (1.77) to Chern 3-form C, we get

$$\langle 0| CC |0\rangle_{q\to 0} \propto \frac{1}{q^2}, \tag{3.30}$$

which means that C-field is a massless 3-form (see chapter 1.4). It is conjectured that YM is a mass-gap theory. So, there are no massless propagating degrees of freedom at low energy. So, θ-vacua at low energy are completely described by the effective theory of C-field. We can integrate all other fields out. The Lagrangian has the form [15, 99]

$$\mathcal{L} = \mathcal{K}(E) + f(\partial E), \tag{3.31}$$

where E is an electric field (1.111) of C, $\mathcal{K}(E)$ is an algebraic function of E and f is some function of electric field derivatives. As all other integrated out particles were massive, derivatives will be suppressed by some scales in the function f. This means that we can consider only the first part of the Lagrangian at the small momentum limit. Considering the fact of YM Hamiltonian being CP-conserving, $\mathcal{K}(E)$ must be a CP-even function because $G\tilde{G}$ is a CP violating term. During the section we take a QCD scale to be unity, so instead of $\frac{E}{\Lambda^2}$ we just write E. The effective Lagrangian at the lowest order must have the form

$$\mathcal{K}(E) = \frac{1}{2}E^2 + ... \tag{3.32}$$

From the equations of motions (1.109) of C-field we get

$$\partial_\mu \frac{\partial \mathcal{K}(E)}{\partial E} = 0. \tag{3.33}$$

This equation is satisfied by an arbitrary constant electric field, namely $E = \theta$. $E = 0$ is not anymore a vacuum around this solution. We should take $E = \theta$ as a vacuum point and consider fluctuations around it. Putting solution back into (3.32) and considering fluctuations around it, we get

$$\mathcal{K} = \theta E + \frac{1}{2}E^2 + ... \tag{3.34}$$

In the language of QCD Lagrangian it gives us the term from (2.101)

$$\propto \theta G\tilde{G}. \tag{3.35}$$

3-forms give us beautiful and simple descriptions for the appearance of such term. In the language of 3-forms, we will have a so-called strong CP problem if a constant solution of an electric field is not prohibited. Solving the strong CP puzzle in this language means forbidding of constant solutions. To make this point clear, let us reconsider correlators (3.29), (3.30). As we see, forbidding a non-zero constant solution in (3.29) gives us a vacuum around zero electric field configuration. This means that (3.30) should not have a massless pole (see, e.g., [100]),

$$\langle 0 | CC | 0 \rangle_{q\to 0} \propto \frac{1}{q^2 - m^2}, \tag{3.36}$$

which will guarantee the theory to have a trivial vacuum with zero electric field. For this procedure, we need to give a mass to the 3-form. As we have seen previously, it can be done by higgsing the 3-form, see the chapter 1.4. So, we should add one propagating degree of freedom in the theory. This can be done by adding a 2-form or by adding a shift invariant scalar field (the same in the dual-language). Shift-invariant scalars usually appear in the theory due to some spontaneous symmetry breaking (SSB). We naturally have such scalar in the theory, the goldstone boson η' corresponding to the break of $U(1)_A$ current. The current and goldstone correspondence (3.26) gives us a unique coupling of η' and the 3-from C,

$$\eta' G\tilde{G}, \tag{3.37}$$

η' transforms in the following way

$$\eta' \to \eta' + c, \tag{3.38}$$

where c is an arbitrary constant. So, the above coupling is invariant under the transformation. The theory is consistent, since a godstone boson is dual to the 2-form (see section 1.4). The above coupling reproduces (3.25) too. Hence, in terms of the electric field and η', Lagrangian has the form

$$\mathcal{L} = \mathcal{K}(E) - \frac{\eta'}{f_\eta} E + \frac{1}{2}(\partial \eta')^2, \tag{3.39}$$

where f_η is a decay constant for η'[†]. Keeping only the leading terms in the expansion (3.32), we will get a simple Lagrangian

$$\mathcal{L} = \frac{1}{2}E^2 - \frac{\eta'}{f_\eta} E + \frac{1}{2}(\partial \eta')^2, \tag{3.40}$$

from which the equations of motion for the corresponding fields read

$$\begin{aligned} \partial_\mu (E - \frac{\eta'}{f_\eta}) &= 0 \\ \Box \eta' &= -\frac{1}{f_\eta} E. \end{aligned} \tag{3.41}$$

The solution for the electric field has the form

$$E = \frac{\eta'}{f_\eta} - \theta, \tag{3.42}$$

where θ is some constant, which should be fixed by boundary conditions. The equation of the motion for η' in that background takes the following form,

$$\Box \eta' + \frac{1}{f_\eta}(\frac{\eta'}{f_\eta} - \theta) = 0. \tag{3.43}$$

[†]since we are not working in canonically normalized fields, this decay constant differs from the one used in the previous section

The equation gives us 3 important results. The η' gets mass from the 2-form, as it was expected from the general considerations (see section 1.4), the mass in above units reproduces (3.28) and the vacuum expectation value (vev) of η' in the presence of $\theta \neq 0$ vacuum is

$$\eta' = f_\eta \theta. \tag{3.44}$$

Reconsideration of (3.42) fields around vev brings the vacuum of the electric field at $E = 0$. We can see that all ingredients are combined in one massive pseudo-scalar degree of freedom. One important point is that all contributions, including instantonic and non-large-N ones of break of chiral symmetry, are included in this mass. To see how an extra mass term will break the picture, let us consider the Lagrangian [101]

$$\mathcal{L} = \frac{1}{2}E^2 - \frac{\eta'}{f_\eta}E + \frac{1}{2}(\partial\eta')^2 - \frac{1}{2}\mu^2\eta'^2, \tag{3.45}$$

where μ is an "extra mass". The equation of motion for the field-E is the same, but the one for η' now has the form

$$\Box\eta' + \frac{1}{f_\eta}(\frac{\eta'}{f_\eta} - \theta) + \mu^2\eta' = 0. \tag{3.46}$$

It is obvious from the above that η' will get vev and the electric field in the vacuum will be

$$E = -\frac{\theta f_\eta^2 \mu^2}{1 + f_\eta^2\mu^2}. \tag{3.47}$$

This means that CP-problem will not be solved. In QCD, usually, all the quarks are assumed to be massive, so the above description is correlated to that.

To solve the strong CP problem in QCD, where all the quarks are massive, we should add a new degree of freedom. The new degree of freedom should work effectively, like η' works in a massless limit. We consider this mechanism in the next section.

3.1.3 PQ Mechanism and the axion

To solve the strong-CP problem, Peccei and Quinn (PQ) suggested an extra $U(1)$ symmetry. It is shifted under phase rotations. Such field shifts the θ-angle, as axial transformations do in the case of the massless quarks.

Schematically, the PQ Lagrangian [12, 62] has an interaction term with a complex field ϕ and a Dirac fermion. That part of the Lagrangian has the following form

$$\mathcal{L} = i\bar{\psi}\not{D}\psi - G\bar{\psi}(1-\gamma_5)\phi\psi + V(|\phi|) + h.c.\,, \tag{3.48}$$

G is a complex coupling constant. To keep the symmetry intact, the phase of the scalar field must be transformed in the same way as the phase of the chiral fermion does. It is clear that, by simultaneous rotations of fermion and scalar phases, θ can be shifted away from the Lagrangian.

In the above system, we should consider spontaneous symmetry breaking to give masses to fermions. Then there should be a goldstone boson a, corresponding to the broken chiral rotation. The boson will be invariant under the shift symmetry,

$$a \to a + c, \tag{3.49}$$

where c is an arbitrary constant. The above constant corresponds to the phase shift of the scalar ϕ. So, under the above transformation, θ shifts linearly too. If we integrate all the heavy fields out, we will be left with the following coupling between a and YM sector,

$$\mathcal{L} \propto \frac{a}{f_a} G\tilde{G}, \tag{3.50}$$

where f_a is the a-field decay constant. Above coupling is exactly similar to (3.37). The field a is called an axion [13, 14]. The axion solves the strong-CP problem in the same manner as η' does in the case of massless quarks.

We may ask, what happens when η' and an axion are present at the same time [101]. We can rewrite the Lagrangian in terms of the electric field

$$\mathcal{L} = \frac{1}{2}E^2 - \frac{\eta'}{f_\eta}E - \frac{a}{f_a}E + \frac{1}{2}(\partial\eta')^2 + \frac{1}{2}(\partial a)^2. \tag{3.51}$$

Here we see that one combination of these fields gets vev while the orthogonal to that combination does not,

$$E = \frac{\eta'}{f_\eta} + \frac{a}{f_a} - \theta\,, \tag{3.52}$$

combine this equation with the ones for a and η', we will get the following result

$$\Box\,(\eta' + a) + \left(\frac{1}{f_\eta} + \frac{1}{f_a}\right)\left(\frac{\eta'}{f_\eta} + \frac{a}{f_a} - \theta\right) = 0. \tag{3.53}$$

The combination $\frac{\eta'}{f_\eta} + \frac{a}{f_a}$ gets vev and ensures that $E = 0$ is the vacuum of the theory. The second combination of the fields will not have any influence on the vacuum structure and will continue to be a free goldstone.

These two alternative solutions of the strong CP problem lead us to the conclusion that all possible solutions can be described in the same manner. Description of 3-forms is very useful in this case. Above solutions have similar forms in this language. To solve the strong CP problem, it is necessary to have some goldstone boson. Then it can be coupled to the 3-form and give mass to it. To get Goldstone, we should have some spontaneous symmetry breaking. There is one more property need to be satisfied by this scalar, it should be a broken phase of some axial symmetry. This is a necessary condition to generate a coupling with Chern 3-form.

In dual Language, we will have the Lagrangian (see chapter 1.4)

$$\mathcal{L} = -\frac{1}{48}F^2 + \frac{1}{2}m^2(C - dB)^2. \tag{3.54}$$

We got the field B from an axion by transforming (1.100). Its normalization is not canonical in the Lagrangian. The mass m and the axion decay constant are related to each other (up to normalization and taking intro account that we took QCD scale to be unity) as

$$m \propto \frac{1}{f_a}, \tag{3.55}$$

As a final note, we should mention, that gravity could have θ-vacua type structure too. Similarly to YM theory, gravity also admits a topological 3-form. The quantity $R\tilde{R}$ is Hodge dual to the 3-form derivative. This quantity is also connected to the axial symmetry (1.69). When we do an axial transformation, this object shifts, as $G\tilde{G}$ does in the YM theory. The above facts motivate us to examine the correlator of $R\tilde{R}$ at the zero momentum limit [15, 101]

$$\langle 0 | R\tilde{R}(q), R\tilde{R}(0) | 0 \rangle_{q \to 0} \overset{?}{\neq} 0.$$

Unfortunately, we cannot calculate this correlator. Since it can be non-trivial, we should consider it carefully. It will play a role during our discussions.

3.2 Consistency of the axion solution to strong-CP problem

In the previous chapter, we reviewed an effective description of θ-vacua [15]. The description can include the axion and η' physics. We should care about UV-sensitivity of the axion. The PQ mechanism is formulated using global symmetry, which may be broken due to quantum effects. So, people argue that the axionic solution could be jeopardized by high dimensional operators, for instance, generated by gravity [102–104]. In [15], it was argued that this could not happen, since the axion has the dual gauge formulation (see the previous section and the section 1.4). As we discussed in the section 1.2 gauge symmetries cannot be broken in self-consistent theories.

So, writing axion in terms of the 2-form gives us an advantage. We can fully control the objects, which can jeopardize our description. Gauge/BRST invariance will forbid all the "arbitrary" corrections, which presences are not clear in the usual formulation of the axion. In this picture, only extra massless 3-forms can spoil the axionic solution, from the all unwanted corrections. As we discussed in the previous chapter, the gravity could have such an object. It appears, that disruption from gravitational Chern-Simons can be easily avoided if the theory contains an additional chiral symmetry that is anomalous regarding the gravity. In fact, the role of a "protector" for axion can be played by a chiral symmetry of a light fermion, such as neutrino [15, 100, 105].

We will complete the story, considering all the possible couplings and explicitly checking UV insensitivity of the axion. We follow our article [1].

3.2.1 Duality

As we discussed in the section 1.4, 2-forms are dual to the massless shift invariant scalars. In other words, the theory

$$\mathcal{L} = \frac{1}{12}(dB)^2, \tag{3.56}$$

where B is a 2-form being dual to a free massless pseudoscalar a,

$$\mathcal{L} = \frac{1}{2}(\partial a)^2\,, \tag{3.57}$$

The duality at the level of free theory is straightforward. However, for some time, the dualities at the level of massive interacting theories were sources of controversy. It has even been suggested [106] that duality ceases to exist at the level of an interacting theory, and that interaction and mass terms break duality.

The above issue first was clarified in [15]. It was proven that the duality still holds for massive and interacting axion field with an arbitrary scalar potential $V(a)$. It appears that non-derivative terms of the axion dualize to a non-trivial kinetic function of a massive 3-form field $C_{\mu\nu\alpha}$, where the 2-form $B_{\mu\nu}$ enters as the longitudinal degree of freedom (a bit short discussion of this was considered in chapter 1.4). This duality does not have discontinuity, neither for the zero coupling nor for the zero-mass limit. The number of degrees of freedom does not change. We will review the duality in details.

Let us consider a theory of a massive pseudo-scalar axion with an arbitrary potential $V(a)$

$$\mathcal{L} = \frac{1}{2}(\partial a)^2 - V(a)\,. \tag{3.58}$$

This theory is dual to the theory of an interacting massive 3-form C with the field strength $F_{\mu\alpha\beta\gamma} = \partial_{[\mu}C_{\alpha\beta\gamma]} = \epsilon_{\mu\alpha\beta\gamma}E$,

$$\mathcal{L} = \Lambda^4\mathcal{K}\left(\frac{E}{\Lambda^2}\right) + \frac{1}{2}m^2C^2\,, \tag{3.59}$$

where Λ and m are parameters of mass dimensionality. The quantity E is an "electric" field (1.111). The $\mathcal{K}\left(\frac{E}{\Lambda^2}\right)$ is a non-derivative function of its argument E, which satisfies,

$$V(a) = \frac{1}{\sqrt{6}}m\Lambda^2\int da\,\mathbf{inv}\,\mathcal{K}'\left(\frac{ma}{\sqrt{6}\Lambda^2}\right) \tag{3.60}$$

where prime denotes a derivative w.r.t the argument and **inv** stands for an inverse function. For the simplest choice, $\mathcal{K}(x) = \frac{1}{2}x^2$ we get quadratic and canonically normalized Lagrangian.

Let us prove the above, starting with the decomposition of the massive field C into transverse and longitudinal modes,

$$C = C^T - dB \tag{3.61}$$

The longitudinal mode (2-form) is the only propagating mode of the massive 3-form field (see 1.4). In this formulation, it also transforms under gauge transformations (1.113). So, the combination is gauge invariant

$$C^T \to C^T + d\Omega\,, \quad B \to B + \Omega\,. \tag{3.62}$$

Here Ω is a gauge shift parameter, which represents an arbitrary two-form.

The above decomposition does not change the theory. In the decomposed form, the theory has the following form

$$\mathcal{L} = \Lambda^4 \mathcal{K}\left(\frac{E}{\Lambda^2}\right) + \frac{1}{2} m^2 (C^T - dB)^2\,. \tag{3.63}$$

Now we can see that the limit $m \to 0$ does not have any discontinuity, the two sectors simply decouple.

In this form, we can perform dualization. We already described the procedure in the section 1.4, but we will repeat it here too due to its importance. The procedure says that we should consider $dB \equiv X$ as a fundamental 3-form and impose Bianchi identity ($dX = 0$) as a constraint through a Lagrange multiplier a,

$$\mathcal{L} = \Lambda^4 \mathcal{K}\left(\frac{E}{\Lambda^2}\right) + \frac{1}{2} m^2 (C^T - X)^2 + \frac{1}{\sqrt{6}} m\, a\, \epsilon_{\mu\alpha\beta\gamma} \partial_\mu X_{\alpha\beta\gamma}\,. \tag{3.64}$$

Integration of X out, gives us the Lagrangian, where instead of the 2-form we have an axion,

$$\mathcal{L} = \Lambda^4 \mathcal{K}\left(\frac{E}{\Lambda^2}\right) + \frac{1}{2} (\partial a)^2 - \frac{1}{\sqrt{6}} m\, a\, E. \tag{3.65}$$

The above Lagrangian is exactly (3.39) for the axion. The gauge symmetry (3.62) of the 2-field is replaced by the global shift symmetry of the axion,

$$a \to a + \text{const}\,. \tag{3.66}$$

To see what the potential C generates we should consider the equations of motion, which read as follows for the field C

$$\partial_\mu \left(\mathcal{K}'\left(\frac{E}{\Lambda^2}\right) - \frac{m}{\sqrt{6}\Lambda^2} a \right) = 0, \tag{3.67}$$

whereas the one of axion is

$$\Box a = \frac{1}{\sqrt{6}} m E\,. \tag{3.68}$$

Solving for E as a function of a from (3.67) and considering (3.68) it follows

$$E(a) = \frac{\sqrt{6}}{m} \frac{dV(a)}{da}\,. \tag{3.69}$$

The above equations establish the relation (3.60). So, finally, we arrive to the theory (3.58) with the potential $V(a)$ determined by the function $\mathcal{K}$ via the relation (3.60). We reviewed the duality part of the article [15].

From the duality, it is clear that we can use the above languages interchangeably. The language of the forms is more reliable, since we have the gauge symmetry (in the case of consistent theories it is protected by the BRST symmetry). So, these symmetries cannot be broken, hence, the solution of the strong-CP problem can be jeopardized only with the small amount of the objects. These objects should have the gauge structure compatible with the description. This point is not transparent in the language of the PQ symmetry, and it is the reason of having brought so many controversies. So, if we choose the dual formulation of the axion, it will be protected from unwanted UV corrections.

We should note, that the example of the $\mathcal{K}(x)$ function can be computed in QCD using the instantons in a dilute gas approximation. The approximation gives the potential $V(a) \propto \cos\left(\frac{\mathrm{a}}{\mathrm{f_a}}\right)$ [58, 107], where f_a is an axion decay constant. The above potential corresponds to [15] the function $\mathcal{K}(x) \propto (x \arcsin x + \sqrt{1-x^2})$. One might worry about the form of the $V(a)$, but in the dual formulation, we see that irrespective of the exact form $\mathcal{K}(x)$, the solution cannot be undone. The higgs effect there is due to the 2-form.

The dual language also makes it transparent why the existence of additional external potential $\tilde{V}(a)$ is un-Higgsing the 3-form and restoring the superselection of θ-vacua. The theory with an additional potential has the following form,

$$\mathcal{L} = \Lambda^4 \mathcal{K}\left(\frac{E}{\Lambda^2}\right) + \frac{1}{2}(\partial a)^2 - \frac{1}{\sqrt{6}} m\, a\, E - \tilde{V}(a)\,. \tag{3.70}$$

The above Lagrangian can be rewritten in a dual language. Here, one axion will be shared among two 3-forms [15]. So, we can rotate these 3-forms in the basis, where one is massless, and the other one is higgsed (similarly to the rotation of W_3 and B to Z and A in the standard model). Hence, the 3-form, corresponding to the QCD vacua, gets a long-range correlator. So, the superselection of θ-vacua is back. We see, clearly, how the extra 3-form can jeopardize the axionic solution. In the next section, we will discuss such issues in details.

3.2.2 UV-in-sensitivity of Dual Axion Mechanism

In this section, we will discuss the other point of the dual formulation of the axion. The axionic solutions of the strong CP problem is insensitive to UV heavy states [15]. If we have some heavy fields at the scale of M_f coupling them to the axion will not change the picture. In particular, if we couple the 3-form with any heavy brane, they cannot regenerate θ-vacua. Apparently, the statement is exact. The above one contradicts the common observations. Usually, heavy objects change the light objects just a bit. In this case, the gauge symmetry protects the solution from such modifications. The electric field E remains exactly zero, as long as new physics brings new massless poles in 3-form correlators. The proof, which relies on integration out of the heavy physics, can be found in [15]. We will not repeat the above steps, but we will cross-check the proof using the examples.

Let us consider an explicit resolution of the brane, like the one in [108]. Let us assume the C-form in the Lagrangian (3.65) to be the QCD Chern-Simons field. Now we add their couplings to a heavy brane coming from some UV-physics. This heavy brane can be resolved in the form of a soliton of a heavy axion b with the effective potential $V(b)$. The theory is described by the following Lagrangian

$$\mathcal{L} = -\frac{1}{48}F^2 - \frac{1}{2\pi}\partial_\mu(qa + q_b b)\epsilon_{\mu\alpha\beta\gamma}C_{\alpha\beta\gamma} + \frac{1}{2}f_a^2(\partial a)^2 + \frac{1}{2}f_b^2(\partial b)^2 - V_b(b)\,, \tag{3.71}$$

where f_a and f_b are decay constants of a and b axions, which are not canonically normalized, and $\mathcal{K}(x) = \frac{1}{2}x^2$. We parameterize the mixing between 3-form and axions with constants q and q_b with the mass dimension-2. The connection between the parameters of (3.71) and (3.65) is straightforward.

The equations of motion show that, despite the presence of the heavy field b, the electric field E corresponding to QCD Chern 3-form vanishes in the vacuum. We can deduce the result from the equation

$$f_a^2\Box a = \frac{q}{2\pi}E\,, \tag{3.72}$$

Since the axion is sourced by E, only $E = 0$ can be the vacuum solution. This insensitivity of the vacuum towards heavy physics (in this case heavy axion b) is transparent in the language of the 3-form Higgs effect. Let us use the exact duality, we can rewrite the potential for axion b as the second 3-form C^b. The Lagrangian reads,

$$\begin{aligned}\mathcal{L} = -\frac{1}{48}F^2 - \frac{1}{2\pi}\partial_\mu(qa + q_b b)\epsilon_{\mu\alpha\beta\gamma}C_{\alpha\beta\gamma} + \frac{1}{2}f_a^2(\partial a)^2 + \\ + \frac{1}{2}f_b^2(\partial b)^2 - \frac{q_{bb}}{2\pi}\partial_\mu b\epsilon_{\mu\alpha\beta\gamma}C^b_{\alpha\beta\gamma} + \Lambda^4\mathcal{K}_b\left(\frac{E^b}{\Lambda^2}\right)\end{aligned} \tag{3.73}$$

where $\mathcal{K}_b$ can be found with the same condition (3.60) w.r.t $V_b(b)$. There are two 3-forms and two scalars in this theory. So, both of the scalars are "eaten" by the 3-forms without leaving any massless degree of freedom. Hence, the CP-violating order parameters are zeros.

The situation is drastically different if the new physics, coupled to ordinary axion, comes in a form of a 3-form without its axion partner (for instance (3.56)),

$$\mathcal{L} = -\frac{1}{48}F^2 - \frac{1}{2\pi}\partial_\mu a\epsilon_{\mu\alpha\beta\gamma}(qC_{\alpha\beta\gamma} + q_b C^b_{\alpha\beta\gamma}) - \frac{1}{48}F_b^2 + \frac{1}{2}f_a^2(\partial a)^2 \tag{3.74}$$

This example was considered in [15] as violating the criterion of absence of massless 3-forms. The new physics in this setup is represented by C_b without an axion partner. Correspondingly, a new massless pole is brought. Integrating out C_b and a, we recover (up to notations) the equation (35) of [15],

$$\mathcal{L} = -\frac{1}{48}F^2 - \frac{1}{48}\frac{q^2}{q_b^2}F_{\mu\alpha\beta\gamma}\frac{m^2}{\Box + m^2}F_{\mu\alpha\beta\gamma}\,, \tag{3.75}$$

where $m^2 = \frac{q_b^2}{4\pi^2 f_a^2}$. In the zero momentum limit ($\Box \to 0$) the second term only corrects normalization of the first term. The effective theory contains a long-range correletor in the form of a massless 3-form field. So, there exists a vacuum solution with a constant E. In this theory, superselection is back, and the CP is broken.

To double-check the above treatment, we can perform analogous integration from (3.73). Assuming $\mathcal{K}_b(x) = \frac{1}{2}x^2$ the effective theory for C is,

$$\mathcal{L} = \frac{1}{2}C_{\alpha\beta\gamma}\left(\Box + m_a^2 + \left(\frac{q_b}{q_{bb}}\right)^2 \frac{m_b^2\Box}{\Box + m_b^2}\right)\Pi_{\alpha\mu}C_{\mu\beta\gamma}\,, \tag{3.76}$$

where $\Pi_{\alpha\mu} = \eta_{\alpha\mu} - \frac{\partial_\alpha\partial_\mu}{\Box}$ is a transverse projector, $m_a^2 = \frac{q^2}{4\pi^2 f_a^2}$ and $m_b^2 = \frac{q_{bb}^2}{4\pi^2 f_b^2}$. The field C has mass, so it cannot have constant electric-field solutions. In this case, there is no vacuum super-selection. We confirm the statement of [15] about the insensitivity of the dual formulation of the axion regarding heavy physics.

We note that an alternative physical way of understanding this UV-insensitivity is that the generation of axion mass from 3-form Higgs effect can be understood in purely topological terms, as discussed in [109]

3.2.3 Resolving branes

We can compute electric-field back reactions on the branes, if we resolve them. This is significant, since it gives us the possibility to address the UV sensitivity in the scenarios where branes source 3-forms. We will consider 3-forms sourced by brane in the form of a soliton, as in [108]. The topological charge of soliton serves as Nöther's charge of the 3-form. In this setup, we can consider back reaction from 3-forms on the brane. We will check the sensitivity of the resolution on the back reaction.

For resolving the brane as a soliton of heavy axion a, we choose the Lagrangian in the following form [108],

$$\mathcal{L} = -\frac{1}{48}F^2 - \frac{q}{2\pi}\partial_\mu a\epsilon_{\mu\alpha\beta\gamma}C_{\alpha\beta\gamma} + \frac{1}{2}f_a^2(\partial a)^2 - V(a), \tag{3.77}$$

the field a is a non-canonically normalized axion, with decay constant f_a. The potential $V(a)$ is chosen in the form

$$V(a) = V_0[1 - cos(a)]\,, \tag{3.78}$$

where V_0 is the constant with the dimension of the energy density. This choice is motivated by (3.1). The parameter q is a coupling between the 3-form and the axion. The connection between the parameters above and (3.70) is straightforward. This is a typical sine-Gordon potential (2.3.4), but our conclusion holds for an arbitrary periodic potential.

Equation of motions for the field C will give

$$-\partial_\mu F_{\mu\alpha\beta\gamma} = J_{\alpha\beta\gamma}, \tag{3.79}$$

where J is a conserved current,

$$J_{\alpha\beta\gamma} = -\frac{q}{2\pi}\epsilon_{\mu\alpha\beta\gamma}\partial_\mu a\,. \tag{3.80}$$

It is conserved due to the topological nature (2.20).

For an exact treatment, we should solve both the equations for C and a simultaneously. However, since the axion is heavy, we will first ignore the back reaction, solve it and then iterate. The field a has a solitonic (2.111) solution,

$$a(z) = 4\tan^{-1} e^{\frac{\sqrt{V_0}(z-z_0)}{f_a}} + 2\pi N, \tag{3.81}$$

which depends on a single coordinate z. This soliton interpolates between the nearest neighbouring minima of the periodic potential,

$$\begin{aligned} a &= 2\pi N, & z &= -\infty \\ a &= 2\pi(N+1), & z &= \infty, \end{aligned} \tag{3.82}$$

where N is an integer. The parameter z_0 is a zero mode (see chapter (2.2.2)) of the soliton. The configuration (3.81) is an exact solution of the equation in the limit $q = 0$, when the back reaction from the 3-form can be ignored.

Substituting the solution (3.81) into the topological current, we get an external source for C. This approximation is valid if we take $\frac{\sqrt{2V_0}}{f_a} \to \infty$ and keep q to be small. The soliton in the approximation is very heavy. So, it acts like an external source for the 3-form, while experiencing no back reaction from it. So, basically, we have

$$a'(z) = 2\pi\delta(z - z_0)\,, \tag{3.83}$$

where $'$ denotes a derivative w.r.t z and the sine-Gordon soliton effectively becomes a delta-function (point-like) source.

The equation of motion of the 3-form (3.79) obviously is gauge redundant. We choose the Coulomb gauge during the discussion, but other gauges obviously give the same result (see the section 1.2 and discussion therein). Using the substitution $C_{\alpha\beta\gamma} = \epsilon_{\mu\alpha\beta\gamma}C_\mu$ and considering (3.83), the equation of motion reduces to

$$\epsilon_{z\alpha\beta\gamma}C''_z = J_{\alpha\beta\gamma} = -q\delta(z - z_0)\epsilon_{z\alpha\beta\gamma}. \tag{3.84}$$

This equation is identical to an equation of an electrostatic field produced by a static point charge in one spatial dimension (see Schwinger model, e.g., [39]).

The solution for C_z is

$$C_z = -\frac{1}{2}q|z - z_0| - 4E_0(z - z_0) + c_2\,, \tag{3.85}$$

where E_0 and c_2 are integration constants. The electric field corresponding to this solution is

$$E = \frac{1}{8} q \, sign(z - z_0) + E_0. \tag{3.86}$$

We see that, the topological charge of the soliton acts as an electric Nöther's charge for the 3-form.

3.2.4 Back Reaction on the branes

During the previous section, we considered the limit, where the soliton is ultra-heavy. In this chapter, we want to compute a back reaction from the electric field on the soliton petrubatively. Intuitively, back reaction should happen, the electric field is different for different sides of the brane. We expect acceleration of the kink. So, the kink should be pushed to the infinity by the electric field.

The only light degree of freedom which survives in the limit is the zero mode of the brane. The zero mode defines the position of the kink. To account the back reaction, we should give time dependence to it,

$$z_0 \to z_0 + \delta z_0(t) \, . \tag{3.87}$$

Inserting this ansatz back into the Lagrangian together with three-form solution, we get an effective Lagrangian for the fluctuation

$$\mathcal{L} = -\frac{1}{48} F^2 - \frac{q}{2\pi} a E + \frac{1}{2} f_a^2 a'^2 \dot{\delta z_0}^2 - \frac{1}{2} f_a^2 a'^2 - V(a). \tag{3.88}$$

In the above Lagrangian, first and last two terms are irrelevant, since one is constant and the other two give a total derivative, which is true due to their topological nature. So, we are left with a simpler Lagrangian,

$$L = \int dz \mathcal{L} = \int dz \left(\frac{1}{2} f_a^2 a'^2 \dot{\delta z_0}^2 - \frac{q}{8\pi} a' C_z \right), \tag{3.89}$$

We should do integration over coordinate z carefully. In the first term we should resolve the square of the delta-function, and explicitly consider the profile (3.81) of the sine-Gordon soliton and take the integral over it. In the second term, we can perform the integration using the delta function approximation (3.83). After the proper normalization, we get the following Lagrangian,

$$L = \frac{1}{2} M \dot{\delta z_0}^2 + \frac{1}{f_a^2} \left(\frac{1}{8} q^2 |\delta z_0| + q E_0 \delta z_0 \right), \tag{3.90}$$

This system describes effectively a point charge with mass $M = 8\frac{\sqrt{V_0}}{f_a}$ and a coordinate δz_0 in one dimensional external electric field [110]. Let us assume for a moment $\delta z_0 > 0$, the solution has the following form

$$\delta z_0(t) = \frac{1}{2 M f_a^2} \left(\frac{1}{8} q^2 + q E_0 \right) t^2 + vt + \delta z_0(0) \, . \tag{3.91}$$

where v is initial velocity. The solution tells us that soliton moves accelerated. Let us choose the frame, where the initial coordinate and the initial velocity are zeros (this choice does not affect our discussion).

We should identify the validity of time-scale of the above approximation. The time-scale after which the effective theory, where we treat the soliton without alteration of the internal structure, breaks down.

In the most conservative approach, the variation of the collective coordinate should not exceed the width of the soliton. So, in this case, the description is valid for sure. This gives the condition,

$$\frac{f_a}{\sqrt{V_0}} \gtrsim |\delta z_0|, \tag{3.92}$$

and applying to the solution (3.91) implies,

$$\frac{f_a}{\sqrt{V_0}} \gtrsim \frac{1}{16\sqrt{V_0} f_a} \left| \frac{1}{8} q^2 + qE_0 \right| t^2 \,. \tag{3.93}$$

This leads us to the following upper bound on the validity of the time-scale,

$$t \lesssim \frac{4f_a}{|q|\sqrt{\left|\frac{1}{8} + \frac{E_0}{q}\right|}} \equiv t_* \,. \tag{3.94}$$

After this time the internal structure of the soliton may be affected by the action of the electric field. Nevertheless, for a long-distance observer the description of the soliton, as of a point particle accelerated by a constant electric field, is still valid. The full relativistic solution for such a particle is well-known [110].

The above scenario can be used in the attractor mechanism of solving the strong-CP problem [99]. We leave this sidenote here, and in the next section we discuss the attractor in a different setup [16, 108]. Since the mechanism is the same, we will avoid the repetition. The important point is UV-insensitivity, which will be discussed now.

In the attractor mechanism, the brane charge q, which controls coupling with the massless field-C should effectively decrease during the process. This requires an effective dependence on the electric field $q(E)$. At the end of the process it should be zero at the critical value E_*. The electric field among the neighboring vacua differs, and we have

$$\Delta E \propto q(E) \,, \tag{3.95}$$

So, at each step the charge decays. As soon as we approach E_*, it will vanish. Obviously, since the process is continuous, it could require a long time to reach E_*. The attractor should not be jeopardized by UV physics, meaning the singularity in the quantity of vacua at the attractor point E_* should be trusted without knowing the internal structure of the brane. During the next sections, we will explicitly verify this by evaluating the scaling of the critical time t_* near the attractor point. Since at the attractor point the electric field is finite E_*, whereas the brane charge $q(E_*)$ vanishes, it is clear from (3.94) that $t_* \to \infty$. So, the closer we are to the attractor vacuum, the longer it takes for the thin-wall approximation to break down.

3.2.5 Effect of Particle Creation

The 3-forms/axions can be used to solve the Hierarchy problem [16, 108]. This problem is a naturalness problem, the inexplicable smallness of the Higgs mass compare to the Planck scale, an ultimate cutoff of the theory. Usually, the Hierarchy problem is approached by low energy supper-symmetry (see, e.g., [39, 40]), or low scale quantum gravity [91, 111], which predicts existence of new physics near the electroweak scale. The mechanism [16, 108] works differently, it does not need new physics at the low energy. The Higgs mass depends effectively on the electric field E. Then the attractor relaxes it to the critical value E_*. Obviously, this scenario is very close to the one, we described in the previous chapter.

Let us consider the brane background and consider the Higgs mass, which depends on the 3-form. Hence, the higgs mass is effectively dependent on the space-time coordinates. The part of the Lagrangian where this information is encoded has the following form

$$\mathcal{L} = \frac{1}{2}(\partial h)^2 - \frac{1}{2}h^2\left(m_h^2 + \frac{F^2}{48M_f^2}\right), \tag{3.96}$$

where h is a Higgs field, m_h is its "bare" mass and M_f is some fundamental scale, which will be assumed to be large.

Considering the solution for electric field (3.86), we get effectively a mass term for Higgs, which depends on a position on z-axis. Basically, since the electric field changes across the brane, so does the Higgs mass. Correspondingly, we have two different masses M_- and M_+,

$$M_\pm^2 = m_h^2 - \frac{1}{128}\frac{q^2}{M_f^2} - \frac{E_0^2}{2M_f^2} \pm \frac{1}{8}\frac{E_0 q}{M_f^2}, \tag{3.97}$$

on the left and right sides of the brane respectively.

We assume that the constant part of the electric field is positive, for the opposite case we could just change the labelling. Let us take some initial time $t = 0$, for which on both sides of the wall the Higgs field is in corresponding vacuum states. In other words, no Higgs particles are excited.

Now, since the wall is moving accelerated (3.91), the electric field changes in time, and the vacuum of the Higgs field changes correspondingly. This leads to a particle creation.

To compute the rate of particle-creation, let us start with a region in the vacuum which corresponds to M_+. After the brane passes the region, the mass decreases to M_-.

The number operator of M_--mass particles is defined in the following way

$$N = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega_-}a_-^+(\vec{p})a_-(\vec{p}), \tag{3.98}$$

where $\omega_\pm^2 = p^2 + M_\pm^2$, and the subscript on the ladder operators has the same meaning. Performing Bogolyubov transformation (3.111) and averaging in the vacuum $|0_+\rangle$ (which is the vacuum in this case), we get the following expression,

$$N = SL\frac{1}{4}\int d^3p\frac{\omega_+}{\omega_-}\left(\frac{\omega_-}{\omega_+} - 1\right)^2, \tag{3.99}$$

where S is the surface area of the brane and L is the distance travelled by it in z direction. Considering

$$\omega_+^2 - \omega_-^2 = \frac{1}{16}\frac{E_0 q}{M_f^2}\,, \tag{3.100}$$

for $M_f^2 \gg \frac{E_0 q}{M_-^2}$ we get

$$\omega_+ \approx \omega_- + \frac{1}{32}\frac{E_0 q}{M_f^2 \omega_-}\,. \tag{3.101}$$

At the lowest order, the produced particle number per unit surface is given by

$$\frac{N}{S} = L\frac{1}{4096}\frac{E_0^2 q^2}{M_f^4}\int d^3p \frac{1}{\omega_-^4}\,. \tag{3.102}$$

After integration we obtain,

$$\frac{N}{S} = L\frac{\pi^2}{4096}\frac{E_0^2 q^2}{M_f^4 M_-}\,. \tag{3.103}$$

Substituting the distance travelled by the domain wall in time t, we get the final result

$$\frac{N}{S} = \left(\frac{1}{16\sqrt{V_0} f_a}\left(\frac{1}{8}q^2 + qE_0\right)t^2 + vt\right)\frac{\pi^2}{4096}\frac{E_0^2 q^2}{M_f^4 M_-}. \tag{3.104}$$

The above expression gives us the amount of Higgs particles created by a passing-by domain wall.

However, we should remember that we can trust the above analysis until the time (3.94). During this time, we get the following number of particles created per unit surface

$$\frac{N}{S} = \frac{f_a}{\sqrt{V_0}}\frac{\pi^2}{4096}\frac{E_0^2 q^2}{M_f^4 M_-}\,. \tag{3.105}$$

Far from the attractor, where the charge q is large, the effect of particle creation can be significant and affect the dynamics of the system. However, not surprisingly, since the effect is proportional to q^2, it is vanishingly small near the attractor point. Again, the attractor behavior is practically unaffected by the high-energy effects.

So, we conclude, that the physics depend on 3-forms is completely insensitive to heavy physics.

3.2.6 * On Bogolyubov Transformations

In this section, we review supplement material for the previous one. We will derive the Bogolyubov transformations in a simple case. Let us consider a free scalar field ϕ, in Schrödinger picture,

$$\phi(\vec{x}) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega(p,t)}(a_t(\vec{p})e^{i\vec{p}\vec{x}} + a_t^+(-\vec{p})e^{i\vec{p}\vec{x}})\,, \tag{3.106}$$

where $\omega(\vec{p},t)$ is the time-dependent frequency. Therefore, ladder operators depend on time too (despite being described in Schrödinger picture). The corresponding canonical momenta have the following form,

$$\pi(\vec{x}) = \int \frac{d^3p}{(2\pi)^3}\frac{-i}{2}(a_t(\vec{p})e^{i\vec{p}\vec{x}} - a_t^+(-\vec{p})e^{i\vec{p}\vec{x}})\,. \tag{3.107}$$

The important point here is that the field and the canonical momenta do not depend on time explicitly. Thus, for any given time we can write

$$\begin{aligned}\frac{1}{2\omega(p,t)}\left(a(\vec{p})_t + a^+(-\vec{p})_t\right) &= \int d^3x e^{-i\vec{p}\vec{x}}\phi(\vec{x})\\ \frac{-i}{2}\left(a(\vec{p})_t - a^+(-\vec{p})_t\right) &= \int d^3x e^{-i\vec{p}\vec{x}}\pi(\vec{x})\,.\end{aligned} \tag{3.108}$$

From this equation immediately follows the following transformation

$$\begin{aligned}a(\vec{p}) &= \frac{1}{2}\left(\frac{\omega}{\omega_0}+1\right)a_0(\vec{p}) + \frac{1}{2}\left(\frac{\omega}{\omega_0}-1\right)a_0^+(-\vec{p})\\ a(-\vec{p})^+ &= \frac{1}{2}\left(\frac{\omega}{\omega_0}+1\right)a_0(-\vec{p})^+ + \frac{1}{2}\left(\frac{\omega}{\omega_0}-1\right)a_0(\vec{p}),\end{aligned} \tag{3.109}$$

where subscript 0 means that the quantities are evaluated at $t_0 = 0$. The quantities without subscript are evaluated at the time, t and all frequencies are evaluated for momentum p. The number operator at time t has the following form

$$N = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega}a^+(\vec{p})a(\vec{p}). \tag{3.110}$$

Rewriting it in terms of operators at time $t_0 = 0$, we get

$$\begin{aligned}N = \frac{1}{4}\int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega}&\left[\left(\frac{\omega}{\omega_0}+1\right)a_0(\vec{p})^+ + \left(\frac{\omega}{\omega_0}-1\right)a_0(-\vec{p})\right]\\ &\left[\left(\frac{\omega}{\omega_0}-1\right)a_0(-\vec{p})^+ + \left(\frac{\omega}{\omega_0}+1\right)a_0(\vec{p})\right],\end{aligned} \tag{3.111}$$

showing that in vacuum, corresponding $t_0 = 0$, after certain time particles will be created.

Chapter 4

Coherent States in Gauge redundant systems

The Coherent states represent mostly classical configurations in quantum descriptions. In QFT, they were introduced by Glauber [112] to describe almost classical electro-magnetic fields. These objects also play an important role to get rid of the infrared divergences, accompanying asymptotic states with the soft photons [113–118]. The coherent states tend to be useful in formulation of the quantum picture of solitons [119] too.
It appears that the S-matrix formulation of the quantum gravity forbids de-Sitter to be a vacuum [18, 21]. Hence, it should be a state built on top of the good vacuum. Such description has been applied to de-Sitter and other cosmological space-times in [7, 17, 120–124]. It was observed, that de-Sitter can be described as a coherent state of the off-shell gravitons [17, 121, 122]. It also appears, that in this case eternal inflation is not possible, due to the quantum breaking [17].
So, the coherent states play an important role, for instance, using them we can bring maximally classical objects in quantum physics. In some cases, they are the only option to describe observed states, e.g., de-Sitter cosmology.
We will consider coherent states in the systems with the gauge redundancy. We will formulate a framework, which gives possibility to build such states and compute evolution. Most importantly, they can extract some interesting information from the gravity.
This chapter follows our article [3].

4.1 Coherent states in QED

The Quantum-electro-dynamics (QED) is a perfect laboratory to study coherent states in gauge redundant systems. The theory has a well-defined classical limit (which we observe every day), and important results can be computed via consideration of classical gauge field interaction with quantum particles. For example, the purely classical electric field can decay into pairs of quanta, via the Schwinger effect [125]. So, QED is a perfect candidate to build and study coherent states.

4.1.1 Coherent States with Global Charge

In theories, where we do not have gauge redundancy, building the coherent states is straightforward and not constraining. First, the theory should be written in Hamiltonian form. For a definiteness, let us consider a free complex scalar field Φ. The theory is described by the following Hamiltonian,

$$\mathcal{H} = \Pi^*\Pi + |\nabla\Phi|^2 + m^2|\Phi|^2, \tag{4.1}$$

where Π is the canonical momenta, corresponding to Φ. In the Hamiltonian formulation, by using the field displacement operator, we can build a coherent state in terms of the canonical fields [126, 127]. So, a particular example of such construction is

$$|C\rangle = e^{-i\int d^3x\left(\Phi_c\hat{\Pi}-\Pi_c\hat{\Phi}+h.c.\right)}|\Omega\rangle\,, \tag{4.2}$$

where $\Phi_c(x)$ and $\Pi_c(x)$ are c-number functions of spatial coordinates, $\hat{\Phi}$ and $\hat{\Pi}$ stand for the field and its canonical momentum operators, respectively. The state $|\Omega\rangle$ is the quantum vacuum of the theory. During the chapter, we will use hat for the operators and subscript c for the corresponding c-numbers. We may skip both hat and subscript in some cases if the context is completely clear. The above state satisfies the following initial conditions by construction,

$$\langle C|\hat{\Phi}|C\rangle(t=0) = \Phi_c(x)\,, \tag{4.3}$$
$$\langle C|\hat{\Pi}|C\rangle(t=0) = \Pi_c(x)\,. \tag{4.4}$$

The Above results can be derived using the canonical commutation relations.

The theory (4.1) admits $U(1)$ symmetry. Hence, it has a corresponding conserved Nöther's (see the section 1.1) charge Q, given by

$$Q = i\int d^3x\,(\Pi\Phi - \Pi^*\Phi^*) \tag{4.5}$$

The expectation value of the charge in the state is conserved and given by

$$\langle C|\hat{Q}|C\rangle = i\int d^3x\,(\Pi_c\Phi_c - \Pi_c^*\Phi_c^*) \equiv Q_c\,. \tag{4.6}$$

It is crucial that the coherent states are not the eigenstates of the global charge operator, and hence the above expression is just an average. So, the charge in this case is conserved as an average quantity. Considering that the vacuum is not charged, acting by the charge on the coherent state, we get

$$\begin{aligned}\hat{Q}|C\rangle = Q_c|C\rangle + e^{-i\int d^3x\left(\Phi_c\hat{\Pi}-\Pi_c\hat{\Phi}+h.c.\right)} \times \\ \times i\int d^3x'\left(\Phi_c\hat{\Pi}+\Pi_c\hat{\Phi}-h.c.\right)|\Omega\rangle \,. \end{aligned} \tag{4.7}$$

The presence of the second term is the reason for the state not being an eigenstate of the charge. This result will have a major impact on the procedure of building physical coherent states in gauge theories. We want to use BRST (chapter 1.2) construction, which relies on the global BRST and ghost number charges and eigenstates.

4.1.2 BRST formulation of QED

Building coherent states in QED require a self-consistent canonical quantization. In the case of QED, in principle different quantization procedures can be used (see e.g., [22]), but we require a clear and consistent physicality condition for the states. So, the natural candidate is the BRST quantization (reviewed in section 1.2). We will still do a very short review for the case of scalar electrodynamics.

The BRST invariant formulation of scalar electrodynamics has the following Lagrangian

$$\mathcal{L} = -\frac{1}{4}\hat{F}_{\mu\nu}^2 + |D_\mu\hat{\Phi}|^2 - m^2|\hat{\Phi}|^2 - \partial_\mu\hat{B}\hat{A}_\mu + \frac{1}{2}\xi\hat{B}^2 + \partial_\mu\hat{\bar{c}}\partial_\mu\hat{c}\,,$$

where $\hat{F}$ is a field strength tensor, D is a covariant derivative with the charge g, $\hat{c}, \hat{\bar{c}}$ are Fadeev-Popov ghosts and $\hat{B}$ denotes the gauge fixing auxiliary scalar. The above theory is invariant under BRST transformations,

$$\begin{aligned}&\delta\hat{A}_\mu = \theta\partial_\mu\hat{c}\,, \quad \delta\hat{\bar{c}} = \theta\hat{B}\,, \quad \delta\hat{c} = \delta\hat{B} = 0\,, \\ &\delta\Phi = ig\theta\hat{c}\hat{\Phi}\,, \quad \delta\Phi^\dagger = -ig\theta\hat{c}\hat{\Phi}^\dagger\,,\end{aligned} \tag{4.8}$$

where θ is a Grassmann parameter of the transformation δ. The conserved BRST charge associated with this symmetry takes the following form,

$$\hat{Q}_B = \int d^3x\left[\hat{c}\left(g\hat{\rho} - \partial_j\hat{E}_j\right) + \hat{B}\hat{\Pi}_{\bar{c}} + \partial_j\left(\hat{c}\hat{E}_j\right)\right]\,; \tag{4.9}$$

where $\hat{\rho} \equiv i(\hat{\Phi}\hat{\Pi} - \hat{\Pi}^\dagger\hat{\Phi}^\dagger)$ is the $U(1)$ charge density, and $\hat{E}_j \equiv \hat{F}_{0j}$ is the electric field operator. The momentum $\hat{\Pi}_{\bar{c}}$ denotes the conjugate momentum of the corresponding ghost field.

From the charge structure, we can deduce, that it enforces the gauss law constraint on physical states. The physical charges should be dressed as it is discussed for instance in [128–130]. So, the BRST framework automatically fulfils the intuitive picture, the charge should carry an electric field.

From the Lagrangian (4.8) the usual (same as in the case of the Lagrangian without the BRST parts) equations of motion follow for the matter field $\hat{\Phi}$. The ghost fields decouple and the Maxwell field satisfies the following equation,

$$\partial_\mu \hat{F}_{\mu\nu} + \hat{J}_\nu - \partial_\nu \hat{B} = 0, \tag{4.10}$$

where J is the current correspond to field Φ. From (4.8) it is clear that the last term is ImQ (1.39). Therefore, between the physical states, we get usual Maxwell equations [28]. The result has an implication for coherent states dynamics [6]. We also note that $\partial \hat{B}$ looks like the conserved current. Hence, using the Maxwell equations, it satisfies a free scalar equation. So, the above construction maintains gauge redundancy. That is the reason of seeing gauge redundancy before the BRST invariance.

4.1.3 Coherent states in QED

The coherent state made of the gauge fields only can be constructed as

$$|A\rangle = e^{-i\hat{f}_A}|\Omega\rangle\,, \tag{4.11}$$

with

$$\hat{f}_A \equiv \int d^3x \left(A_j^c \hat{E}_j - E_j^c \hat{A}_j + A_0^c \hat{B} - B^c \hat{A}_0 \right)\,. \tag{4.12}$$

The c-number objects specify the initial field-configurations. We can compute the initial expectation values for one-point functions easily. Naively, we get the result,

$$\langle A|\hat{A}_\mu|A\rangle = A_\mu^c\,, \quad \langle A|\hat{E}_j|A\rangle = E_j^c\,, \quad \langle A|\hat{B}|A\rangle = B^c\,; \tag{4.13}$$

The other fields have vanishing initial expectation values.

We should act on the state by the BRST charge to determine the physicality condition. Considering the vacuum annihilation by the charge, we are left with the following condition

$$\int d^3x \left(-\hat{c}\partial_j E_j^c + \hat{\Pi}_{\bar{c}} B^c + \partial_j \left(\hat{c} E_j^c \right) \right) |\Omega\rangle = 0\,. \tag{4.14}$$

To derive the expression we used properties of $e^{-i\hat{f}_A}$ (displacement) operator. We stress that the expression is exact. Since, the charge $\hat{Q}_B$ is conserved, we can always use equal time commutators in such computations. The last (boundary) term of (4.14) vanishes not only for localized electric field configurations, but also for more general ones. This point will be discussed in details in the following section. The above condition (4.14) is satisfied for (in the case of localized solution)

$$\partial_j E_j^c = 0\,, \qquad B^c = 0\,. \tag{4.15}$$

This is equivalent to satisfy the chargeless Gauss' law. Naively, introduction of the non-trivial B^c (with the special construction, relating it to the electric field) seems possible, but we should consider the following identity

$$\langle A|\, \hat{B}\, |A\rangle = \langle A|\, \{Q, \hat{\bar{c}}\}\, |A\rangle\,, \tag{4.16}$$

which is true due to (4.8). So, BRST requires vanishing expectation value of $\hat{B}$ [6].

Now, let us move to the matter fields. As we discussed, they should be dressed with appropriate electric fields. A naive attempt would be to dress the state (4.2) with the coherent photons. Such state has the following form

$$|A\rangle \otimes |C\rangle \,. \tag{4.17}$$

This assumption does not work (except in the very specific limit [6]), which is evident from the equation

$$\hat{Q}_B\{|A\rangle \otimes |C\rangle\} = \hat{Q}_B e^{-i\int d^3x \left(A^c_j \hat{E}_j - E^c_j \hat{A}_j + A^c_0 \hat{B} - B^c \hat{A}_0\right)} \\ \times e^{-i\int d^3x \left(\Phi_c \hat{\Pi} - \Pi_c \hat{\Phi} + h.c.\right)} |\Omega\rangle \neq 0 \,, \tag{4.18}$$

The inability to dress the scalar coherent state with a coherent electromagnetic configuration is connected to the fact that the former is not an eigenstate of $U(1)$ charge; see e.g., (4.7). So, such constructions cannot be made BRST invariant by dressing them with coherent photons. We should construct the gauge invariant operators instead (operators with a zero BRST charge).

The idea of defining gauge invariant matter degrees of freedom goes back to Dirac [128] (see also [129]). The treatment suggests compensation of phase rotation of the matter operators and gauge field shift under gauge transformations in the following form

$$\hat{\Phi}_g = \hat{\Phi} \cdot \exp\left(-ig\frac{1}{\nabla^2}\partial_j \hat{A}_j\right) \,, \tag{4.19}$$

$$\hat{\Pi}_g = \hat{\Pi} \cdot \exp\left(+ig\frac{1}{\nabla^2}\partial_j \hat{A}_j\right) \,. \tag{4.20}$$

The subscript g indicates the gauge invariance. Similar non-local operators $\exp(-ig\frac{1}{\Box}\partial_\mu A^\mu)$ were used to maintain gauge invariance in the case of anomalous symmetries [131].

The redefined operators satisfy canonical commutation relations

$$[\hat{\Phi}_g(t,x), \hat{\Pi}_g(t,y)] = i\delta^{(3)}(x-y) \,. \tag{4.21}$$

Moreover, it is straightforward that these quantities are BRST invariant

$$[\hat{Q}_B, \hat{\Phi}_g(x)] = [\hat{Q}_B, \hat{\Pi}_g(x)] = 0 \,. \tag{4.22}$$

Based on this observation, we can construct dressed coherent states for matter fields to the analogy of (4.2), but with the dressed operators $\hat{\Phi}_g$ and $\hat{\Pi}_g$. The coherent state will have the following form,

$$|C_g\rangle = e^{-i\int d^3x \left(\Phi_c \hat{\Pi}_g - \Pi_c \hat{\Phi}_g + h.c.\right)} |\Omega\rangle \,. \tag{4.23}$$

The state satisfies the physicality conditions for arbitrary $\Phi_c(x)$ and $\Pi_c(x)$. To establish the connection to the classical field configurations, let us consider their expectation values at the initial time

$$\langle C_g|\hat{\Phi}_g|C_g\rangle(t=0) = \Phi_c\,, \quad \langle C_g|\hat{\Pi}_g|C_g\rangle(t=0) = \Pi_c\,, \tag{4.24}$$

$$\partial_j\langle C_g|\hat{E}_j|C_g\rangle(t=0) = ig\left(\Phi_c\Pi_c - \Pi_c^*\Phi_c^*\right)\,, \tag{4.25}$$

$$\langle C_g|\hat{A}_\mu|C_g\rangle(t=0) = 0\,. \tag{4.26}$$

The state can be further supplemented with the coherent photon configuration similar to the gauge field coherent states discussed above. Putting all the ingredients together, we get a physical coherent state

$$|C_g, A\rangle = e^{-i\int d^3x\left(\Phi_c\hat{\Pi}_g - \Pi_c\hat{\Phi}_g + h.c.\right)} \times e^{-i\int d^3x\left(A_j^c\hat{E}_j - E_j^c\hat{A}_j + A_0^c\hat{B}\right)}|\Omega\rangle\,. \tag{4.27}$$

In the above expression, the order of the exponentials matters. To reproduce (4.24) we should not change the ordering, or we should do a long computation and redefine classical initial quantities. Physicality condition fixes E_j^c as a transverse, hence (4.25) is not altered. The only modification is the generation of the initial time expectation value for $\hat{A}_\mu$.

The dynamics of the system can be obtained from Heisenberg equations of motions, which can be obtained from Hamiltonian equations. The Hamiltonian of the system reads,

$$\begin{aligned}\hat{H} = \int d^3x \Big[&\frac{1}{2}\hat{E}_j^2 + \frac{1}{4}\hat{F}_{ij}^2 + |\hat{\Pi}|^2 + |D_j\hat{\Phi}|^2 + m^2|\hat{\Phi}|^2 \\ &+\hat{B}\partial_j\hat{A}_j - \frac{\xi}{2}\hat{B}^2 + \partial_j\left(\hat{A}_0\hat{E}_j - \hat{B}\hat{A}_j\right) \\ &+\hat{A}_0\left(-\partial_j\hat{E}_j + ig\left(\hat{\Phi}\hat{\Pi} - \hat{\Phi}^\dagger\hat{\Pi}^\dagger\right)\right) \\ &+\hat{\Pi}_c\hat{\Pi}_{\bar{c}} + \partial_j\hat{\bar{c}}\,\partial_j\hat{c}\Big]\,.\end{aligned} \tag{4.28}$$

The Heisenberg's operator equations for the gauge sector give us the following equations for 1-point functions

$$\partial_0\langle\hat{A}_0\rangle = \partial_j\langle\hat{A}_j\rangle, \tag{4.29}$$

$$\partial_0\langle\hat{A}_j\rangle = \langle\hat{E}_j\rangle + \partial_j\langle\hat{A}_0\rangle\,, \tag{4.30}$$

$$\partial_0\langle\hat{E}_j\rangle - \partial_i\langle\hat{F}_{ij}\rangle = ig\langle\hat{\Phi}^\dagger\hat{D}_j\hat{\Phi} - h.c.\rangle\,, \tag{4.31}$$

$$\partial_j\langle\hat{E}_j\rangle = ig\langle\hat{\Phi}\hat{\Pi} - \hat{\Pi}^\dagger\hat{\Phi}^\dagger\rangle\,, \tag{4.32}$$

with $\langle\ldots\rangle$s are denoted the expectation value of the Heisenberg equation of motions in the coherent state (4.27). As a note, except the first one, all the other equations restore the gauge invariant form of the Maxwell equations. The first equation gives us a sign that we should work in the framework where a particular gauge is fixed. Meaning, when we use the classical Maxwell equations, we always fix the gauge. The above equation tells us the gauge we should work in. We can always alter gauge fixing, and we get the different gauge fixing condition. Hence, at the classical level, we see manifestation of the gauge invariance. The terms of (4.31) and (4.32) evaluated over a coherent state, will contain both classical and quantum non-linearities. The quantum contributions can be quantified by

$$S_j \equiv i\langle \hat{\Phi}^\dagger \hat{D}_j \hat{\Phi} - (\hat{D}_j \hat{\Phi})^\dagger \hat{\Phi} \rangle - \bar{J}_j \,, \tag{4.33}$$

$$S_0 \equiv i\langle \hat{\Phi}\hat{\Pi} - \hat{\Pi}^\dagger \hat{\Phi}^\dagger \rangle - \bar{\rho} \,, \tag{4.34}$$

where $\bar{J}_\mu$ is defined as the current constructed out of 1-point functions, i.e.

$$(-i)\bar{J}_j \equiv \langle \hat{\Phi}^\dagger \rangle \left(\partial_j \langle \hat{\Phi} \rangle - ig \langle A_j \rangle \langle \hat{\Phi} \rangle \right) - h.c. \,, \tag{4.35}$$

$$(-i)\bar{\rho} \equiv \langle \hat{\Phi} \rangle \langle \hat{\Pi} \rangle - \langle \hat{\Pi}^\dagger \rangle \langle \hat{\Phi}^\dagger \rangle \,. \tag{4.36}$$

As a result, (4.31) and (4.32) become

$$\partial_0 \langle \hat{E}_j \rangle - \partial_i \langle \hat{F}_{ij} \rangle = g\bar{J}_j + gS_j \,, \tag{4.37}$$

$$\partial_j \langle \hat{E}_j \rangle = g\bar{\rho} + gS_0 \,, \tag{4.38}$$

Without S_0 and S_j, the above are classical equations of motions for the theory. The computation of these quantum terms requires the knowledge of 2-point and 3-point functions. However, they can be evaluated explicitly following [127] in the coherent state up to a desirable orders in $\hbar$ and g. The explicit form of the equation for the scalar field is useful to discuss the particle production. It will be provided during the next chapters.

4.1.4 Classical Charges

In this section, we discuss the possibility of having a fundamentally classical sources in QED. This is an electromagnetic analogue of the cosmological constant in the gravity [132]. To add classical sources to QED, we should add the following Lagrangian,

$$\Delta\mathcal{L} = -\hat{A}_\mu J^{\rm cl}_\mu \,, \quad \text{with} \quad \partial_\mu J^{\rm cl}_\mu = 0 \,; \tag{4.39}$$

where the current $J^{\rm cl}_\mu$ is a predetermined c-number function. Its presence does not alter the BRST transformation properties (4.8). In this case, however, the Lagrangian density is invariant up to a total derivative

$$\delta\left(\mathcal{L} + \Delta\mathcal{L}\right) = -\partial_\mu \left(\theta \hat{c} J^{\rm cl}_\mu \right) \,. \tag{4.40}$$

As a result the Nöether's charge needs to be modified correspondingly, giving us

$$\hat{Q}^J_B = \hat{Q}_B + \int d^3x \; \hat{c} J^{\rm cl}_0 \,, \tag{4.41}$$

with $\hat{Q}_B$ denoting the BRST charge in the absence of the classical source (4.9).

In the limit, where quantum effects are completely neglected, $J_{\rm cl}$ will source a classical electromagnetic field configuration. But since the quantum effects will enter slowly, its presence could start particle production, i.e., back-reaction. To treat the corresponding state as a vacuum, the only legitimate case is to ignore completely the back-reaction [18]. In the gravity, similar considerations have profound consequences. For instance, validity of the S-matrix description forces us to consider the de-Sitter as a coherent state build up on top of Minkowski. In our case, similarly, we should consider to build up the state around the vacuum of the theory without a classical source (4.28). The coherent state of the classical charge in the literature was discussed (see, e.g., [133, 134]) in different contexts.

The total BRST charge $\hat{Q}_B^J$ commutes with the total Hamiltonian

$$\hat{H}_J = \hat{H} + \int d^3x \hat{A}_\mu J_\mu^{\rm cl}. \tag{4.42}$$

So to construct a consistent coherent state of electromagnetic field, we should satisfy

$$\hat{Q}_B^J |J\rangle = 0\,. \tag{4.43}$$

Following the above arguments, we should use the vacuum $|\Omega\rangle$ of $\hat{H}$ as a basis. Then the state,

$$|J\rangle = e^{-i\int d^3x \left(A_j^c \hat{E}_j - E_j^c \hat{A}_j + A_0^c \hat{B}\right)} |\Omega\rangle \tag{4.44}$$

constructed analogous to the pure photon state, satisfies the physicality condition (4.43) if

$$\int d^3x \left[\hat{c}\left(-\partial_j E_j^c + J_0^{\rm cl}\right) + \partial_j\left(\hat{c} E_j^c\right)\right] |\Omega\rangle = 0\,. \tag{4.45}$$

This requires satisfaction of the Gauss' law

$$\partial_j E_j^c = J_0^{\rm cl}\,, \tag{4.46}$$

as long as the boundary term of (4.45) vanishes. The latter is trivially satisfied by localized classical sources. Still, it can be applicable to more general cases. We will consider it in the next section.

It must be stressed that the vacuum $|\Omega\rangle$ is a ground state of $\hat{H}$ and annihilated by $\hat{Q}_B$, not by $\hat{Q}_B^J$. Consequently, the physicality condition satisfied by it is not preserved by the Hamiltonian flow generated by $\hat{H}_J$. So, in the theory of $\hat{H}_J$ the vacuum $|\Omega\rangle$, is not a physical state, nevertheless $|J\rangle$ is.

Next, we want to ask how good the state $|J\rangle$ is. First, the equations (4.31) and (4.32) are altered by the classical source. Second, we know that Schwinger pair creation may happen. So, the process of our interest is the pair production of $\Phi\Phi^\dagger$ and the back-reaction on 1-point function of the electric field. It appears that the pair will be created until the electric field is completely screened, or the energy capacity is empty. Further detailed discussion regarding this can be found in the next chapter.

4.1.5 Infinitely Homogeneous Classical Source

During the last chapter regarding QED, we discussed the homogeneous constant charge, similar to the cosmological constant in gravity [132].

We take the model from the previous chapter and choose the following current $J^{\rm cl}_\mu = \delta^\mu_0 \rho = \delta^\mu_0 const$. The coherent state in this case should satisfy the physicality condition (4.45). As we mentioned, the condition leads to the Gauss' law for the electric field E^c_j. For which the boundary term (4.45) should vanish. Now, let us look how it happens in this case. Naively, we may think that the boundary term is not vanishing.

The electric field in this case is linear in the coordinates to satisfy the Gauss' constraint. In this case, the field grows towards infinity rapidly. We can deduce that Schwinger pairs are created rapidly too, and the electric field is discharged. The solution should not be maintained in the quantum theory. Let us still consider a slightly regularized version of the above scenario, a spherical boundary, which is taken to infinity.

The linear solution of the E^c_j has the following form,

$$E^c_j = \frac{\rho}{3} x_j \,; \tag{4.47}$$

This quantity diverges in infinity. Nevertheless, the above-mentioned boundary term vanishes. Let us consider the ghost field, which can be decomposed into creation and annihilation operators

$$\hat{c}(x, t_0) = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2k}} \left(\hat{a}_k e^{i\vec{k}\cdot\vec{x}} + \hat{b}^\dagger_k e^{-i\vec{k}\cdot\vec{x}} \right) . \tag{4.48}$$

If we further consider that the vacuum is annihilated by $\hat{a}_k$, we are led to the following equivalence

$$\int d^3x \; \partial_j (\hat{c} E^c_j) |\Omega\rangle = 0 \;\Leftrightarrow\; \int d^3x \; \partial_j \left(e^{-i\vec{k}\cdot\vec{x}} x_j \right) = 0 \,. \tag{4.49}$$

The above expression simplifies to

$$\frac{\partial}{\partial k_j} \left(k_j \delta^{(3)}(k) \right) = 0 \,, \tag{4.50}$$

which holds as one of the properties of Dirac's delta-function. Therefore, the coherent state of the electromagnetic field, produced by the external source in question, is consistent with the physicality conditions of BRST quantization. The above treatment is not a complete proof of the physicality of the above state. For instance, there can be an issue regarding the unboundedness of the Hamiltonian from below. We shall set aside the issues and discuss the dynamical aspects of the above state.

To simplify the discussion, let us take $A^c_\mu = 0$ in (4.44), which corresponds to $\langle \hat{A}_\mu \rangle = 0$ at the initial time. We are left with the following coherent state

$$|\rho\rangle = e^{i \int d^3x \left(E^c_j \hat{A}_j \right)} |\Omega\rangle \,. \tag{4.51}$$

Next, we are interested in computing the leading quantum corrections to the dynamics of 1-point expectation values for the gauge sector, resulting from the quantum pair creation of scalar particles. As it was demonstrated explicitly in [127], the relevant dynamics follows from (4.29)-(4.32) (with inclusion of additional classical source) by retaining only the expectation values of bilinear operators in S. For the setup in question, the equations take the following form,

$$\partial_0\langle\hat{A}_0\rangle = \partial_j\langle\hat{A}_j\rangle, \tag{4.52}$$

$$\partial_0\langle\hat{E}_j\rangle - \partial_i\langle\hat{F}_{ij}\rangle = ig\langle\rho|\hat{\Phi}^\dagger D_j^{\rm cl}\hat{\Phi} - h.c.|\rho\rangle\,, \tag{4.53}$$

$$\partial_j\langle\hat{E}_j\rangle - \rho = ig\langle\rho|\hat{\Phi}\hat{\Pi} - h.c.|\rho\rangle\,, \tag{4.54}$$

where $D_\mu^{\rm cl} \equiv \partial_\mu - igA_\mu^{\rm cl}$. The terms on the right-hand side of (4.53) and (4.54) are quantum, since 2-point functions are of order $\hbar$ in the leading order. Therefore, we need to find the relevant correlators at the tree level. The required dynamics of $\hat{\Phi}$ follows from solving the following semi-classical equation

$$\left(D_\mu^{\rm cl}D_\mu^{\rm cl} + m^2\right)\hat{\Phi}(x,t) = 0, \tag{4.55}$$

with the classical electromagnetic configuration for our setup, given by $A_j^{\rm cl} = \frac{1}{3}\rho t x_j$, $A_0^{\rm cl} = \frac{1}{2}\rho t^2$. This equation needs to be solved with appropriate initial conditions, which for our coherent state corresponds to

$$\langle\rho|\hat{\Phi}(x,0)\hat{\Phi}(y,0)|\rho\rangle = \langle\Omega|\hat{\Phi}(x,0)\hat{\Phi}(y,0)|\Omega\rangle\,. \tag{4.56}$$

In other words, even though the dynamics is given by the background-dependent equation of motion, the initial conditions for the mode functions are the ones in charge-free vacuum. See [126, 127] for the discussion of this part and implications for perturbative dynamics.

The time dependence of the background field appearing in (4.55) induces particle creation. During the process, the 1-point function of the scalar remains zero, but the 2-point function will show a non-trivial dynamics. The one is the first quantum correction to the dynamics of the 1-point functions of the gauge sector, as per (4.53) and (4.54). So, we will get a back-reaction on the electromagnetic field. The Schwinger pair creation will screen the field. But the process has a threshold, the creation of the pair is suppressed beyond it. The radius reads,

$$R \sim \frac{m^2}{g\rho}\,. \tag{4.57}$$

So, we can read the boundary validity from the radius. A semi-classical approximation is not valid after it, and the uniform charge undergoes quantum breaking.

4.2 Coherent states in Linear gravity

In this section, we will study coherent states in linear gravity. We will introduce the cosmological constant at the linear level to study consistency of the de-Sitter space. We also add scalar matter pertubatively, which tends to be exact in a large-N limit (see the section 2.6). We also note that, de-Sitter space as a coherent state in the linear gravity was built [17, 121] based on a correspondence between the classical de Sitter metric of linear Einstein theory and the source-free solution of the massive Fierz-Pauli theory [132]. Here we will build the states for the massless gravity using the BRST treatment, as we did for QED in the previous sections. We will use up and down indices to avoid sign confusions.

4.2.1 BRST formulation of the Linear gravity

A free spin-2 field in Minkowski space has the following BRST invariant formulation,

$$\mathcal{L} = \frac{1}{2}(\partial_\alpha \hat{h}_{\mu\nu})^2 - \frac{1}{2}(\partial_\alpha \hat{h})^2 + \partial_\alpha \hat{h} \partial_\mu \hat{h}^{\mu\alpha} - \partial_\mu \hat{h}^{\mu\alpha} \partial_\nu \hat{h}^\nu_\alpha \\ -\partial_\mu \hat{B}_\nu \left(\hat{h}^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu}\hat{h} \right) + \frac{1}{2}\xi \hat{B}^2_\mu + \partial_\alpha \hat{\bar{C}}_\mu \partial^\alpha \hat{C}^\mu \,, \tag{4.58}$$

which is a linearized version of the framework developed in [135]. It follows from Einstein's theory of gravity in the limit of infinite Planck's mass. Here $\hat{h}_{\mu\nu}$, $\hat{C}_\mu$ and $\hat{B}_\mu$ are the massless spin-2 (graviton) field, Fadeev-Popov ghost vectors and an auxiliary vector, respectively, ξ is a gauge fixing parameter.

The theory is invariant under the BRST transformation,

$$\delta \hat{h}_{\mu\nu} = \theta \left(\partial_\mu \hat{C}_\nu + \partial_\nu \hat{C}_\mu \right) \,, \tag{4.59}$$

$$\delta \hat{\bar{C}}_\mu = \theta \hat{B}_\mu \,, \tag{4.60}$$

where θ is a Grassmann variable. The rest of the fields are transforming trivially.

Analogous to the previous section, the construction of states is performed within the canonical Hamiltonian framework. Following the ADM formalism [136], we supplement the Lagrangian (4.58) with boundary terms appropriate for removing the time-derivatives of $\hat{h}_{00}$ and $\hat{h}_{0j}$. So, the conjugate momentum of $\hat{h}_{ij}$ reduces to

$$\hat{\pi}_{ij} = \partial_0 h_{ij} - \delta_{ij}\partial_0 h_{kk} + 2\delta_{ij}\partial_k h_{k0} - \partial_i h_{j0} - \partial_j h_{i0} \,. \tag{4.61}$$

Its BRST transformation follows from (4.59) and is given by

$$\delta \hat{\pi}_{ij} = 2\theta \left(\nabla^2 \delta_{ij} - \partial_i \partial_j \right) \hat{C}_0 \,. \tag{4.62}$$

Similarly to the previous sections, we construct the physical states in a BRST invariant fashion. A pure-graviton coherent state, free of ghosts, can be built as

$$|h\rangle = e^{-i\int d^3x \left(h^c_{ij}\hat{\pi}_{ij} - \pi^c_{ij}\hat{h}_{ij} - \Pi^\mu_c \hat{B}_\mu \right)} |\Omega\rangle \,, \tag{4.63}$$

where $\Pi^\mu \equiv -h^{0\mu} + \frac{1}{2}\eta^{0\mu}h$. c-number functions set the initial expectation values of the corresponding fields and $|\Omega\rangle$ is the Minkowski vacuum. As in the case of QED, $B^c_\mu = 0$ due to the consistency (will be discussed in details in [6]). After imposing BRST condition, we get

$$\begin{aligned}\hat{Q}_B\,|h\rangle &= e^{-i\int d^3x\left(h^c_{ij}\hat{\pi}_{ij}-\pi^c_{ij}\hat{h}_{ij}-\Pi^\mu_c\hat{B}_\mu\right)}\\ &\times\int d^3x\left(-2h^c_{ij}(\nabla^2\delta_{ij}-\partial_i\partial_j)\hat{C}_0+2\pi^c_{ij}\partial_i\hat{C}_j\right)|\Omega\rangle\,. \end{aligned} \tag{4.64}$$

The consistency of the state requires the above expression to vanish. Upon integrating the second line by parts we are led to the relations similar to the classical constraints of the linear gravity

$$\left(\nabla^2\delta_{ij}-\partial_i\partial_j\right)h^c_{ij}=0\,,\quad \text{and}\quad \partial_i\pi^c_{ij}=0\,, \tag{4.65}$$

which must be satisfied by the physical configuration along with the following boundary conditions

$$\int d^3x\;\partial_i\left(\pi^c_{ij}\hat{C}_j\right)|\Omega\rangle=0\,, \tag{4.66}$$

$$\int d^3x\;\partial_i\left(\left[\partial_j\hat{C}_0-\hat{C}_0\partial_j\right](\delta_{ij}h^c_{kk}-h^c_{ij})\right)|\Omega\rangle=0\,. \tag{4.67}$$

Notice that the field $h^c_{0\mu}$ is unrestricted, since the corresponding operators represent Lagrange multipliers. Similar to A_0 in electrodynamics (they just change ImQ [6]. Hence, the physics is not altered). The equation (4.65) represents the gravitational charge-free Gauss' law, while (4.66) and (4.67) are equivalent to the boundary term of (4.14) and are automatically satisfied for configurations that vanish at the boundary/at the infinity.

4.2.2 The cosmological constant λ

In the previous section, we demonstrated the construction procedure of a coherent state from the free gravitons over the Minkowski vacuum. In this section we will add a classical source, the cosmological constant, to the theory and repeat the same steps. This scenario is close to infinitely homogeneous source in the QED (see the section (4.1.5)). So, we should alter the gravitational Lagrangian, adding the following term

$$\Delta\mathcal{L}=-\lambda\hat{h}\,. \tag{4.68}$$

This addition does not alter the expressions for canonical momenta, but it gives the additional contribution to the Nöether's charge. Since the above is invariant under the BRST transformations (4.59) only up to a total derivative,

$$\hat{Q}^\lambda_B=\hat{Q}_B+\int d^3x2\lambda\hat{C}_0\,. \tag{4.69}$$

We must define the physical Hilbert space w.r.t this charge, since it commutes with the Hamiltonian of the system in the presence of the cosmological constant ($[\hat{H},\hat{Q}^\lambda_B]=0$ while $[\hat{H},\hat{Q}_B]\neq 0$, where $\hat{Q}_B$ is the BRST charge without the cosmological constant).

The above discussion is an exact analogue of QED with classical charges (see the section 4.1.4). Here, we should build BRST invariant charges on top of the non-BRST invariant vacuum. The role of such vacuum is played by Minkowski space. We do not present the form of $\hat{Q}_B$, due to the lack of necessity. To build such a state, we need a classical soluton with a cosmological constant source. The solution is given by [132]

$$h_{ij} = -\frac{\lambda}{6}\left(t^2\delta_{ij} + x_i x_j\right), \quad h_{00} = h_{0j} = 0\,, \tag{4.70}$$

which, for $\lambda > 0$, represents the short-scale approximation of the de Sitter space-time. In fact, switching to the dimensionless metric $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}/M_{\rm pl}$, it is straightforward that (4.70) follows from the latter in $M_{\rm pl} \to \infty$ limit while λ is fixed. So, (4.58) is supplemented by (4.68) from the fully nonlinear Einstein-Hilbert action with the cosmological constant. It should be noted that in this limit the curvature of the spacetime vanishes,

$$H^2 \simeq \frac{\lambda}{M_{\rm pl}} \to 0\,. \tag{4.71}$$

This is equivalent to taking the Hubble radius to infinity. Thus, (4.70) is a good approximation of the de Sitter spacetime only at deep sub-horizon scales.

Using the above solution (4.70) let us build the coherent state on top of the Minkowski at $t = 0$ as

$$|h\rangle = e^{-i\int d^3x\ \left(h^c_{ij}\hat{\pi}_{ij} + \frac{1}{2}h^c_{kk}\hat{B}_0\right)}|\Omega\rangle\,, \tag{4.72}$$

$$\text{with} \quad h^c_{ij} = -\frac{\lambda}{6}x_i x_j\,. \tag{4.73}$$

It is straightforward to see that in the presence of the cosmological constant the first equation of (4.65) is modified as

$$\left(\nabla^2\delta_{ij} - \partial_i\partial_j\right)h^c_{ij} - \lambda = 0\,. \tag{4.74}$$

We should note, again, that $|\Omega\rangle$ is not annihilated by $\hat{Q}^\lambda_B$. The constraint (4.74) is readily satisfied by (4.70) and the boundary conditions (4.66) and (4.67) persist without modification. The first one is trivially satisfied, and the second one is satisfied in the same manner as (4.49).

4.2.3 Scalar matter coupled to gravity

In this section, we will couple gravity with the scalar matter field. We will consider only the highest orders in the ${M_{\rm pl}}^{-1}$ expansion. Let us introduce coupling of graviton to a scalar field. The corresponding Lagrangian has the form,

$$\mathcal{L} = \mathcal{L}_\phi + \mathcal{L}_h - \frac{1}{2M_{\rm pl}}\hat{h}_{\mu\nu}\hat{T}^{\mu\nu} + \mathcal{O}({M_{\rm pl}}^{-2})\,, \tag{4.75}$$

where $\mathcal{L}_\phi$ is a Lagrangian of the scalar field in the absence of gravity, $\mathcal{L}_h$ is given by (4.58) describing the free propagation of the graviton (including gauge fixing and ghost parts [135]) and $\hat{T}_{\mu\nu}$ stands for the energy-momentum tensor of the first two terms of (4.75). The gravitons' contribution to the energy momentum tensor requires modification of the BRST transformation (4.59) by adding non-linear terms. These terms do not contribute to the effects of order ${M_{\rm pl}}^{-1}$ that we are after (we can think in terms of the large-N (2.175)). Hence, we will only consider in the energy-momentum the scalar contribution.

The transformation property of the scalar field is given by,

$$\delta\hat{\phi} = \frac{\theta}{M_{\rm pl}}\hat{C}^\mu\partial_\mu\hat{\phi}\,, \tag{4.76}$$

which takes the following form when rewritten in terms of canonical variables,

$$\delta\hat{\phi} = \frac{\theta}{M_{\rm pl}}\left(\hat{\Pi}_\phi\hat{C}_0 - \partial_j\hat{\phi}\hat{C}_j\right) + \mathcal{O}({M_{\rm pl}}^{-2}) \tag{4.77}$$

where $\hat{\Pi}_\phi$ is a canonical momentum of the scalar.

Analogous to electrodynamics (4.19), the BRST invariant dressing of the scalar coherent state requires replacement of the scalar field operators by their gauge invariant counterparts.

An invariant version of $\hat{\phi}(x^\mu)$ is simply $\hat{\phi}(x^\mu + \epsilon^\mu)$, with ϵ, a function of the graviton field, transforming as

$$\delta\epsilon^\mu = -\frac{\theta}{M_{\rm pl}}C^\mu\,. \tag{4.78}$$

We can find the ϵ up to the first nontrivial order in ${M_{\rm pl}}^{-1}$. It is given by (similar objects were found in [137, 138]),

$$\epsilon_0 = -\frac{1}{4M_{\rm pl}}\frac{1}{\nabla^2}\hat{\pi}_{kk} + \mathcal{O}({M_{\rm pl}}^{-2})\,. \tag{4.79}$$

$$\epsilon_j = \frac{1}{M_{\rm pl}}\frac{1}{\nabla^2}\partial_i\left(\hat{h}_{ij} - \frac{1}{2}\delta_{ij}\hat{h}_{kk}\right) + \mathcal{O}({M_{\rm pl}}^{-2})\,. \tag{4.80}$$

The transformation properties can be verified using (4.59) and (4.62). Therefore, a coherent state of the scalar field that satisfies the physicality condition of being annihilated by the BRST charge can be constructed as

$$|C\rangle = e^{i\int d^3x \Pi^c_\phi(x)\hat{\phi}(x^\mu+\epsilon^\mu)}|\Omega\rangle\,, \tag{4.81}$$

with $\Pi^c_\phi(x)$ being an arbitrary function of the spatial coordinates and x^0 appearing in the argument of $\hat{\phi}$ sets the initial time (which we take $x^0 = 0$). The above state, is not the general one. In this case, the initial value of $\hat{\phi}$ is zero. We avoid writing the general one, since it is technical and does not give us any non-trivial information. The configuration still has kinetic energy, and hence it sources gravity, or more specifically up to order ${M_{\rm pl}}^{-2}$ we get

$$\langle C|\hat{\Pi}_\phi|C\rangle(t=0) = \Pi^c_\phi + \mathcal{O}({M_{\rm pl}}^{-2})\,, \tag{4.82}$$

$$\nabla^2\langle C|\hat{h}_{ij}|C\rangle(t=0) = \frac{\delta_{ij}}{8M_{\rm pl}}{\Pi^c_\phi}^2 + \mathcal{O}({M_{\rm pl}}^{-2})\,. \tag{4.83}$$

Notice, that (4.83) is the equation one would expect in Newtonian limit, for an arbitrary $\Pi^c_\phi(x)$ while having $\phi_c(x) = 0$.

4.2.4 Consistency of the linear de-Sitter

Not a long ago, it was observed, that de Sitter space-time is inconsistent with the S-matrix formulation. The first glimpse of it is the Gibbons-Hawking radiation [139]. Since the state is dynamical, it cannot be a vacuum. It should be more like a destabilized saturon as the Gross-Neveu saturons (section 2.5). So, in principle, due to its classical nature, it is best described by coherent state on top of the Minkowski [18], which apparently has quantum breaking [17]. The quantum breaking time $t_Q \sim {M_{\rm pl}}^2/H^3$ is finite for finite $M_{\rm pl}$. So, the system should exhibit a graceful exit, we cannot have an eternal inflation [17, 18]. The same argument destroys the possibility of having superselection of θ-vacua (section 3.1) [63], since $\theta \neq 0$ has non-zero vacuum energy (3.1).

So, we may ask if our construction of the de-Sitter-space in the linear gravity is self-consistent. Our construction was done in the linear theory, which cannot have any Gibbons-Hawking type radiation. So, the short answer is yes, but we should check it carefully. One should consider the following arguments together:

- The spin-2 theory is a self-consistent limit of GR,

  $$M_{\rm pl} \to \infty\,, \quad \lambda = \text{finite}\,. \tag{4.84}$$

  In this limit all non-linearities vanish, hence, the solution (4.70) and the corresponding coherent state (4.72) is perfectly well-defined.

- In the theory, the Hubble scale is vanishing (4.71), despite having non-zero cosmological constant λ. Hence, the quantum break time t_Q is infinite, i.e., there is no quantum breaking in this case.

- The effects of the de-Sitter are not experienced by gravitons, nor by any other fields. Let us consider a matter with the energy-momentum tensor $T_{\mu\nu}$ coupled to the gravity $h_{\mu\nu}$. The coupling suppressed by $M_{\rm pl}$ vanishes for solution $h_{\mu\nu}$ in the limit (4.84).

So, we can see that our construction of the coherent state is totally self-consistent. One may still argue that it is a totally decoupled theory, and decoupling makes it trivial and self-consistent. To address this question, we should consider Gibbons-Hawking radiation in the context of large-N gravity (see the section 2.6). That will be our topic of the next chapter.

4.2.5 Gibbons-Hawking radiation in linear gravity

In terms of the coherent state, the Gibbons-Hawking radiation can be understood as a decay (S matrix process) of the coherent state [17, 121], very similar to the decay of the saturon (see the section 2.5.4) The rate of the decay [17, 121] is estimated by

$$\Gamma \sim H\,, \tag{4.85}$$

and the power of the emitted radiation,

$$P \sim H^2\,, \tag{4.86}$$

which is in agreement with Gibbons-Hawking radiation of temperature H. In the articles [17, 121] is argued, that the state slowly looses the coherence, due to the decay process. The process leads to entanglement. So, eventually, the process becomes completely quantum, leading to the quantum break time [17].

Now, let us consider what happens in a large-N limit of the gravity (section 2.6). Let us consider the large-N Lagrangian (2.175) and add the cosmological term (4.68). In this case, the decay rate (4.85) is enhanced by the factor N,

$$\Gamma \sim H\,N\,. \tag{4.87}$$

Of course, in the present limit Γ diverges due to the infinite Hubble volume. Nevertheless, the particle production rate per unit volume is given by,

$$\frac{\Gamma}{V} \sim H^4 N = \left(\frac{\lambda}{\Lambda_{\text{gr}}}\right)^2\,, \tag{4.88}$$

which is finite. The result can be understood from the corpuscular perspective. The rate (4.87) has similarities with the rate (2.167). The gravitons, as saturons constituencies do, should re-scatter to give us creation of particles (like (2.166)). The rate accounts for a birth of pair of scalars. The rate of the process is given by

$$\Gamma \sim H\alpha_{\text{gr}}^2 N_{\text{gr}}^2\, N\,, \tag{4.89}$$

where $\alpha_{\text{gr}} = H^2/{M_{\text{pl}}}^2$ is the coupling between the coherent state gravitons and scalars and $N_{\text{gr}} = {M_{\text{pl}}}^2/H^2$ is the occupation number of gravitons in the coherent state per Hubble volume. Since $\alpha_{\text{gr}} N_{\text{gr}} = 1$, the two compensate each other. So, we get enhancement by a factor N, resulting with the finite rate per volume (4.88).

So, in this limit, despite having zero temperature, the quanta still can be created. Now, let us discuss if the coherent state undergoes the quantum breaking, consistency of our construction. The back-reaction on the coherent state from the particle-creation vanishes, despite the finite production rate per unit volume. This happens because the mean number density of the constituents is infinite in the limit (2.173) and their frequencies ($\sim H$) are zeros. The above renders to infinite quantum break time. The quantum break time [18, 19] can be estimated as,

$$t_Q \sim \frac{{M_{\text{pl}}}^2}{H^3 N} \sim \frac{\Lambda_{\text{gr}}^2}{H^3}\,, \tag{4.90}$$

which is infinite in a large-N limit. So, the coherent state in the linear theory is eternal. Since it has practically infinite coherence of gravitons compare to created quanta, the finite process cannot jeopardize its coherence.

Here we see, the power of the coherent states, in the BRST formulation we can immediately check their consistency and use them in various processes. Using the semi-classical physics directly not always gives us a self-consistent picture. Meantime, we see a significant simplification due to large-N physics, practically the only limit where the cosmological constant is consistent.

Conclusions

In this thesis, using the power of gauge redundancy and large-N physics, we have studied certain physically motivated systems.

The first important aspect is implication for a recently discovered phenomenon of saturation. It has been argued [8–10] that certain properties which previously have been attributed exclusively to black holes (such as the area-law entropy, thermal-like evaporation, long time-scale for information retrieval) go well-beyond gravity and are shared by a large class of object called Saturons. These are the entities that saturate a certain new bound (2.134) on the microstate degeneracy. It has been also shown in the same articles, that Saturons appear in variety of ordinary quantum field theories. They come in the form of solitons, baryons, instantons and other objects. The study of saturons is highly motivated, as this is an important step towards understanding the physics of black holes.

In the present thesis, based on publication [2], we have shown that saturons exist in Gross-Neveu model in the form of the fermion bound states of maximal entropy. This model is extremely useful as it is calculable in large N and the spectrum and the S-matrix are known. We have shown that N-fermion bound state, which carries the maximal entropy, is a saturon and exhibits all the above properties of a black hole. This is remarkable, especially in the light of the fact that the Gross-Neveu model is calculable and is a UV-complete asymptotically-free theory. The demonstrated connection between the properties of a black hole and of saturated N-fermion bound state also provides support for black hole N-portrait [7]. This is a microscopic theory, according to which a black hole represents a saturated bound-state of N gravitons.

Our second result is about a gauge redundant description of axion. This description was formulated in [15]. In it, the axion is replaced by an antisymmetric 2-form $B_{\mu\nu}$. The global shift symmetry of axion is substituted by the gauge redundancy of QCD, under which $B_{\mu\nu}$ shifts. The elimination of the strong-CP violation by axion is described as a Higgs effect, during which the Chern-Simons 3-form of QCD becomes massive by eating up the $B_{\mu\nu}$ axion. As a result, the topological susceptibility of QCD vacuum vanishes and the vacuum becomes exact CP-conserving.

By the power of gauge-redundancy, this formulation is immune against UV-corrections that would ordinarily break the global Peccei-Quinn symmetry explicitly.

In the present thesis, based on publication [1], we gave various consistency checks to UV-insensitivity of the gauge formulation of axion. We confirm that the formulation is indeed insensitive towards any heavy physics. This has important implications for axion physics, both at the fundamental level and phenomenologically.

The third direction of our research is in formulating the BRST invariant description of the de Sitter space in the form of the coherent state of gravitons.

It has been suggested some time ago [17] that quantum gravity cannot admit the de-Sitter vacuum, since such a state inevitably undergoes a phenomenon of quantum breaking. The impossibility of de Sitter as of vacuum is fully supported by the S-matrix formulation of quantum gravity [18]. Thus, the conclusion was that in quantum gravity, de-Sitter must be viewed as an excited N-graviton state on top of the S-matrix vacuum of Minkowski. The most natural candidate is coherent state of gravitons [17].

The previous constriction of de-Sitter coherent state, did not explicitly address the question of BRST invariance. Instead, the gauge redundancy was implemented by Stückelberg fields of Fierz-Pauli graviton.

In this thesis, based on publication [3], we have explicitly addressed the question of BRST invariance of coherent states in gauge-redundant theories.

We have constructed the BRST invariant coherent states of photons and gravitons in scalar electrodynamics and gravity, respectively. We have seen that, consistently, the state accounts for the classical electric and gravitational fields of the sources.

In particular, we have explicitly constructed a BRST invariant coherent state of gravitons. At the level of expectation values, this state reproduced the classical de-Sitter metric. We have also studied certain aspects of BRST invariant de-Sitter coherent state in a special regime of quantum gravity, called species regime [19]. This regime is obtained by coupling gravity to N_{sp} particle species. This lowers the gravitational cutoff to the so-called species scale $M_P/\sqrt{N_{sp}}$ [19]. Next, by taking the $N_{sp} \to \infty$ while keeping the species scale finite, one can effectively nullify the infinite class of quantum gravitational corrections. This simplifies the analysis significantly. We studied the de-Sitter coherent state of graviton in this limit, and showed that the evaporation rate per Hubble volume can be kept finite even though the Gibbons-Hawking temperature vanishes. Our research on de-Sitter coherent state is highly motivated for fundamental as well as observational aspects of gravity and cosmology. In particular, the study of the properties of quantum state of de-Sitter is crucial for understanding physics of dark energy and of inflationary cosmology.

Appendix A

Index notations

A.1 Space-time

We use the mostly minus metric tensor,

$$\eta_{\mu\nu} = diag(1, -1, ..., -1) \tag{A.1}$$

We use lower index notation. The repeating down indices should be summed via the metric tensor

$$A_\mu B_\mu = \sum_{\mu\nu} \eta^{\mu\nu} A_\mu B_\nu, \tag{A.2}$$

In this notation the contraction of two four-dimensional epsilon tensors is given by

$$\epsilon_{\mu\nu\alpha\beta}\epsilon_{\mu\nu\alpha\beta} = -24. \tag{A.3}$$

We also skip writing explicitly the contractions, e.g.,

$$\begin{aligned} F^2 &= F_{\mu\nu}F_{\mu\nu} \\ JC &= J_{\mu\nu\alpha}C_{\mu\nu\alpha} \\ \not{D} &= \gamma_\mu D_\mu. \end{aligned} \tag{A.4}$$

The tilde is defined via

$$\tilde{F} = \epsilon F. \tag{A.5}$$

A.2 Internal

We do not write explicitly the action of the group on elements, for instance, we write index free

$$\not{D}\psi, \tag{A.6}$$

instead of

$$\not{\partial}\psi_j + ig\not{A}^A T_j^{Ak}\psi_k. \tag{A.7}$$

Still, in some cases, the indices are written explicitly. We also do not write contraction of the adjoint index when it is clear, e.g.

$$G^A \tilde{G}^A = G\tilde{G} \tag{A.8}$$

Acknowledgements

Foremost, I would like to express my sincere gratitude to Prof. Gia Dvali for giving me a great opportunity to work with him and his group as a PhD student, for his continuous support during the research, insightful comments and very important advises. I am also thankful to him for the reference letters and providing extremely helpful advises for finding the postdoc position. Moreover, I want to thank him for his remarks and suggestions regarding the thesis.

I want to thank Dr. Lasha Berezhiani for the possibility to work on a very interesting project with him, his friendly and crucial comments, and the constant support during the whole PhD time. I am also deeply thankful for his support with reference letters and advises during the application procedure for the postdoc position.

I would like to express my gratitude to Prof. Goran Senjanovic for very insightful discussions, being the second referee of my thesis and for his support with recommendation letters and advises during the applications of a postdoc position.

My next thank goes to my former master students Andrea Paolini and Anja Stuhlfauth. It was an honour for me to co-supervise their theses. I am thankful to them for getting a great experience of supervision and an interesting collaboration.

I want to acknowledge very insightful discussions about the strong-CP problem in our group with Prof. Gia Dvali, Dr. Georgios Karananas, Dr. Lasha Berezhiani, Prof. Goran Senjanovic and Prof. Cesar Gomez.

My genuine thanks go to all committee members for their acceptance of being in the thesis committee.

I would like to express my deep appreciation to the IMPRS program for funding my PhD studies.

I am also grateful to our group for the friendly, professional and helpful atmosphere. I am especially delighted that I acquired friends during my PhD study.

I deeply appreciate the member of the administrations at MPI and ASC for their efficiency and kindness.

I express the special thanks to my wife Natalia Nebulishvili for proofreading of my thesis, and also for her continuous support along the PhD study.

I want to thank my family members and all friends who supported and wished all the best for me, especially my mother, who always was very supportive to me, during her life.

Bibliography

[1] O. Sakhelashvili, "Consistency of the dual formulation of axion solutions to the strong CP problem", Phys. Rev. D **105**, 085020 (2022), arXiv:2110.03386 [hep-th].

[2] G. Dvali and O. Sakhelashvili, "Black-hole-like saturons in Gross-Neveu", Phys. Rev. D **105**, 065014 (2022), arXiv:2111.03620 [hep-th].

[3] L. Berezhiani, G. Dvali, and O. Sakhelashvili, "de Sitter space as a BRST invariant coherent state of gravitons", Phys. Rev. D **105**, 025022 (2022), arXiv:2111.12022 [hep-th].

[4] G. Dvali, O. Sakhelashvili, and A. Paolini, "In preparation",

[5] G. Dvali and O. Sakhelashvili, "In preparation",

[6] L. Berezhiani, G. Dvali, and O. Sakhelashvili, "In preparation",

[7] G. Dvali and C. Gomez, "Black Hole's Quantum N-Portrait", Fortsch. Phys. **61**, 742–767 (2013), arXiv:1112.3359 [hep-th].

[8] G. Dvali, "Entropy Bound and Unitarity of Scattering Amplitudes", JHEP **03**, 126 (2021), arXiv:2003.05546 [hep-th].

[9] G. Dvali, "Unitarity Entropy Bound: Solitons and Instantons", Fortsch. Phys. **69**, 2000091 (2021), arXiv:1907.07332 [hep-th].

[10] G. Dvali, "Area Law Saturation of Entropy Bound from Perturbative Unitarity in Renormalizable Theories", Fortsch. Phys. **69**, 2000090 (2021), arXiv:1906.03530 [hep-th].

[11] D. J. Gross and A. Neveu, "Dynamical Symmetry Breaking in Asymptotically Free Field Theories", Phys. Rev. D **10**, 3235 (1974).

[12] R. D. Peccei and H. R. Quinn, "CP Conservation in the Presence of Instantons", Phys. Rev. Lett. **38**, 1440–1443 (1977).

[13] S. Weinberg, "A New Light Boson?", Phys. Rev. Lett. **40**, 223–226 (1978).

[14] F. Wilczek, "Problem of Strong P and T Invariance in the Presence of Instantons", Phys. Rev. Lett. **40**, 279–282 (1978).

[15] G. Dvali, "Three-form gauging of axion symmetries and gravity", (2005), arXiv:hep-th/0507215.

[16] G. Dvali and A. Vilenkin, "Cosmic attractors and gauge hierarchy", Phys. Rev. D **70**, 063501 (2004), arXiv:hep-th/0304043.

[17]G. Dvali, C. Gomez, and S. Zell, "Quantum Break-Time of de Sitter", JCAP **06**, 028 (2017), arXiv:1701.08776 [hep-th].

[18]G. Dvali, "S-Matrix and Anomaly of de Sitter", Symmetry **13**, 3 (2020), arXiv:2012.02133 [hep-th].

[19]G. Dvali, "Quantum Gravity in Species Regime", (2021), arXiv:2103.15668 [hep-th].

[20]G. Dvali and M. Redi, "Black Hole Bound on the Number of Species and Quantum Gravity at LHC", Phys. Rev. D **77**, 045027 (2008), arXiv:0710.4344 [hep-th].

[21]G. Dvali, "On S-Matrix Exclusion of de Sitter and Naturalness", (2021), arXiv:2105.08411 [hep-th].

[22]M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory* (Addison-Wesley, Reading, USA, 1995).

[23]S. Weinberg, *The quantum theory of fields*, Vol. 1 (Cambridge University Press, 1995).

[24]S. Weinberg, *The quantum theory of fields*, Vol. 2 (Cambridge University Press, 1996).

[25]L. D. Faddeev and V. N. Popov, "Feynman Diagrams for the Yang-Mills Field", Phys. Lett. B **25**, edited by J.-P. Hsu and D. Fine, 29–30 (1967).

[26]B. Lautrup, "CANONICAL QUANTUM ELECTRODYNAMICS IN COVARIANT GAUGES", (1967).

[27]N. Nakanishi, "Covariant Quantization of the Electromagnetic Field in the Landau Gauge", Prog. Theor. Phys. **35**, 1111–1116 (1966).

[28]T. Kugo and I. Ojima, "Local Covariant Operator Formalism of Non-Abelian Gauge Theories and Quark Confinement Problem", Progress of Theoretical Physics Supplement **66**, 1–130 (1979), eprint: https://academic.oup.com/ptps/article-pdf/doi/10.1143/PTPS.66.1/5402525/66-1.pdf.

[29]C. Becchi, A. Rouet, and R. Stora, "Renormalization of the Abelian Higgs-Kibble Model", Commun. Math. Phys. **42**, 127–162 (1975).

[30]M. Z. Iofa and I. V. Tyutin, "Gauge Invariance of Spontaneously Broken Nonabelian Theories in the Bogolyubov-Parasiuk-HEPP-Zimmerman Method", Teor. Mat. Fiz. **27**, 38–47 (1976).

[31]K. Fujikawa, "Path Integral Measure for Gauge Invariant Fermion Theories", Phys. Rev. Lett. **42**, 1195–1198 (1979).

[32]K. Fujikawa, "Path Integral for Gauge Theories with Fermions", Phys. Rev. **D21**, [Erratum: Phys. Rev.D22,1499(1980)], 2848 (1980).

[33]K. Fujikawa and H. Suzuki, *Path integrals and quantum anomalies* (2004).

[34]M. F. Atiyah and I. M. Singer, "The Index of elliptic operators. 1", Annals Math. **87**, 484–530 (1968).

[35]E. J. Weinberg, *Classical solutions in quantum field theory: Solitons and Instantons in High Energy Physics*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Sept. 2012).

[36]V. P. Nair, *Quantum field theory: A modern perspective* (2005).

[37]G. Dvali, *Advanced quantum field theory (lecture course)*, 2017.

[38]F. Quevedo, S. Krippendorf, and O. Schlotterer, "Cambridge Lectures on Supersymmetry and Extra Dimensions", (2010), arXiv:1011.1491 [hep-th].

[39]M. Shifman, *Advanced topics in quantum field theory.: A lecture course* (Cambridge Univ. Press, Cambridge, UK, Feb. 2012).

[40]S. Weinberg, *The quantum theory of fields. Vol. 3: Supersymmetry* (Cambridge University Press, June 2013).

[41]S. R. Coleman, "1/N", in 17th International School of Subnuclear Physics: Pointlike Structures Inside and Outside Hadrons (Mar. 1980).

[42]A. V. Manohar, "Large N QCD", in Les Houches Summer School in Theoretical Physics, Session 68: Probing the Standard Model of Particle Interactions (Feb. 1998), pp. 1091–1169, arXiv:hep-ph/9802419.

[43]Y. Frishman and J. Sonnenschein, "Non-Perturbative Field Theory – From Two Dimensional Conformal field Theory to QCD in Four Dimensions", (2010), arXiv:1004.4859 [hep-th].

[44]R. F. Dashen, B. Hasslacher, and A. Neveu, "Semiclassical Bound States in an Asymptotically Free Theory", Phys. Rev. D **12**, 2443 (1975).

[45]G. 't Hooft, "A Planar Diagram Theory for Strong Interactions", Nucl. Phys. B **72**, edited by J. C. Taylor, 461 (1974).

[46]S. Coleman and E. Weinberg, "Radiative corrections as the origin of spontaneous symmetry breaking", Phys. Rev. D **7**, 1888–1910 (1973).

[47]A. M. Polyakov, "Particle Spectrum in Quantum Field Theory", JETP Lett. **20**, edited by J. C. Taylor, 194–195 (1974).

[48]M. Nakahara, *Geometry, topology and physics* (2003).

[49]R. F. Dashen, B. Hasslacher, and A. Neveu, "Nonperturbative Methods and Extended Hadron Models in Field Theory 2. Two-Dimensional Models and Extended Hadrons", Phys. Rev. D **10**, 4130–4138 (1974).

[50]R. Jackiw and C. Rebbi, "Solitons with Fermion Number 1/2", Phys. Rev. D **13**, 3398–3409 (1976).

[51]M. B. Hindmarsh and T. W. B. Kibble, "Cosmic strings", Rept. Prog. Phys. **58**, 477–562 (1995), arXiv:hep-ph/9411342 [hep-ph].

[52]G. 't Hooft, "Magnetic Monopoles in Unified Gauge Theories", Nucl. Phys. B **79**, edited by J. C. Taylor, 276–284 (1974).

[53]A. A. Belavin, A. M. Polyakov, A. S. Schwartz, and Y. S. Tyupkin, "Pseudoparticle Solutions of the Yang-Mills Equations", Phys. Lett. B **59**, edited by J. C. Taylor, 85–87 (1975).

[54]A. I. Vainshtein, V. I. Zakharov, V. A. Novikov, and M. A. Shifman, "Abc of instantons", Phys. Usp. **25**, 195–215 (1982).

[55]G. 't Hooft, "Computation of the Quantum Effects Due to a Four-Dimensional Pseudoparticle", Phys. Rev. **D14**, [,70(1976)], 3432–3450 (1976).

[56]C. W. Bernard, N. H. Christ, A. H. Guth, and E. J. Weinberg, "Pseudoparticle parameters for arbitrary gauge groups", Phys. Rev. D **16**, 2967–2977 (1977).

[57]R. Rajaraman, *SOLITONS AND INSTANTONS. AN INTRODUCTION TO SOLITONS AND INSTANTONS IN QUANTUM FIELD THEORY* (1982).

[58]S. Coleman, *Aspects of symmetry: selected erice lectures* (Cambridge University Press, 1985).

[59]L. D. Landau and E. M. Lifshits, *Quantum Mechanics: Non-Relativistic Theory*, Vol. v.3, Course of Theoretical Physics (Butterworth-Heinemann, Oxford, 1991).

[60]C. G. Callan Jr., R. F. Dashen, and D. J. Gross, "The Structure of the Gauge Theory Vacuum", Phys. Lett. B **63**, edited by J. C. Taylor, 334–340 (1976).

[61]R. Jackiw and C. Rebbi, "Vacuum Periodicity in a Yang-Mills Quantum Theory", Phys. Rev. Lett. **37**, edited by J. C. Taylor, 172–175 (1976).

[62]R. D. Peccei, "QCD, strong CP and axions", J. Korean Phys. Soc. **29**, S199–S208 (1996), arXiv:hep-ph/9606475 [hep-ph].

[63]G. Dvali, C. Gomez, and S. Zell, "A Proof of the Axion?", (2018), arXiv:1811.03079 [hep-th].

[64]R. F. Dashen, B. Hasslacher, and A. Neveu, "Nonperturbative Methods and Extended Hadron Models in Field Theory 1. Semiclassical Functional Methods", Phys. Rev. D **10**, 4114 (1974).

[65]R. F. Dashen, B. Hasslacher, and A. Neveu, "Particle spectrum in model field theories from semiclassical functional integral techniques", Phys. Rev. D **11**, 3424–3450 (1975).

[66]G. Dvali, "Bounds on quantum information storage and retrieval", Phil. Trans. A. Math. Phys. Eng. Sci. **380**, 20210071 (2021), arXiv:2107.10616 [hep-th].

[67]G. Dvali, G. F. Giudice, C. Gomez, and A. Kehagias, "UV-Completion by Classicalization", JHEP **08**, 108 (2011), arXiv:1010.1415 [hep-ph].

[68]G. Dvali, C. Gomez, and A. Kehagias, "Classicalization of Gravitons and Goldstones", JHEP **11**, 070 (2011), arXiv:1103.5963 [hep-th].

[69]G. Dvali, C. Gomez, R. S. Isermann, D. Lüst, and S. Stieberger, "Black hole formation and classicalization in ultra-Planckian 2→N scattering", Nucl. Phys. B **893**, 187–235 (2015), arXiv:1409.7405 [hep-th].

[70]M. B. Voloshin, "Estimate of the onset of nonperturbative particle production at high-energy in a scalar theory", Phys. Lett. B **293**, 389–394 (1992).

[71]A. S. Gorsky and M. B. Voloshin, "Nonperturbative production of multiboson states and quantum bubbles", Phys. Rev. D **48**, 3843–3851 (1993), arXiv:`hep-ph/9305219`.

[72]L. S. Brown, "Summing tree graphs at threshold", Phys. Rev. D **46**, R4125–R4127 (1992), arXiv:`hep-ph/9209203`.

[73]E. N. Argyres, R. H. P. Kleiss, and C. G. Papadopoulos, "Amplitude estimates for multi - Higgs production at high-energies", Nucl. Phys. B **391**, 42–56 (1993).

[74]J. M. Cornwall, "On the High-energy Behavior of Weakly Coupled Gauge Theories", Phys. Lett. B **243**, 271–278 (1990).

[75]H. Goldberg, "Breakdown of perturbation theory at tree level in theories with scalars", Phys. Lett. B **246**, 445–450 (1990).

[76]E. Witten, "Baryons in the 1/n Expansion", Nucl. Phys. B **160**, 57–115 (1979).

[77]J. D. Bekenstein, "Black holes and entropy", Phys. Rev. D **7**, 2333–2346 (1973).

[78]S. W. Hawking, "Particle Creation by Black Holes", Commun. Math. Phys. **43**, edited by G. W. Gibbons and S. W. Hawking, [Erratum: Commun.Math.Phys. 46, 206 (1976)], 199–220 (1975).

[79]J. D. Bekenstein, "A Universal Upper Bound on the Entropy to Energy Ratio for Bounded Systems", Phys. Rev. D **23**, 287 (1981).

[80]D. N. Page, "Information in black hole radiation", Phys. Rev. Lett. **71**, 3743–3746 (1993), arXiv:`hep-th/9306083`.

[81]G. Dvali, O. Kaikov, and J. S. V. Bermúdez, "How special are black holes? Correspondence with objects saturating unitarity bounds in generic theories", Phys. Rev. D **105**, 056013 (2022), arXiv:`2112.00551 [hep-th]`.

[82]G. Dvali and R. Venugopalan, "Classicalization and unitarization of wee partons in QCD and gravity: The CGC-black hole correspondence", Phys. Rev. D **105**, 056026 (2022), arXiv:`2106.11989 [hep-th]`.

[83]F. Gelis, E. Iancu, J. Jalilian-Marian, and R. Venugopalan, "The Color Glass Condensate", Ann. Rev. Nucl. Part. Sci. **60**, 463–489 (2010), arXiv:`1002.0333 [hep-ph]`.

[84]G. Dvali, F. Kühnel, and M. Zantedeschi, "Vortexes in Black Holes", (2021), arXiv:`2112.08354 [hep-th]`.

[85]A. B. Zamolodchikov and A. B. Zamolodchikov, "Exact S Matrix of Gross-Neveu Elementary Fermions", Phys. Lett. B **72**, 481–483 (1978).

[86]A. Addazi, M. Bianchi, and G. Veneziano, "Glimpses of black hole formation/evaporation in highly inelastic, ultra-planckian string collisions", JHEP **02**, 111 (2017), arXiv:`1611.03643 [hep-th]`.

[87]G. Dvali, "A Microscopic Model of Holography: Survival by the Burden of Memory", (2018), arXiv:1810.02336 [hep-th].

[88]G. Dvali, L. Eisemann, M. Michel, and S. Zell, "Black hole metamorphosis and stabilization by memory burden", Phys. Rev. D **102**, 103523 (2020), arXiv:2006.00011 [hep-th].

[89]G. Dvali and C. Gomez, "Black Hole's 1/N Hair", Phys. Lett. B **719**, 419–423 (2013), arXiv:1203.6575 [hep-th].

[90]G. Dvali, "Non-Thermal Corrections to Hawking Radiation Versus the Information Paradox", Fortsch. Phys. **64**, 106–108 (2016), arXiv:1509.04645 [hep-th].

[91]G. Dvali, "Black Holes and Large N Species Solution to the Hierarchy Problem", Fortsch. Phys. **58**, 528–536 (2010), arXiv:0706.2050 [hep-th].

[92]C. Vafa and E. Witten, "Parity conservation in quantum chromodynamics", Phys. Rev. Lett. **53**, 535–536 (1984).

[93]G. 't Hooft, "Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking", NATO Sci. Ser. B **59**, 135–157 (1980).

[94]G. 't Hooft, "Symmetry Breaking Through Bell-Jackiw Anomalies", Phys. Rev. Lett. **37**, [,226(1976)], 8–11 (1976).

[95]G. 't Hooft, "How instantons solve the u(1) problem", Physics Reports **142**, 357 –387 (1986).

[96]E. Witten, "Current Algebra Theorems for the U(1) Goldstone Boson", Nucl. Phys. B **156**, 269–283 (1979).

[97]G. Veneziano, "U(1) Without Instantons", Nucl. Phys. B **159**, 213–224 (1979).

[98]M. Luscher, "The Secret Long Range Force in Quantum Field Theories With Instantons", Phys. Lett. B **78**, 465–467 (1978).

[99]G. Dvali, "A Vacuum accumulation solution to the strong CP problem", Phys. Rev. D **74**, 025019 (2006), arXiv:hep-th/0510053.

[100]G. Dvali, S. Folkerts, and A. Franca, "How neutrino protects the axion", Phys. Rev. D **89**, 105025 (2014), arXiv:1312.7273 [hep-th].

[101]G. Dvali, "Topological Origin of Chiral Symmetry Breaking in QCD and in Gravity", (2017), arXiv:1705.06317 [hep-th].

[102]M. Kamionkowski and J. March-Russell, "Planck scale physics and the Peccei-Quinn mechanism", Phys. Lett. B **282**, 137–141 (1992), arXiv:hep-th/9202003.

[103]R. Holman, S. D. H. Hsu, T. W. Kephart, E. W. Kolb, R. Watkins, and L. M. Widrow, "Solutions to the strong CP problem in a world with gravity", Phys. Lett. B **282**, 132–136 (1992), arXiv:hep-ph/9203206.

[104]S. M. Barr and D. Seckel, "Planck scale corrections to axion models", Phys. Rev. D **46**, 539–549 (1992).

[105] G. Dvali and L. Funcke, "Small neutrino masses from gravitational θ-term", Phys. Rev. D **93**, 113002 (2016), arXiv:1602.03191 [hep-ph].

[106] R. Kallosh, A. D. Linde, D. A. Linde, and L. Susskind, "Gravity and global symmetries", Phys. Rev. D **52**, 912–935 (1995), arXiv:hep-th/9502069.

[107] C. G. Callan Jr., R. F. Dashen, and D. J. Gross, "Toward a Theory of the Strong Interactions", Phys. Rev. D **17**, edited by M. A. Shifman, 2717 (1978).

[108] G. Dvali, "Large hierarchies from attractor vacua", Phys. Rev. D **74**, 025018 (2006), arXiv:hep-th/0410286.

[109] G. Dvali, R. Jackiw, and S.-Y. Pi, "Topological mass generation in four dimensions", Phys. Rev. Lett. **96**, 081602 (2006), arXiv:hep-th/0511175.

[110] L. D. Landau and E. M. Lifschits, *The Classical Theory of Fields*, Vol. Volume 2, Course of Theoretical Physics (Pergamon Press, Oxford, 1975).

[111] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, "The Hierarchy problem and new dimensions at a millimeter", Phys. Lett. B **429**, 263–272 (1998), arXiv:hep-ph/9803315.

[112] R. J. Glauber, "Coherent and incoherent states of the radiation field", Phys. Rev. **131**, 2766–2788 (1963).

[113] V. Chung, "Infrared Divergence in Quantum Electrodynamics", Phys. Rev. **140**, B1110–B1122 (1965).

[114] T. W. B. Kibble, "Coherent Soft-Photon States and Infrared Divergences. I. Classical Currents", J. Math. Phys. **9**, 315–324 (1968).

[115] T. W. B. Kibble, "Coherent soft-photon states and infrared divergences. ii. mass-shell singularities of green's functions", Phys. Rev. **173**, 1527–1535 (1968).

[116] T. W. B. Kibble, "Coherent soft-photon states and infrared divergences. iii. asymptotic states and reduction formulas", Phys. Rev. **174**, 1882–1901 (1968).

[117] T. W. B. Kibble, "Coherent soft-photon states and infrared divergences. iv. the scattering operator", Phys. Rev. **175**, 1624–1640 (1968).

[118] P. P. Kulish and L. D. Faddeev, "Asymptotic conditions and infrared divergences in quantum electrodynamics", Theor. Math. Phys. **4**, 745 (1970).

[119] G. Dvali, C. Gomez, L. Gruending, and T. Rug, "Towards a Quantum Theory of Solitons", Nucl. Phys. B **901**, 338–353 (2015), arXiv:1508.03074 [hep-th].

[120] G. Dvali and C. Gomez, "Black Holes as Critical Point of Quantum Phase Transition", Eur. Phys. J. C **74**, 2752 (2014), arXiv:1207.4059 [hep-th].

[121] G. Dvali and C. Gomez, "Quantum Compositeness of Gravity: Black Holes, AdS and Inflation", JCAP **01**, 023 (2014), arXiv:1312.4795 [hep-th].

[122] G. Dvali and C. Gomez, "Quantum Exclusion of Positive Cosmological Constant?", Annalen Phys. **528**, 68–73 (2016), arXiv:1412.8077 [hep-th].

[123] L. Berezhiani, "On Corpuscular Theory of Inflation", Eur. Phys. J. C **77**, 106 (2017), arXiv:1610.08433 [hep-th].

[124] L. Berezhiani and M. Trodden, "How Likely are Constituent Quanta to Initiate Inflation?", Phys. Lett. B **749**, 425–430 (2015), arXiv:1504.01730 [hep-th].

[125] J. Schwinger, "On gauge invariance and vacuum polarization", Phys. Rev. **82**, 664–679 (1951).

[126] L. Berezhiani and M. Zantedeschi, "Evolution of coherent states as quantum counterpart of classical dynamics", Phys. Rev. D **104**, 085007 (2021), arXiv:2011.11229 [hep-th].

[127] L. Berezhiani, G. Cintia, and M. Zantedeschi, "Background-field method and initial-time singularity for coherent states", Phys. Rev. D **105**, 045003 (2022), arXiv:2108.13235 [hep-th].

[128] P. A. M. Dirac, "Gauge invariant formulation of quantum electrodynamics", Can. J. Phys. **33**, 650 (1955).

[129] E. Bagan, M. Lavelle, and D. McMullan, "Charges from dressed matter: Construction", Annals Phys. **282**, 471–502 (2000), arXiv:hep-ph/9909257.

[130] H. Hirai and S. Sugishita, "Dressed states from gauge invariance", JHEP **06**, 023 (2019), arXiv:1901.09935 [hep-th].

[131] G. Dvali and M. Redi, "Phenomenology of 10^{32} Dark Sectors", Phys. Rev. D **80**, 055001 (2009), arXiv:0905.1709 [hep-ph].

[132] G. Dvali, S. Hofmann, and J. Khoury, "Degravitation of the cosmological constant and graviton width", Phys. Rev. D **76**, 084006 (2007), arXiv:hep-th/0703027.

[133] W. Mück, "Counting Photons in Static Electric and Magnetic Fields", Eur. Phys. J. C **73**, 2679 (2013), arXiv:1310.6909 [hep-th].

[134] W. Mück, "Photons in a Ball", Eur. Phys. J. C **75**, 585 (2015), arXiv:1510.04490 [hep-th].

[135] T. Kugo and I. Ojima, "Subsidiary Conditions and Physical S Matrix Unitarity in Indefinite Metric Quantum Gravitational Theory", Nucl. Phys. B **144**, 234–252 (1978).

[136] R. Arnowitt, S. Deser, and C. W. Misner, "Dynamical structure and definition of energy in general relativity", Phys. Rev. **116**, 1322–1330 (1959).

[137] W. Donnelly and S. B. Giddings, "Diffeomorphism-invariant observables and their nonlocal algebra", Phys. Rev. D **93**, [Erratum: Phys.Rev.D 94, 029903 (2016)], 024030 (2016), arXiv:1507.07921 [hep-th].

[138] M. B. Fröb and W. C. C. Lima, "Propagators for gauge-invariant observables in cosmology", Class. Quant. Grav. **35**, 095010 (2018), arXiv:1711.08470 [gr-qc].

[139] G. W. Gibbons and S. W. Hawking, "Cosmological Event Horizons, Thermodynamics, and Particle Creation", Phys. Rev. D **15**, 2738–2751 (1977).