Electrical control of a spin qubit in InSb nanowire quantum dots: Strongly suppressed spin relaxation in high magnetic field

Suzana Miladić,[1] Pavle Stipsić,[1] Edib Dobardžić,[1] and Marko Milivojević[2,3]
[1]*Faculty of Physics, University of Belgrade, Studentski trg 12, 11001 Belgrade, Serbia*
[2]*NanoLab, QTP Center, Faculty of Physics, University of Belgrade, Studentski trg 12, 11001 Belgrade, Serbia*
[3]*Department of Theoretical Physics and Astrophysics, Faculty of Science, P. J. Šafárik University, Park Angelinum 9, 040 01 Košice, Slovak Republic*

(Received 20 November 2019; revised manuscript received 2 February 2020; accepted 3 April 2020; published 22 April 2020)

In this paper we investigate the impact of gating potential and magnetic field on phonon induced spin relaxation rate and the speed of the electrically driven single-qubit operations inside the InSb nanowire spin qubit. We show that a strong g factor and high magnetic field strength lead to the prevailing influence of electron-phonon scattering due to deformation potential, considered irrelevant for materials with a weak g factor, like GaAs or Si/SiGe. In this regime we find that spin relaxation between qubit states is significantly suppressed due to the confinement perpendicular to the nanowire axis. We also find that maximization of the number of single-qubit operations that can be performed during the lifetime of the spin qubit requres single quantum dot gating potential.

DOI: 10.1103/PhysRevB.101.155307

I. INTRODUCTION

Spin of an electron confined in a semiconductor quantum dot (QD) can act as a carrier of quantum information [1] and a building block of quantum computers. In order to manipulate electron spin, usage of the external magnetic [2,3] and electric [4–6] field was suggested. Although spin control by means of a magnetic field is straightforward, electrical control of spin qubit through electric-dipole spin resonance (EDSR) is technologically more desirable [7–10].

Spin-orbit coupling (SOC) plays an essential role in the EDSR spin qubit scheme, since it allows transitions between qubit states using the spin-independent driving, such as electric-dipole interaction. On the other hand, the presence of SOC induces undesired phonon mediated transitions between qubit states [11–21]. In order to suppress the coupling to phonons, approaches like the optimal design of QDs [22,23] or the control of system size [24] was suggested.

Relaxation rates are dependent on the full three-dimensional QD potential, but in most cases contribution of the confinement along the direction(s) perpendicular to the substrate in which QDs are embedded can be neglected. Assuming magnetic fields up to several tesla, this reduction is justified in material with a weak effective Landé g factor. A typical example that satisfies this assumption are lateral GaAs QDs [25], while in the opposite direction lies an InSb nanowire, having two orders of magnitude stronger g factor [26]. Having also very strong SOC, spin qubits in InSb nanowires [27–31] have attracted much attention due to the observed [28] fast electric-dipole induced transition between qubit states, whose speed is equal to the strength of Rabi frequency.

Since both Rabi frequency and phonon induced relaxation rates are dependent on the magnetic field orientation and strength, design of the gating potential, and SOC, there is a wide range of possibility to tune their strength, with the goal of obtaining as much as possible single-qubit operations during its lifetime.

In this paper we search for the optimal regime in which electrical control of the InSb spin qubit can be achieved. We analyze both single and double quantum dot (DQD) potential and discuss its positive features and negative drawbacks on the spin qubit. In the case of double quantum dot potential, there is the possibility to tune the distance between the dots and to analyze the effects of the asymmetric gating potential. Also, we address the situations in which full three-dimensional confinement has nontrivial influence on spin relaxation rates. We will show that scattering by deformation potential dominates in this regime. Finally, to offer a quantitative insight into the spin qubit quality, we define a figure of merit as the ratio of Rabi frequency and the overall spin relaxation rate and discuss the obtained results in terms of this measure.

This paper is organized as follows. In Sec. II the single-electron Hamiltonian model of the InSb nanowire is introduced. In Sec. III we start with the definition of Rabi frequency and phonon induced spin relaxation rate between spin qubit states. After that, we independently study their dependence on tunable parameters of the system. Using the obtained results, quality of the spin qubit is discussed with the help of the figure of merit as a quantitative measure. In the end we finish the paper with a short conclusion and the impact of the presented results.

II. NANOWIRE SPIN QUBIT MODEL

We start with the Hamiltonian describing the electron confined in an InSb nanowire [30]

$$H = \frac{p^2}{2m^*} + V(x) + H_{so} + H_z, \qquad (1)$$

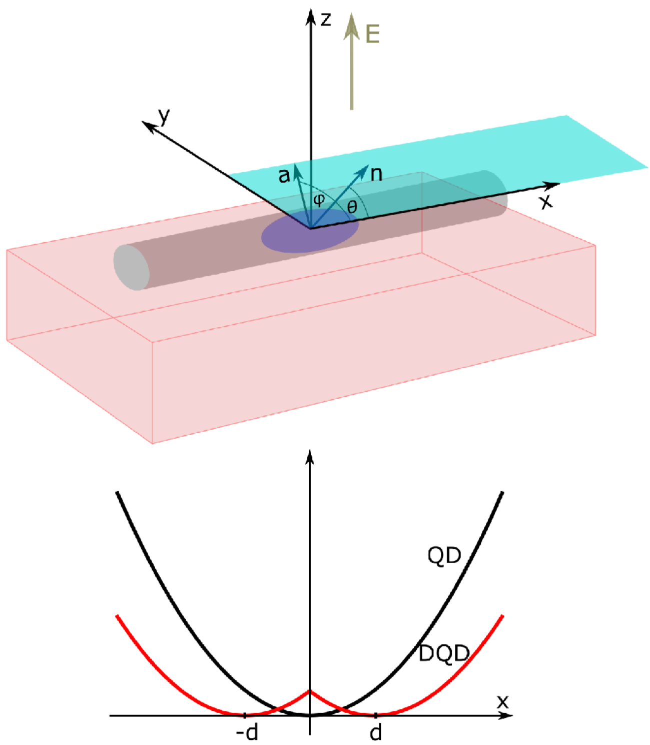


FIG. 1. (Upper panel) *Nanowire QD—schematic view.* Electron dynamics along the nanowire (x) axis is described by the Hamiltonian H, given in Eq. (1). Angle between the nanowire x axis and magnetic field direction $\mathbf{n} = (\cos\theta, \sin\theta, 0)$ is equal to θ, while the spin-orbit vector $\mathbf{a} = (\cos\varphi, \sin\varphi, 0)$ builds an angle φ with the x axis. (Lower panel) Confining potential used in Eq. (1): QD and DQD potential. In the case of a DQD potential [Eq. (6)] symmetric confinement is depicted ($\omega_L = \omega_R$), with distance between the dots equal to $2d$.

where m^* is the effective mass, $p = -i\hbar\partial/\partial_x$ momentum in x direction, $V(x)$ is the gating potential used to localize the electron, while $H_{\rm so}$ represents the spin-orbit interaction Hamiltonian consisting of two terms: Dresselhaus [32] and Rashba [33]. The presence of the Dresselhaus SOC is due to the material in which an electron is embedded. On the other hand, Rashba SOC appears when an electric field E in the z direction is applied (see Fig. 1). In an InSb nanowire, a spin-orbit interaction Hamiltonian is equal to [30]

$$H_{\rm so} = (\alpha_D\sigma_x + \alpha_R\sigma_y)p, \quad (2)$$

where σ_x and σ_y are Pauli matrices, while α_D and α_R are Dresselhaus and Rashba spin-orbit coupling strengths. Suitable change of parameters α_R and α_D with $\alpha = \sqrt{\alpha_D^2 + \alpha_R^2}$ and $\varphi = \arctan(\alpha_R/\alpha_D)$ allows us to write Eq. (2) as

$$H_{\rm so} = \alpha\boldsymbol{a}\cdot\boldsymbol{\sigma}p, \quad (3)$$

using the unit spin-orbit vector $\boldsymbol{a} = (\cos\varphi, \sin\varphi, 0)$ and the vector $\boldsymbol{\sigma}$ made of Pauli matrices. Finally, $H_{\rm z}$ is the Zeeman term, describing the coupling of spin and magnetic field

$$H_{\rm z} = \frac{g}{2}\mu_B\boldsymbol{B}\cdot\boldsymbol{\sigma}, \quad (4)$$

where g is the effective Landé factor, μ_B is the Bohr magneton, while $\boldsymbol{B} = B\boldsymbol{n}$ is the applied magnetic field in the plane of the substrate, building an angle θ with the growth x axis of the nanowire (see the upper panel of Fig. 1). In this work a magnetic field is considered to be in-plane to minimize the orbital effects [22,34–36]. In Appendix A we have shown that for B up to 3 T, orbital effects of a magnetic field are small and can be neglected.

Typical gating that confines a single electron in experimental setups [37] can be modeled as a harmonic oscillator quantum dot (QD) [38] or double quantum dot (DQD) [29] potential. Corresponding potentials are equal to (see the lower panel of Fig. 1 as an illustration)

$$V^{\rm QD}(x) = \tfrac{1}{2}m^*\omega^2x^2, \quad (5)$$

$$V^{\rm DQD}(x) = \tfrac{1}{2}m^*\min\{\omega_L^2(x+d)^2, \omega_R^2(x-d)^2\}. \quad (6)$$

In the case of a QD potential, the only degree of freedom is the harmonic potential frequency ω, while in the DQD case frequencies ω_L and ω_R can be tuned, as well as the distance $2d$ between the dots. Since DQD potential allows asymmetric confinement, we introduce asymmetry parameter δ, equal to the ratio of frequencies in the left and right dot, $\delta = \omega_L/\omega_R$. Impact of the DQD confinement will be discussed in terms of δ, $2d$, and $\omega_R = \omega$ (more detailed explanation can be found in Sec. III A).

The Hamiltonian of the electron in different potential types and magnetic field strengths can be solved using the numerical diagonalization [39], although perturbative approaches in the study of spin qubit properties are common [21,27,30]. In this work we follow the numerical approach; the numerical procedure used in obtaining the eigenvalues and eigenvectors of the Hamiltonian given in Eq. (1) is explained in Appendix B. In order to successfully diagonalize the Hamiltonian, orbital $x_0 = \sqrt{\hbar/m^*\omega}$ and spin-orbit $x_{\rm so} = \hbar/m^*\alpha$ lengths are defined. In our calculations we have used $m^* = 0.014\ m_e$ [29], $x_0 = 30$ nm [29], and $x_{\rm so} = 165$ nm [40] parameters for both QD and DQD potentials (recall that $\omega_R = \omega$ in the DQD case), related to the experimental reports on InSb nanowires. On the other hand, we have used g factor in bulk InSb material, $g = -51.3$ [41], being in the range of the experimentally reported values [38,42]. Initial check of the numerical recipe presented in Appendix B were exact analytical results obtained in the special case of the infinite square well [43]. In this case we were able to reproduce the results concerning the angular dependence of the energy splitting between Zeeman sublevels, Rabi frequency, and the relaxation rate.

The nanowire Hamiltonian [Eq. (1)] describes the single-electron dynamics in the x direction only. To ensure the validity of the one-dimensional approximation and to suppress the dynamics in the yz plane, a much stronger yz plane confinement than in the x direction is needed. In this case, a wave function along both directions, y and z, will correspond to the respective ground state. To take into the account the wire geometry of the system, the same confinement length $y_0 = z_0 = 10$ nm in the $y(z)$ direction is assumed. We model the confinement potential as harmonic [39], to which the ground state wave function $\psi(y) = e^{-y^2/2y_0^2}/\sqrt{\sqrt{\pi}y_0}$ corresponds. In the z direction an additional potential eEz ($z > 0$; $z = 0$

corresponds to the position of the substrate) is present due to the applied electric field. Finally, the substrate acts as an infinite potential barrier for the confined electron, forbidding him to propagate in the $z < 0$ region [44]. The ground state $\psi(z)$ of the Hamiltonian in the z direction is found using the same numerical method as for the Hamiltonian in the x direction. Thus, the ground state wave function in the yz plane is equal to $\Psi(y, z) = \psi(y)\psi(z)$.

III. EDSR AND SPIN RELAXATION IN NANOWIRE SPIN QUBIT

In order to achieve electrical control of the nanowire spin qubit, an oscillating electric field in the x direction should be switched on, resulting in the Rabi Hamiltonian $H_R = eE_0 x \cos(\omega_E t)$. When the applied electric field is in resonance with our quantum system, Rabi frequency Ω_{01} is defined as

$$\Omega_{01} = \frac{eE_0}{h} |\langle 0|x|1\rangle|, \tag{7}$$

measuring the speed of the single-qubit rotations. In Eq. (7) states $|0\rangle$ and $|1\rangle$ correspond to the ground and first excited state of the single electron Hamiltonian H, while $e|\langle 0|x|1\rangle|$ is the dipole matrix element. We are particularly interested in the case where qubit states are Zeeman sublevels of the orbital ground state, since in this regime strength of the Rabi frequency can be manipulated by changing the magnetic field orientation [30].

Besides providing the opportunity to electrically control the spin qubit, SOC triggers the undesired phonon induced transition between qubit states, setting up a limit on the qubit lifetime. Rate of spin relaxation can be determined from the Fermi golden rule

$$\Gamma_{01} = \frac{2\pi}{\hbar} \sum_{\nu\mathbf{q}} |M_\nu(\mathbf{q})|^2 |\langle \psi_0|e^{i\mathbf{q}\cdot\mathbf{r}}|\psi_1\rangle|^2 \delta(\Delta E_{01} - \hbar\omega_{\nu\mathbf{q}}). \tag{8}$$

Transition is triggered by acoustic phonons of energy $\hbar\omega_{\nu\mathbf{q}}$ that correspond to the energy separation between qubit states, $\Delta E_{01} = |E_0 - E_1|$. We assume a linear dispersion relation of acoustic phonons with respect to the intensity of wave vector $\mathbf{q}$, $\omega_{\nu\mathbf{q}} = c_\nu|\mathbf{q}|$, yielding $|\mathbf{q}| = \Delta E_{01}/\hbar c_\nu$.

Next, three different geometric factors $|M_\nu(\mathbf{q})|^2$ entering spin relaxation rates originate from different types of electron-phonon scattering: electron-longitudinal phonon scattering due to the deformation potential [45]

$$|M_{\mathrm{LA-DP}}(\mathbf{q})|^2 = \frac{\hbar D^2}{2\rho c_{\mathrm{LA}} V} |\mathbf{q}|, \tag{9}$$

electron-longitudinal phonon scattering due to the piezoelectric field [45]

$$|M_{\mathrm{LA-PZ}}(\mathbf{q})|^2 = \frac{32\pi^2\hbar(eh_{14})^2}{\epsilon^2\rho c_{\mathrm{LA}} V} \frac{(3q_x q_y q_z)^2}{|\mathbf{q}|^7}, \tag{10}$$

where h_{14} is piezoelectric constant, and electron-transverse phonon scattering due to the piezoelectric field [45]

$$|M_{\mathrm{TA-PZ}}(\mathbf{q})|^2 = 2\frac{32\pi^2\hbar(eh_{14})^2}{\epsilon^2\rho c_{\mathrm{TA}} V} \times \left| \frac{q_x^2 q_y^2 + q_x^2 q_z^2 + q_y^2 q_z^2}{|\mathbf{q}|^5} - \frac{(3q_x q_y q_z)^2}{|\mathbf{q}|^7} \right|. \tag{11}$$

Finally, spin relaxation rates are dependent on the transition matrix element $|\langle \psi_0|e^{i\mathbf{q}\cdot\mathbf{r}}|\psi_1\rangle|^2$ which depends on the full three-dimensional confinement. In order to divide the contribution of confinements along the nanowire axis and the yz plane, we write the transition matrix element as $|\langle 0|e^{iq_x x}|1\rangle|^2 |T_{yz}|^2$, where $|\langle 0|e^{iq_x x}|1\rangle|^2$ is the contribution along the nanowire direction, while

$$|T_{yz}|^2 = \left| \iint dydz |\Psi(y, z)|^2 e^{i(q_y y + q_z z)} \right|^2 \tag{12}$$

represents scattering in a plane perpendicular to the nanowire axis.

The role of $|T_{yz}|^2$ in the spin relaxation rate depends on the regime in which spin qubit operates. At low magnetic fields, when $|\mathbf{q}|z_0 \ll 1$ and $|\mathbf{q}|y_0 \ll 1$, dipole approximation $e^{i\mathbf{q}\cdot\mathbf{r}} \approx 1 + i\mathbf{q}\cdot\mathbf{r}$ is valid [22] and $|T_{yz}|^2$ can be replaced with $(1 + |\mathbf{q}|^2 z_0^2 \cos^2\theta) \approx 1$, implying that one-dimensional approximation is justified. However, at higher magnetic fields, dipole approximation is not valid and confinement in the yz direction can play a significant role. To determine its role in the spin relaxation rate, we have numerically calculated $|T_{yz}|^2$ beyond the dipole approximation.

Magnetic field strengths for which the system operates outside of the dipole approximation ($|\mathbf{q}|y_0 \geqslant 1$) can be roughly estimated; assuming energy separation between qubit states proportional to $g\mu_B B$, Fermi golden rule determines phonon wave number $|\mathbf{q}| = g\mu_B B/(\hbar c_\lambda)$, where $c_{\mathrm{LA}} = 3800$ m/s [46] and $c_{\mathrm{TA}} = 1900$ m/s [47], giving us magnetic field strengths for the electron-phonon scattering in the longitudinal (0.084 T) and transverse (0.042 T) direction above which we are outside of the dipole approximation.

Before we continue, we provide necessary parameters for the calculation of the spin relaxation rate: $eh_{14} = 1.41 \times 10^9$ eV/m [45], $\epsilon = 16.5$, $D = 7$ eV [48], $\rho = 5775$ kg/m^3 [49].

A. Rabi frequency

We start the discussion of obtained results with the analysis of Rabi frequency dependence on the parameters of interest.

In Fig. 2(a), dependence of Ω_{01} (in eE_0x_0/h units) on $\theta - \varphi$ and magnetic field strength is presented for the QD confinement potential. Our results confirm the expected π periodic behavior with respect to $\theta - \varphi$ [30]. Depending on the magnetic field strength, results can be divided into two classes. In the first class qubit states represent Zeeman sublevels of the orbital ground state; in this regime zero Rabi frequency can be found for special magnetic field orientations ($\theta - \varphi = 0, \pi$), since these qubit states have orthogonal spin components. In the second class, magnetic field strengths have led to rearrangement of energy levels, such that qubit states originate from the ground and the first excited orbital state. In this situation, an orbital qubit is constructed, with a very weak dependence of Ω_{01} on $\theta - \varphi$ ($\Omega_{01} \neq 0$ in the orbital qubit regime for any $\theta - \varphi$). Critical magnetic field value B_c of spin to orbital qubit transition is almost independent on $\theta - \varphi$ and can be easily determined from the eigenspectrum analysis. Alternatively, for $\theta - \varphi = 0, \pi$, abrupt switch of Ω_{01} from zero to the nonzero value at B_c is a fingerprint of the

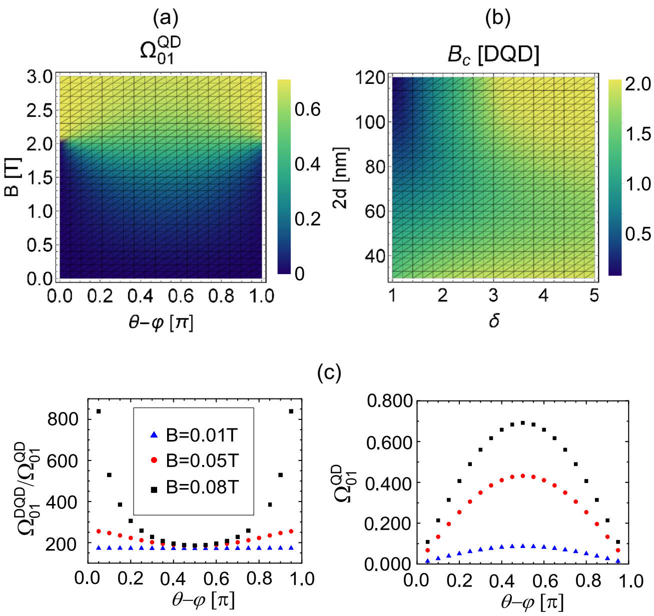


FIG. 2. (a) Dependence of Rabi frequency Ω_{01}^{QD} (in eE_0x_0/h units) on $\theta - \varphi \in (0, \pi)$ and $B \in (0, 3)$ T for QD gating potential. (b) In the case of DQD confinement, dependence of B_c on the asymmetry parameter $\delta \in (1, 5)$ and distance between the dots $2d \in (30, 120)$ nm is given. (c) Dependence of the ratio $\Omega_{01}^{DQD}/\Omega_{01}^{QD}$ on $\theta - \varphi \in (0.05, 0.95)\pi$ and magnetic field strengths $B = 0.01$ T, $B = 0.05$ T, and $B = 0.08$ T is presented for the symmetric DQD potential; distance between the dots is equal to $2d = 120$ nm. For the same angle range and magnetic field values Ω_{01}^{QD} in eE_0x_0/h units is presented.

transition. In the case of the QD potential, we extract the critical magnetic field value $B_c \approx 2.04$ T.

Gating with DQD potential gives a qualitatively similar dependence of Ω_{01} on B and $\theta - \varphi$. Being interested in the qualitative comparison of the impacts of QD and DQD potentials, we first establish a basis for comparison between them. To this end, we assume the same frequency of the QD potential and the right dot of the DQD potential, $\omega = \omega_R$, and vary the asymmetry parameter δ and the distance between the dots $2d$. For highly asymmetric DQD confinement and the large interdot distance, the electron will reside on only one dot, i.e., this potential is effectively the same as the single QD potential. The qualitative similarity of the single and double QD potential is checked through the comparison of the probability density of the ground and first excited state (qubit states); similar probability density profiles of the qubit states directly correspond to the similar Rabi frequency values of the two systems. Using the numerical comparison of the probability densities and the Rabi frequency in the case of QD and DQD potential, it can be concluded that for $2d \geqslant 120$ nm and $\delta \geqslant 5$ there is no effective difference between the results arising from two potentials. In other words, one should use $\delta < 5$ and $2d < 120$ nm to test the genuine effects of the DQD potential.

Figure 2(b) depicts the dependence of B_c in the DQD case on $\delta \in (1, 5)$ and $2d \in (30, 120)$ nm. When compared to the B_c value in the QD case, drastically lower values are found, especially in the case of symmetric confinement with well separated left and right QD. As an example, critical magnetic field value $B_c \approx 0.085$ T for the symmetric DQD confinement with $2d = 120$ nm is roughly 24 times smaller than in the QD case.

Lower B_c for the symmetric DQD confinement is followed by at most factor 3 increase of $\Omega_{01}(B_c^{DQD})$, when compared to $\Omega_{01}(B_c^{QD})$. This slight increase, followed by lower B_c below which symmetric DQD operates, indicates a steeper rise of Rabi frequency for symmetric DQD confinements and the possibility to induce an even bigger difference between Ω_{01}^{DQD} and Ω_{01}^{QD} for the optimal magnetic field configuration. To investigate this possibility, we have performed a numerical analysis of the Rabi frequency ratio $\Omega_{01}^{DQD}/\Omega_{01}^{QD}$ for a wide range of DQD confinements and different magnetic field strengths/orientations, such that both systems operate as spin qubits. Our results confirm that symmetric DQD confinement maximally enhances this ratio when operating at magnetic field strengths close to B_c for the DQD potential, while the field orientation should be chosen such that $\theta - \varphi$ is close to 0 or π. In order to illustrate this conclusion, in the left panel of Fig. 2(c) we present the ratio $\Omega_{01}^{DQD}/\Omega_{01}^{QD}$ for $2d = 120$ nm and $\delta = 1$ in the DQD case, assuming field orientations $\theta - \varphi \in (0.05, 0.95)\pi$ and magnetic field strengths $B = 0.01$ T, $B = 0.05$ T, and $B = 0.08$ T ($B_c^{DQD} \approx 0.085$ T for this setup). Since angles $\theta - \varphi = 0, \pi$ should be excluded from the analysis because they correspond to zero Rabi frequency, we have restricted our plots to a $\theta - \varphi$ region smaller than π [see the right panel of Fig. 2(c) for the Ω_{01}^{QD} values], obtaining the highest ratio of around 800. It should be noticed that for angles closer to $0/\pi$ even bigger ratios (10^4) can be obtained, but at the cost of lowering the value of Rabi frequency.

B. Spin relaxation

Another important component for determining spin qubit quality is the spin relaxation rate. Similarly as Rabi frequency, Γ_{01} is dependent on the magnetic field and gating potential. However, Γ_{01} can be additionally dependent on the confinement in yz plane. In order to compare the influence of three-dimensional confinement with the confinement along the nanowire axis solely, we define one-dimensional approximation of the relaxation rate Γ_{01}^{1D} by changing the transition matrix element $|\langle\psi_0|e^{i\mathbf{q}\cdot\mathbf{r}}|\psi_1\rangle|^2$ with $|\langle 0|e^{iq_x x}|1\rangle|^2$ in Eq. (8).

It has been known that in lateral GaAs QDs spin relaxation rates are dominated by piezoelectric field [50,51]. In our case, we wish to analyze the influence of each relaxation channel; thus, the overall spin relaxation rate will be divided into three contributions:

$$\Gamma_{01} = \Gamma_{01}^{LA-DP} + \Gamma_{01}^{LA-PZ} + \Gamma_{01}^{TA-PZ}, \tag{13}$$

each dependent on a different geometric factor, see Eqs. (9)–(11).

Before presenting the numerical results, conclusions independent on the choice of gating potentials are provided. First, Γ_{01} shows oscillatory dependence on the $\theta - \varphi$ angle, being equal to zero for $\theta - \varphi = 0, \pi$ and reaching the maximum for $\theta - \varphi = \pi/2$ in the spin qubit regime [21]. Second, for weak magnetic field strengths ($B < 0.1$ T), piezoelectric fields dominate relaxation rates. At the same time, yz confinement can be ignored.

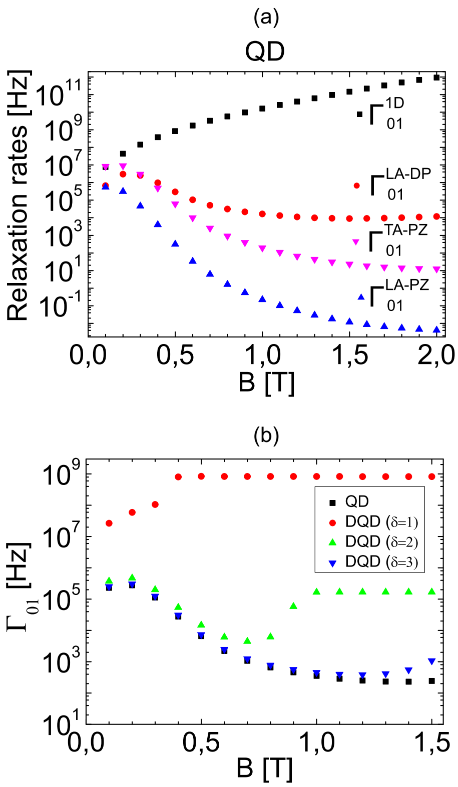


FIG. 3. (a) Dependence of the relaxation rates on the magnetic field strength $B \in (0.1, 2)$ T for $\theta - \varphi = \pi/2$. Red circles represent the contribution of deformation potential in the scattering rates, while inverted pink (blue) triangles show the contribution of piezoelectric field for the electron-phonon scattering in the transverse (longitudinal) direction. Finally, black squares represent relaxation rates in the one-dimensional approximation, in which the contribution of the confinement perpendicular to the nanowire axis is neglected. (b) Dependence of Γ_{01} on the magnetic field strength $B \in (0.1, 1.5)$ T in the case of QD and DQD confinement potential. Magnetic field orientation is chosen such that $\theta - \varphi = 0.05\pi$. In the DQD case, the distance between the dots is set at 90 nm, while the asymmetry parameter is varied.

To explore a new type of behavior accessible in InSb spin qubits, we focus our attention on stronger magnetic fields and investigate its impact on each relaxation channel and one-dimensional approximation of the total relaxation rate Γ_{01}^{1D}. We start from the QD potential. In Fig. 3(a), dependence of relaxation rates on $B \in (0.1, 2)$ T for the fixed angle $\theta - \varphi = \pi/2$ is given [52]. Red circles represent the contribution of deformation potential, pink inverse (blue) triangles denote the impact of piezoelectric field in the electron-phonon scattering along the transverse (longitudinal) direction. Graphs show that relaxation rate $\Gamma_{01}^{\mathrm{LA-PZ}}$ can safely be ignored, while $\Gamma_{01}^{\mathrm{LA-DP}}$ and $\Gamma_{01}^{\mathrm{TA-PZ}}$ have nontrivial influence on Γ_{01}. For weak magnetic fields $\Gamma_{01}^{\mathrm{TA-PZ}}$ term is dominant, while for large magnetic fields $\Gamma_{01}^{\mathrm{LA-DP}}$ should be considered solely [39]. A different influence of $\Gamma_{01}^{\mathrm{TA-PZ}}$ and $\Gamma_{01}^{\mathrm{LA-DP}}$ lies in the opposite behavior of the corresponding geometric factors: $|M_{\mathrm{TA-PZ}}(\mathbf{q})|^2$ $[|M_{\mathrm{LA-DP}}(\mathbf{q})|^2]$ is inversely (directly) proportional to the energy splitting between the Zeeman levels and decreases (increases) with the magnetic field rise.

Contribution of the yz plane confinement on the spin relaxation rate can be determined by comparing the Γ_{01}^{1D} with relaxation rate channels. The comparison is illustrated in Fig. 3(a), clearly demonstrating that one-dimensional approximation of the spin relaxation rate is valid only for weak magnetic fields, below 0.1 T. At higher fields, due to the strong g factor of the InSb material, both $|\mathbf{q}|y_0$ and $|\mathbf{q}|z_0$ are greater than one, triggering the effects of the yz plane confinement for each relaxation rate channel. Thus, suppressed spin relaxation represents a fingerprint of a material with a strong g factor.

In the case of DQD potentials, dependence of B_c on the form of gating presents a serious limitation on the regimes that can be accessed. For example, if the B_c value is sufficiently weak, $B_c < 0.1$ T, the spin qubit operates under the dominant influence of the piezoelectric field. A strong magnetic field regime is beneficial for spin qubit operation due to strong Rabi frequency and suppressed spin relaxation. In order to operate in this regime, asymmetric DQD potential should be used. To compare the influence of QD and DQD potential on Γ_{01}, in Fig. 3(b), we plot the dependence of the spin relaxation rate in the case of QD and DQD confinement on the magnetic field strength $B \in (0.1, 1.5)$ T, assuming $\theta - \varphi = \pi/2$ and $2d =$ 90 nm. Besides the symmetric $\delta = 1$ confinement, asymmetric DQD confinements ($\delta = 2, 3$) were analyzed as well. The presented results show that DQD gating leads to increased relaxation rates, when compared to the QD potential. This difference is minimized for highly asymmetric gating potentials. Note that B independent Γ_{01} values suggest that orbital qubit is created: energy difference between the states with the same spin component (representing the orbital qubit states in our case) is independent on B and triggers phonons on the same energy, leading to the observed effect. Consequently, these points should be excluded from the spin qubit analysis.

Finally, we emphasize that in the special case of the asymmetric DQD potential with $\delta = 1.5$ a similar trend of the spin relaxation rate is ascertained [21], i.e., after the increase of the spin relaxation rate in the dominant regime of the piezoelectric field, suppression of spin relaxation is observed, followed by the increase up to magnetic field independent saturation value [see the green triangles in Fig. 3(b) as a comparison].

C. Spin qubit quality

Quantitative estimate of the spin qubit quality can be given with the help of the figure of merit ξ [22],

$$\xi = \frac{\Omega_{01}}{\Gamma_{01} + \Gamma_{\mathrm{o}}}, \tag{14}$$

measuring the number of qubit operations that can be implemented during the qubit lifetime. In Eq. (14) Γ_{o} represents relaxation rate of decay channels different from phonons. To divide the contribution of phonons from them, we rewrite ξ in terms of the phonon figure of merit $\xi_{\mathrm{ph}} = \Omega_{01}/\Gamma_{01}$ and relative influence of other channels with respect to phonons $\Gamma_{\mathrm{o}}/\Gamma_{01}$. Thus,

$$\xi = \frac{\xi_{\mathrm{ph}}}{1 + \frac{\Gamma_{\mathrm{o}}}{\Gamma_{01}}}. \tag{15}$$

We first analyze ξ_{ph} for the QD confinement. Neglecting the weak magnetic field regime [53], in Fig. 4 we

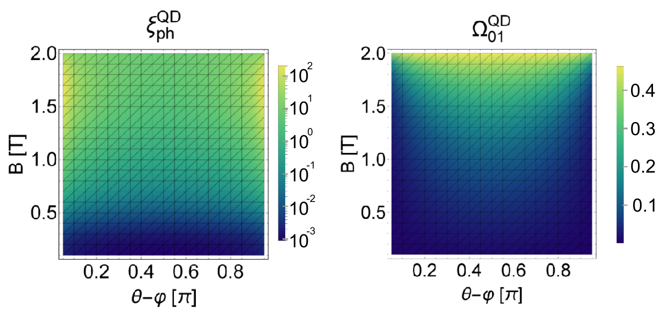


FIG. 4. For the QD confining potential, dependence of the figure of merit ξ_{ph}^{QD} (given in dimensionless unit 7.25 $\frac{m}{V} \times E_0$) and Rabi frequency (in eE_0x_0/h units) on the relative angle $\theta - \varphi \in (0.05, 0.95)\pi$ and magnetic field strength $B \in (0.1, 2)$ T is presented.

present the dependence of ξ_{ph}^{QD} on $B \in (0.1, 2)$ T and $\theta - \varphi \in (0.05, 0.95)\pi$. The restricted $\theta - \varphi$ domain plotted is due to the *a priori* exclusion of $\theta - \varphi = 0, \pi$ values ($\Gamma_{01}^{QD} = 0$ in these situations). Plots show that to maximal value of ξ_{ph}^{QD} correspond relative angles $\theta - \varphi = 0.05\pi, 0.95\pi$. This result suggests that for $\theta - \varphi$ closer to 0 or π than presented even bigger ξ_{QD} values can be obtained, at the cost of lowering the Rabi frequency. In other words, Γ_{01}^{QD} has a steeper decline to zero than Ω_{01}^{QD}, when $\theta - \varphi$ goes from $\pi/2$ to 0 or π.

Magnetic field orientation isotropy of Γ_o [51] implies that shift from $\theta - \varphi = \pi/2$ increases $\Gamma_o/\Gamma_{01}^{QD}$ also. Thus, in order to maximize ξ, optimization of both ξ_{ph}^{QD} and $\Gamma_o/\Gamma_{01}^{QD}$ is needed. Since at high magnetic fields phonon induced relaxation dominates [51], deviation of $\theta - \varphi$ from $\pi/2$ improves the spin qubit quality until $\Gamma_o/\Gamma_{01}^{QD}$ drops below 1. This sets up the optimal magnetic field orientation.

Finally, we compare the impacts of DQD and QD potentials on the spin qubit quality. As discussed in Sec. III A, Rabi frequency in the DQD case can be three orders of magnitude greater than in the QD case. Enhanced Rabi frequency suggests that SOC effects are more pronounced; thus, phonon induced spin relaxation rate should be enhanced. When compared to the QD case, an increase of Γ_{01}^{DQD} followed by the negative trend of ξ_{ph}^{DQD} ensures that spin qubit quality decreases; symmetric DQD confinements give the poorest results, while highly asymmetric DQD potentials provide similar values as for QD gating.

IV. CONCLUSIONS

We have investigated the influence of gating potentials, magnetic field strength and orientation on Rabi frequency and spin relaxation rate in a single electron InSb nanowire spin qubit. Due to the strong Landé g factor, we were able to show that InSb spin qubit can operate in the regime in which deformation potential of acoustic phonons dominate relaxation rate. Qualitatively new behavior of spin relaxation rate comes from the confinement perpendicular to the nanowire axis, offering a new regime in which spin qubit can successfully operate. We have shown that gating potential has a crucial role in enabling such a situation, additionally pointing out simple harmonic potential as beneficial for the optimal definition of a spin qubit. Although presented for InSb nanowire spin qubits, conclusions remain valid for spin qubits in other materials with a strong g factor. Thus, modifications of g due to different effects, e.g., strong in-plane magnetic field [54], do not interfere with the conclusions stated in this work.

ACKNOWLEDGMENTS

This research was funded by the Ministry of Education, Science, and Technological Development of the Republic of Serbia under projects ON171035 and ON171027 and the National Scholarship Programme of the Slovak Republic (ID 28226).

APPENDIX A: DERIVATION OF THE EFFECTIVE ONE-DIMENSIONAL HAMILTONIAN

Here we derive the effective one-dimensional Hamiltonian H of the electron in an InSb quantum wire, by averaging the three-dimensional kinetic energy term T_{3D} and two-dimensional spin-orbit Hamiltonian H_{so}^{2D} over y and z direction. Thus, we start from the three-dimensional Hamiltonian

$$H_{3D} = T_{3D} + V(x) + H_{so}^{2D} + H_z, \tag{A1}$$

where $T_{3D} = \sum_{i=x,y,z} P_i^2/2m^*$ ($P_i = p_i + eA_i$),

$$H_{so}^{2D} = \alpha_R(P_x\sigma_y - P_y\sigma_x) + \alpha_D(P_x\sigma_x - P_y\sigma_y), \tag{A2}$$

while $V(x)$ and H_z are the gating potential and the Zeeman term, defined in Eq. (4) and Eqs. (5) and (6), respectively. The choice of the vector potential components $A_x = -Bz\sin\theta$, $A_y = 0$, $A_z = -By\cos\theta$ is such that it corresponds to the applied in-plane magnetic field $\mathbf{B} = B(\cos\theta, \sin\theta, 0)$. After averaging the kinetic energy operator over the y and z direction using the ground state wave function $\Psi(y,z) = \psi(y)\psi(z)$, we get

$$\langle T\rangle = \frac{p_x^2}{2m^*} - \frac{eB\langle z\rangle\sin\theta}{m^*}p_x + \left[\frac{\langle p_y^2\rangle}{2m^*} + \frac{\langle(p_z - eBy\cos\theta)^2\rangle}{2m^*} + \frac{e^2B^2\sin^2\theta\langle z^2\rangle}{2m^*}\right]. \tag{A3}$$

In the previous equation, only the first and second term affect the dynamics in the x direction, while all terms in the square brackets can be considered the constant shift of energy and, therefore, can be neglected.

Next, effective one-dimensional spin-orbit interaction Hamiltonian is equal to

$$\begin{aligned}\langle H_{so}\rangle &= \alpha_R((p_x - eB\langle z\rangle\sin\theta)\sigma_y - \langle p_y\rangle\sigma_x)\\ &\quad+\alpha_D((p_x - eB\langle z\rangle\sin\theta)\sigma_x - \langle p_y\rangle\sigma_y)\\ &= (p_x - eB\langle z\rangle\sin\theta)(\alpha_R\sigma_y + \alpha_D\sigma_x),\end{aligned} \tag{A4}$$

where we have used the fact that expectation value of the momentum p_y, $\langle p_y\rangle = \int_{-\infty}^{\infty}\Psi^*(y,z)p_y\Psi(y,z)$, is explicitly equal to zero.

A further simplification of the effective Hamiltonian can be made by neglecting the term $eB\langle z\rangle\sin\theta\, p_x/m^*$ from Eq. (A3) and $eB\langle z\rangle\sin\theta$ from Eq. (A4). Assuming that intensity of p_x

is proportional to $\hbar/x_0$, magnetic field dependent terms can be neglected if the relation

$$\frac{\hbar}{x_0} \gg eB\langle z\rangle \tag{A5}$$

is satisfied. More concretely, when the $\hbar/x_0$ is for a factor of 10 stronger than the magnetic field dependent term, orbital effects of the magnetic field are small and can be discarded. In our calculations, the magnetic field strengths of interest are up to 3 T, yielding the relation for the z expectation value

$$\langle z\rangle \leqslant 0.1\frac{\hbar}{ex_0 \times 3\ \mathrm{T}} \tag{A6}$$

that has to be satisfied to successfully operate in this regime. As discussed in Sec. II, the wave function $\psi(z)$ is dependent on the strength of the applied electric field E: with the increase of the electric field strength $\langle z\rangle$ increases. In other words, the strength of the electric field is limited from above. Numerical estimate for the critical value of electric field is 6.5×10^6 V/m, going to be used in our numerical calculations. Under these assumptions, the effective one-dimensional Hamiltonian resembles the one defined in Eq. (1), used in the rest of the paper.

APPENDIX B: NUMERICAL SOLUTION OF THE ONE-DIMENSIONAL SCHRÖDINGER EQUATION

In order to find eigenvectors and eigenenergies of the Hamiltonian H, given in Eq. (1), numerical diagonalization is performed. After defining orbital and spin-orbit lengths as x_0 and $x_{\mathrm{so}} = \hbar/m\alpha$, respectively, such that $x = x_0 u$, where u is dimensionless variable, H can be written in the following form:

$$H = \frac{\hbar^2}{2m^* x_0^2} H_{\mathrm{red}}. \tag{B1}$$

Eigenvectors of H are the same as of H_{red}, while eigenvalues of H and H_{red} differ for the factor $\hbar^2/2m^* x_0^2$, having the energy units. The benefits of using H_{red} instead of H stems from the transfer into dimensionless units, more suitable for numerical manipulation. The concrete form of H_{red} is equal to

$$H_{\mathrm{red}} = -\frac{d^2}{du^2} - 2i\frac{x_0}{x_{\mathrm{so}}}\mathbf{a}\cdot\boldsymbol{\sigma}\frac{d}{du} + V_{\mathrm{eff}}(u) + g_{\mathrm{eff}}\boldsymbol{n}\cdot\boldsymbol{\sigma}, \tag{B2}$$

where g_{eff} and $V_{\mathrm{eff}}(u)$ are effective Landé factor and effective potential, respectively,

$$g_{\mathrm{eff}} = g\frac{m^* x_0^2 \mu_B B}{\hbar^2}, \qquad V_{\mathrm{eff}}(u) = \frac{2m^* x_0^2}{\hbar^2} V(x_0 u), \tag{B3}$$

while vectors $\mathbf{a}$ and $\mathbf{n}$ are spin-orbit and magnetic field unit vectors, respectively, defined in the main text. The form of effective potential depends on the choice of gating potential (5) and (6), while effective Landé factor is linearly dependent on the magnetic field strength B.

To numerically solve the eigenproblem of H_{red}, orbital space is discretized with an uniform grid. First and second derivative of a wave function are approximated by finite difference uniform grid formulas [55]

$$\frac{d\psi(u)}{du} = \frac{\psi_{-4}}{280h} - \frac{4\psi_{-3}}{105h} + \frac{\psi_{-2}}{5h} - \frac{4\psi_{-1}}{5h} - \frac{\psi_4}{280h} + \frac{4\psi_3}{105h} - \frac{\psi_2}{5h} + \frac{4\psi_1}{5h} + O(h^8), \tag{B4}$$

$$\frac{d^2\psi(u)}{du^2} = -\frac{\psi_{-4}}{560h^2} + \frac{8\psi_{-3}}{315h^2} - \frac{\psi_{-2}}{5h^2} + \frac{8\psi_{-1}}{5h^2} - \frac{205\psi_0}{72h^2} - \frac{\psi_4}{560h^2} + \frac{8\psi_3}{315h^2} - \frac{\psi_2}{5h^2} + \frac{8\psi_1}{5h^2} + O(h^8), \tag{B5}$$

with accuracy to the h^8 order, where h is the uniform grid step. By definition, $\psi_{\pm n} = \psi(u \pm nh)$ represent wave functions shifted in the left/right $(-/+)$ direction of the coordinate space for nh.

Uniform grid formulas allow us to represent the Hamiltonian as a square matrix. Effective potential is represented as a diagonal matrix, while matrix representation of the first and second order derivative have nondiagonal terms in addition. Since H_{red} is dependent on spin degrees of freedom also, the orbital part of the Hamiltonian is trivially extended in the spin space. Also, the Zeeman Hamiltonian is trivially extended in the orbital space, while the matrix form of the spin-orbit Hamiltonian is obtained as a tensor product of the first derivative matrix and spin Hamiltonian $\mathbf{a}\cdot\boldsymbol{\sigma}$.

In the QD case, harmonic potential is centered at $u = 0$, while in the case of DQD potential numerical calculations assumed each QD center range from $u = \pm 1/2$ to $u = \pm 2$. We have checked that for all studied situations the choice of u from the interval $(-8, 8)$ is enough to capture the smooth decline of the orbital wave function to 0 at $u = \pm 8$. Also, the division of the orbital space into $N = 2000$ parts was enough to ensure convergence of the results, i.e., for the increase of N to 4000 the relative difference between the results is below 10^{-4}.

[1] C. H. Bennett and D. P. DiVincenzo, Nature (London) **404**, 247 (2000).

[2] F. H. L. Koppens, C. Buizert, K. J. Tielrooij, I. T. Vink, K. C. Nowack, T. Meunier, L. P. Kouwenhoven, and L. M. K. Vandersypen, Nature (London) **442**, 766 (2006).

[3] D. Press, T. D. Ladd, B. Zhang, and Y. Yamamoto, Nature (London) **456**, 218 (2008).

[4] V. N. Golovach, M. Borhani, and D. Loss, Phys. Rev. B **74**, 165319 (2006).

[5] K. C. Nowack, F. H. L. Koppens, Yu. V. Nazarov and L. M. K. Vandersypen, Science **318**, 1430 (2007).

[6] R. Brunner, Y.-S. Shin, T. Obata, M. Pioro-Ladriére, T. Kubo, K. Yoshida, T. Taniyama, Y. Tokura, and S. Tarucha, Phys. Rev. Lett. **107**, 146801 (2011).

[7] E. Kawakami, P. Scarlino, D. R. Ward, F. R. Braakman, D. E. Savage, M. G. Lagally, M. Friesen, S. N. Coppersmith, M. A. Eriksson, and L. M. K. Vandersypen, Nat, Nanotechnol. **9**, 666 (2014).

[8] K. Takeda, J. Yoneda, T. Otsuka, T. Nakajima, M. R. Delbecq, G. Allison, Y. Hoshi, N. Usami, K. M. Itoh, S. Oda, T. Kodera, and S. Tarucha, npj Quantum Inf. **4**, 54 (2018).
[9] D. V. Khomitsky, L. V. Gulyaev, and E. Ya. Sherman, Phys. Rev. B **85**, 125312 (2012).
[10] D. V. Khomitsky, E. A. Lavrukhina, and E. Ya. Sherman, Phys. Rev. B **99**, 014308 (2019).
[11] A. V. Khaetskii and Y. V. Nazarov, Phys. Rev. B **61**, 12639 (2000).
[12] A. V. Khaetskii and Y. V. Nazarov, Phys. Rev. B **64**, 125316 (2001).
[13] D. Mozyrsky, S. Kogan, V. N. Gorshkov, and G. P. Berman, Phys. Rev. B **65**, 245213 (2002).
[14] C. Tahan, M. Friesen, and R. Joynt, Phys. Rev. B **66**, 035314 (2002).
[15] E. Ya. Sherman and D. J. Lockwood, Phys. Rev. B **72**, 125340 (2005).
[16] P. Stano and J. Fabian, Phys. Rev. B. **72**, 155410 (2005).
[17] F. Baruffa, P. Stano, and J. Fabian, Phys. Rev. Lett. **104**, 126401 (2010).
[18] X. Linpeng, T. Karin, M. V. Durnev, R. Barbour, M. M. Glazov, E. Ya. Sherman, S. P. Watkins, S. Seto, and K.-M. C. Fu, Phys. Rev. B **94**, 125401 (2016).
[19] V. N. Stavrou, J. Phys.: Condens. Matter **29**, 485301 (2017).
[20] V. N. Stavrou, J. Phys.: Condens. Matter **30**, 455301 (2018).
[21] Z.-H. Liu, R. Li, X. Hu, and J. Q. You, Sci. Rep. **8**, 2302 (2018).
[22] O. Malkoc, P. Stano, and D. Loss, Phys. Rev. B **93**, 235413 (2016).
[23] J. I. Climente, A. Bertoni, G. Goldoni, M. Rontani, and E. Molinari, Phys. Rev. B **75**, 081303(R) (2007).
[24] D. Chaney and P. A. Maksym, Phys. Rev. B **75**, 035323 (2007).
[25] E. Reyes-Gómez, N. Porras-Montenegro, C. A. Perdomo-Leiva, H. S. Brandi, and L. E. Oliveira, J. Appl. Phys. **104**, 023704 (2008).
[26] H. A. Nilsson, P. Caroff, C. Thelander, M. Larsson, J. B. Wagner, L.-E. Wernersson, L. Samuelson, and H. Q. Xu, Nano Lett. **9**, 3151 (2009).
[27] M. Trif, V. N. Golovach, and D. Loss, Phys. Rev. B **77**, 045434 (2008).
[28] S. Nadj-Perge, S. M. Frolov, E. P. A. M. Bakkers, and L. P. Kouwenhoven, Nature (London) **468**, 1084 (2010).
[29] S. Nadj-Perge, V. S. Pribiag, J. W. G. van den Berg, K. Zuo, S. R. Plissard, E. P. A. M. Bakkers, S. M. Frolov, and L. P. Kouwenhoven, Phys. Rev. Lett. **108**, 166801 (2012).
[30] R. Li, J. Q. You, C. P. Sun, and F. Nori, Phys. Rev. Lett. **111**, 086805 (2013).
[31] R. Li and J. Q. You, Phys. Rev. B **90**, 035303 (2014).
[32] G. Dresselhaus, Phys. Rev. **100**, 580 (1955).
[33] E. I. Rashba, Fiz. Tv. Tela (Leningrad) **2**, 1224 (1960); Sov. Phys. Solid State **2**, 1109 (1960).
[34] M. Raith, P. Stano, and J. Fabian, Phys. Rev. B **83**, 195318 (2011).
[35] M. Raith, P. Stano, and J. Fabian, Phys. Rev. B **86**, 205321 (2012).
[36] S. Wu, L. Cheng, H. Yu, and Q. Wang, Phys. Lett. A **382**, 1922 (2018).
[37] F. N. M. Froning, M. K. Rehmann, J. Ridderbos, M. Brauns, F. A. Zwanenburg, A. Li, E. P. A. M. Bakkers, D. M. Zumbühl, and F. R. Braakman, Appl. Phys. Lett. **113**, 073102 (2018).
[38] D. Fan, S. Li, N. Kang, P. Caroff, L. B. Wang, Y. Q. Huang, M. T. Deng, C. L. Yu, and H. Q. Xu, Nanoscale **7**, 14822 (2015).
[39] M. P. Nowak and B. Szafran, Phys. Rev. B **89**, 205412 (2014).
[40] Z.-H. Liu, O. Entin-Wohlman, A. Aharony, and J. Q. You, Phys. Rev. B **98**, 241303(R) (2018).
[41] M. Cardona, N. E. Christensen, and G. Fasol, Phys. Rev. B **38**, 1806 (1988).
[42] H. A. Nilsson, O. Karlström, M. Larsson, P. Caroff, J. N. Pedersen, L. Samuelson, A. Wacker, L. E. Wernersson, and H. Q. Xu, Phys. Rev. Lett. **104**, 186804 (2010).
[43] Z.-H. Liu and R. Li, AIP Adv. **8**, 075115 (2018).
[44] R. de Sousa and S. Das Sarma, Phys. Rev. B **68**, 155330 (2003).
[45] J. I. Climente, A. Bertoni, G. Goldoni, and E. Molinari, Phys. Rev. B **74**, 035313 (2006).
[46] S. E. Hebboul and J. P. Wolfe, Phys. Rev. B **34**, 3948 (1986).
[47] O. Madelung, Ed., *Intrinsic Properties of Group IV Elements and III-V, II-VI and I-VII Compounds*, Landolt-Bornstein, New Series Vol. 22 (Springer, Berlin, 1987), Chap. 2.15, p. 130.
[48] C. F. Destefani and S. E. Ulloa, Phys. Rev. B **72**, 115326 (2005).
[49] I. M. Tsidilkovskii and K. M. Demchuk, Phys. Status Solidi B **44**, 293 (1971); K. M. Demchuk and I. M. Tsidilkovskii, *ibid.* **82**, 59 (1977); K. Tukioka, Jpn. J. Appl. Phys. **30**, 212 (1991).
[50] J. L. Cheng, M. W. Wu, and C. Lü, Phys. Rev. B **69**, 115318 (2004).
[51] L. C. Camenzind, L. Yu, P. Stano, J. D. Zimmerman, A. C. Gossard, D. Loss, and D. M. Zumbühl, Nat. Commun. **9**, 3454 (2018).
[52] Although dependence of relaxation rates on the magnetic field strength are presented for the special $\theta - \varphi = \pi/2$ magnetic field orientation, conclusion remain valid for different $\theta - \varphi$ angles.
[53] Low fields correspond to weak Zeeman splitting. In this case low temperatures are needed, such that Zeeman splitting is above the thermal broadening; temperature dependence analysis is beyond the scope of the presented work.
[54] P. Stano, C.-H. Hsu, M. Serina, L. C. Camenzind, D. M. Zumbühl, and D. Loss, Phys. Rev. B **98**, 195314 (2018).
[55] B. Fornberg, Math. Comput. **51**, 699 (1988).

Control of a spin qubit in a lateral GaAs quantum dot based on symmetry of gating potential

Pavle Stipsić [1] and Marko Milivojević [2,3]

[1] Faculty of Physics, University of Belgrade, Studentski trg 12, 11001 Belgrade, Serbia
[2] NanoLab, QTP Center, Faculty of Physics, University of Belgrade, Studentski trg 12, 11001 Belgrade, Serbia
[3] Department of Theoretical Physics and Astrophysics, Faculty of Science, P. J. Šafárik University, Park Angelinum 9, 040 01 Košice, Slovak Republic

(Received 14 January 2020; revised manuscript received 25 February 2020; accepted 10 March 2020; published 6 April 2020)

We study the influence of quantum dot symmetry on the Rabi frequency and phonon-induced spin relaxation rate in a single-electron GaAs spin qubit. We find that anisotropic dependence on the magnetic field direction is independent of the choice of the gating potential. Also, we discover that relative orientation of the quantum dot, with respect to the crystallographic frame, is relevant in systems with $\mathbf{C}_{1\mathrm{v}}$, $\mathbf{C}_{2\mathrm{v}}$, or $\mathbf{C}_n$ ($n \neq 4r$) symmetry. To demonstrate the important impact of the gating potential shape on the spin qubit lifetime, we compare the effects of an infinite-wall equilateral triangle, square, and rectangular confinement with the known results for the harmonic potential. In the studied cases, enhanced spin qubit lifetime is revealed, reaching almost six orders of magnitude increase for the equilateral triangle gating.

DOI: 10.1103/PhysRevB.101.165302

I. INTRODUCTION

Every quantum two-level system can act as the quantum bit, a basic unit of quantum information processing [1,2]. Among different solid-state implementations of the qubit system [3–6], single-electron spin in a semiconductor quantum dot (QD) can be used to achieve the task. In order to manipulate spins of charge carriers embedded inside a semiconductor material electrically, through electric dipole spin resonance (EDSR) [7], the presence of spin-orbit interaction (SOI) is obligatory.

Besides its positive effect in EDSR-based schemes [8–16], SOI enables the electron-phonon coupling-mediated transitions between the qubit states [17–20], affecting the spin qubit lifetime. To suppress the coupling to phonons, different approaches like the optimization of the QD design [21,22] or control of the system size [23] were suggested. The observed anisotropy of the spin relaxation rate on the in-plane magnetic field orientation [24] offered another playground for fine-tuning the spin qubit's desired properties. In circular QDs, this is the only degree of freedom accessible in the optimization of the spin qubit, while for the elliptical confining potential [22,25–27] orientation of the QD potential with respect to the crystallographic frame can be used as the tuning parameter.

Evidently, different symmetry of the gating potential [28] is the main reason for the observed behavior. But to what extent can the potential symmetry alter the basic properties of the electrically controlled spin qubit? To address this question, we have performed a general analysis valid for the lateral GaAs QD system with $\mathbf{C}_{n\mathrm{v}}$ or $\mathbf{C}_n$ symmetry of the gating potential. Besides the expected anisotropy on the magnetic field orientation, we were able to find potential symmetries for which the QD orientation with respect to the crystallographic frame can act as another control parameter of the spin qubit characteristics. With our theory, we offer a simple and efficient way to determine the impact of the gating potential on the Rabi frequency and spin relaxation rate. This is shown in the example of anisotropic and isotropic harmonic potential, as well as for the infinite-wall equilateral triangle, square, and rectangular potential.

This paper is organized as follows. In Sec. II we define a single-electron GaAs spin qubit model. In Sec. III we define the dipole moment of the electrically controlled spin qubit that describes both the Rabi frequency and SOI-induced spin relaxation rate mediated by acoustic phonons. In Sec. IV we present the main results of the paper: analytical expressions for the dipole moment in the case of the gating potential with $\mathbf{C}_{n\mathrm{v}}$ or $\mathbf{C}_n$ symmetry. In Sec. V, to illustrate the impact of the gating potential on the spin qubit lifetime we use the obtained expressions to compare the influence of the harmonic confinement with an infinite-wall equilateral triangle, square, and rectangular potential. In Sec. VI we give our conclusions.

II. DYNAMICS OF THE LATERAL QD

We start with the Hamiltonian describing the lateral dynamics of a single-electron in the GaAs material,

$$H = H_0 + H_\mathrm{z} + H_\mathrm{so} = \frac{p_x^2 + p_y^2}{2m^*} + V(x, y) + H_\mathrm{z} + H_\mathrm{so}, \quad (1)$$

where p_x and p_y are the momentum operators, m^* is the effective mass ($m^* = 0.067m_e$ for GaAs, m_e is the electron mass), while $V(x, y)$ is the gating potential used to localize the electron in a QD. In the lateral system, symmetries that can be present are the n-fold rotational symmetry and the vertical mirror plane symmetry σ_v. For simplicity, we assume that σ_v coincides with the yz plane of the QD coordinate frame (see Fig. 1). Thus, we assume a general form of the orbital Hamiltonian H_0 that has a $\mathbf{C}_{n\mathrm{v}}$ or $\mathbf{C}_n$ ($n = \infty$ also) symmetry.

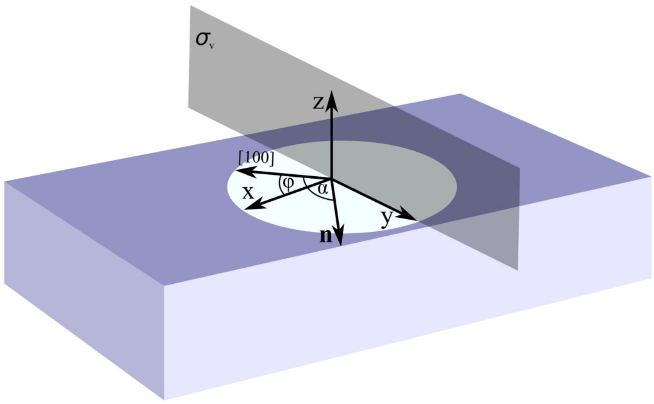


FIG. 1. A schematic view of the GaAs lateral QD. The y axis of the QD reference frame coincides with the vertical mirror plane symmetry σ_v. We define the angle between the chosen x axis and the crystallographic [100] axis as φ. The magnetic field is aligned along the $\mathbf{n}$ direction, forming an angle α with the [100] direction.

Due to the symmetry, eigenenergies and eigenvectors of H_0 can be classified according to the irreducible representations (IRs) of a given point-group symmetry.

Besides H_0, in Eq. (1) the Zeeman term H_z appears, describing the coupling of spin and magnetic field:

$$H_z = g\mu_B \boldsymbol{B} \cdot \boldsymbol{s}, \tag{2}$$

where g is the effective Landé factor ($g \approx -0.44$ for GaAs), μ_B is the Bohr magneton, $\boldsymbol{s} = 1/2\boldsymbol{\sigma}$ is the electron's spin, and $\boldsymbol{B} = B\boldsymbol{n}$ is the in-plane magnetic field forming an angle α with the crystallographic [100] axis. In Eq. (1) we have neglected the orbital effects of the in-plane magnetic field. This is a reasonable assumption for the magnetic field strength weaker than a few teslas [29]. In the case of the magnetic field applied in the z direction, orbital effects would be much more pronounced [29].

Eigenstates of $H_0 + H_z$ can be written in a direct product form $|\Psi_i\pm\rangle = |\Psi_i\rangle \otimes |\pm\rangle$, where $|\Psi_i\rangle$ corresponds to the eigenvectors of the Hamiltonian H_0 with an energy ϵ_i, while $|\pm\rangle$ represents eigenvectors of H_z with spin projection parallel or antiparallel to the magnetic field direction and an eigenenergy $\pm g\mu_B B/2$, respectively. The effect of H_z on the eigenspectra of H_0 can be seen as the splitting of H_0 eigenenergies into two branches with an energy difference $|g|\mu_B B$. In this work, we assume that $|g|\mu_B B$ is much weaker than the energy difference between the ground and the first excited state of the orbital Hamiltonian H_0.

Besides H_0 (H_z) that acts trivially in the spin (orbital) space, the SOI Hamiltonian does not commute with $H_0 + H_z$. It consists of two terms, Dresselhaus [30] and Rashba [31]: the Dresselhaus term exists due to the bulk inversion asymmetry of the structure, while the Rashba term is present when an electric field perpendicular to the growth direction is applied. The form of spin-orbit coupling is dependent on the structure's symmetry. For GaAs, having the zincblende structure, the SOI Hamiltonian is equal to

$$H_{so} = 2\alpha_d\left(p_y^c s_y - p_x^c s_x\right) + 2\alpha_r\left(p_x^c s_y - p_y^c s_x\right), \tag{3}$$

where α_r and α_d are Rashba and Dresselhaus coupling constants, while p_x^c and p_y^c are momentum operators in the [100] and [010] crystallographic directions, respectively. The electron spin is locked to the crystal momentum, since the potential trap confines electron of the crystal. Thus, an electron in a QD inherits the features of the crystal for which the crystal momentum is only appropriately defined. However, we have the choice to define the x axis of our coordinate frame independently on the crystallographic [100] direction. Assuming that the angle between them is φ, p_x^c and p_y^c should be written in terms of momentum operators in the chosen frame: $p_x^c = p_x \cos\varphi - p_y \sin\varphi$, $p_y^c = p_x \sin\varphi + p_y \cos\varphi$.

The spin-orbit Hamiltonian can be written in a different form using the Rashba $l_r = \hbar^2/2m\alpha_r$ and Dresselhaus $l_d = \hbar^2/2m\alpha_d$ precession lengths:

$$H_{so} = \hbar\left(\frac{p_y^c s_y - p_x^c s_x}{m^* l_d} + \frac{p_x^c s_y - p_y^c s_x}{m^* l_r}\right). \tag{4}$$

To compare the ratio of the spin-orbit precession length and the orbital confinement length l, we redefine l_r and l_d in terms of the overall spin-orbit length l_{so} and the spin-orbit angle ν:

$$l_d^{-1} = l_{so}^{-1} \sin\nu, \quad l_r^{-1} = l_{so}^{-1} \cos\nu. \tag{5}$$

Since we assume no doping of the GaAs material [32], l_{so} can be considered constant. Moreover, the relation $l_{so} \gg l$ [29,33] is satisfied in GaAs QDs, meaning that SOI can be treated as a perturbation.

Without SOI, qubit states can be defined as $|\Psi_0\pm\rangle = |\Psi_0\rangle \otimes |\pm\rangle$, where $|\Psi_0\rangle$ corresponds to the ground state of the spin-independent Hamiltonian H_0. Because SOI can be treated on the level of a perturbation, we calculate first-order corrections of the qubit states due to spin-orbit coupling. Since it is known that the standard perturbation technique badly incorporates the spin-orbit-induced corrections [34,35], we follow the procedure explained in Ref. [22]: the Hamiltonian H is transformed using the unitary operator $U = \exp(\mathrm{i}\mathbf{n}_{so} \cdot \boldsymbol{s})$, defined with the help of the position-dependent spin-orbit vector $\mathbf{n}_{so} = l_{so}^{-1}(r_1 \sin\nu + r_2 \cos\nu, -r_1 \cos\nu - r_2 \sin\nu, 0)$:

$$UHU^\dagger = H_0 + H_z + H_{so}^{eff}. \tag{6}$$

The unitary operator U does not change the orbital and Zeeman Hamiltonian. On the other hand, the SOI Hamiltonian H_{so} is transformed into

$$H_{so}^{eff} = g\mu_B(\mathbf{n}_{so} \times \mathbf{B}) \cdot \boldsymbol{s} - \frac{\hbar^2}{4m^* l_{so}^2}(1 + 2l_z s_z \cos 2\nu), \tag{7}$$

where $l_z = -\mathrm{i}(r_1\partial_{r_2} - r_2\partial_{r_1})$ is the orbital angular momentum. Using H_{so}^{eff}, the first-order correction of the qubit states can be written as

$$\delta|\Psi_0\sigma'\rangle = U \sum_{i\neq 0,\sigma''} \frac{\langle\Psi_i\sigma''|H_{so}^{eff}|\Psi_0\sigma'\rangle}{\epsilon_0 - \epsilon_i + \frac{\sigma'-\sigma''}{2} g\mu_B B}|\Psi_i\sigma''\rangle, \tag{8}$$

where the sum over $i \neq 0$ corresponds to all orbital eigenvectors $|\Psi_i\rangle$ different from the ground state $|\Psi_0\rangle$, while $\sigma'' = \pm$.

The lateral QD model is valid if the electron dynamics in the z direction is suppressed; i.e., an electron is always in the ground state. Thus, we assume that confinement length in the z direction is much stronger than in the xy plane. The Hamiltonian describing the quantum confinement in the z

direction is equal to $H(z) = p_z^2/2m^* + V(z)$, where $V(z) = eE_0 z$ for $z \geqslant 0$ and $V(z) = \infty$ for $z < 0$. To this Hamiltonian corresponds the following ground state (for $z > 0$) [36]:

$$\Psi_0(z) = 1.4261\sqrt{\chi}\mathrm{Ai}(\chi z - 2.3381), \quad (9)$$

where Ai is the Airy function, while $\chi = (2m^* eE_0/\hbar^2)^{1/3}$ is the inverse of the characteristic length $z_0 = 1.5587/\chi$ in the z direction.

In order to simplify the notation, in the rest of the paper we assume that $|\uparrow\rangle$ and $|\downarrow\rangle$ represent SOI corrected qubit states in the xy plane, while $|\Psi_\uparrow\rangle = |\uparrow\rangle\Psi_0(z)$ and $|\Psi_\downarrow\rangle = |\downarrow\rangle\Psi_0(z)$ correspond to wavefunctions of the qubit states in three dimensions.

III. RABI FREQUENCY AND PHONON-INDUCED SPIN RELAXATION RATE

Electrical control of the spin qubit is possible by applying the in-plane oscillating electric field $\mathbf{E}\cos\omega t$, resulting in the Rabi Hamiltonian $H_R = e\mathbf{E}\cdot\mathbf{r}\cos(\omega t)$. The Rabi frequency, measuring the speed of the single-qubit rotations, is equal to $\Omega = e/\hbar|\mathbf{E}\cdot\langle\uparrow|\mathbf{r}|\downarrow\rangle|$, where

$$\mathbf{d}_{\uparrow\downarrow} = \langle\uparrow|\mathbf{r}|\downarrow\rangle \quad (10)$$

is the dipole moment (in e units), present due to the SOI-induced spin mixing mechanism. Misalignment of the applied field direction and the dipole moment leads to a trivial suppression of the Rabi frequency. Since it is beneficial to increase the Rabi frequency as much as possible, the electric field should be applied in the direction of the dipole moment. Thus, for fixed $|\mathbf{E}|$, the maximal value $\max(\Omega) = \Omega_{\uparrow\downarrow}$ of the Rabi frequency

$$\Omega_{\uparrow\downarrow} = \frac{e}{\hbar}|\mathbf{E}||\mathbf{d}_{\uparrow\downarrow}| \quad (11)$$

is completely dependent on the strength of the dipole moment.

Since spin-phonon interaction in semiconductor QDs is irrelevant [17], unlike donor-bound electrons in direct band-gap semiconductors [37], only electron-phonon-induced transition between the qubit states should be considered in the study of spin relaxation. Electron-phonon coupling is triggered by the SOI-induced admixture mechanism, being highly dependent on the symmetry of the gating potential [37]. We determine the rate of spin relaxation at $T = 0$ from the Fermi golden rule,

$$\Gamma_{\uparrow\downarrow} = \frac{2\pi}{\hbar}\sum_{\nu\mathbf{q}}|M_\nu(\mathbf{q})|^2|\langle\Psi_\uparrow|e^{i\mathbf{q}\cdot\mathbf{r}_c}|\Psi_\downarrow\rangle|^2\delta(\epsilon_{\uparrow\downarrow} - \hbar\omega_{\nu\mathbf{q}}), \quad (12)$$

assuming the dominant contribution of acoustic phonons, having an energy $\hbar\omega_{\nu\mathbf{q}}$, equal to the level separation between the qubit states, $\epsilon_{\uparrow\downarrow} = |g|\mu_B B$. For magnetic field strengths up to a few teslas, relevant for this work, the linear dependence of phonon frequencies on the crystal wave vector length can be used, $\omega_{\nu\mathbf{q}} = c_\nu|\mathbf{q}|$, giving us $|\mathbf{q}| = |g|\mu_B B/\hbar c_\nu$ [38].

The geometric factor $|M_\nu(\mathbf{q})|^2$ is dependent on the phonon mode, longitudinal acoustic (LA) or transverse acoustic (TA). The longitudinal geometric factor [39]

$$|M_{\mathrm{LA}}(\mathbf{q})|^2 = \frac{\hbar D^2}{2\rho c_{\mathrm{LA}}V}|\mathbf{q}| + \frac{32\pi^2\hbar(eh_{14})^2}{\epsilon^2\rho c_{\mathrm{LA}}V}\frac{(3q_xq_yq_z)^2}{|\mathbf{q}|^7} \quad (13)$$

depends on both D and h_{14}, representing the deformation and piezoelectric constant, respectively. On the other hand, the transverse geometric factor [39]

$$|M_{\mathrm{TA}}(\mathbf{q})|^2 = 2\frac{32\pi^2\hbar(eh_{14})^2}{\epsilon^2\rho c_{\mathrm{TA}}V} \times\left|\frac{q_x^2q_y^2 + q_x^2q_z^2 + q_y^2q_z^2}{|\mathbf{q}|^5} - \frac{(3q_xq_yq_z)^2}{|\mathbf{q}|^7}\right| \quad (14)$$

is dependent on the piezoelectric constant solely. Other parameters for the GaAs material are [22,34] $c_{\mathrm{LA}} = 5290$ m/s, $c_{\mathrm{TA}} = 2480$ m/s, $\rho = 5300$ kg/m^3, $D = 7$ eV, $eh_{14} = 1.4 \times 10^9$ eV/m, and $\epsilon = 12.9$.

Finally, in Eq. (12) both the lateral and the z-direction confinement enter the relaxation rate through the scattering matrix element $|\langle\Psi_\uparrow|e^{i\mathbf{q}\cdot\mathbf{r}_c}|\Psi_\downarrow\rangle|^2$. We employ the dipole approximation $e^{i\mathbf{q}\cdot\mathbf{r}_c} \approx 1 + i\mathbf{q}\cdot\mathbf{r}_c$, justified for magnetic field strengths below a few teslas.

To summarize, the phonon-induced relaxation rate can be divided into three separate channels: the deformation phonons $\Gamma_{\uparrow\downarrow}^{\mathrm{def}}$, the longitudinal piezoelectric phonons $\Gamma_{\uparrow\downarrow}^{\mathrm{piez,LA}}$, and the transverse piezoelectric phonons $\Gamma_{\uparrow\downarrow}^{\mathrm{piez,TA}}$. In GaAs QDs, $\Gamma_{\uparrow\downarrow}^{\mathrm{piez,TA}}$ is the dominant relaxation channel, being two orders of magnitude stronger than $\Gamma_{\uparrow\downarrow}^{\mathrm{piez,LA}} + \Gamma_{\uparrow\downarrow}^{\mathrm{def}}$ in the dipole approximation regime. Thus, we can identify the total relaxation rate with $\Gamma_{\uparrow\downarrow}^{\mathrm{piez,TA}}$ [40]:

$$\Gamma_{\uparrow\downarrow} = \frac{256\pi(eh_{14})^2(|g|\mu_BB)^3}{105c_{\mathrm{TA}}^5\rho\hbar^4\epsilon^2}\left(1 + \frac{7}{33}K_{\mathrm{TA}}^2z_0^2\right)|\mathbf{d}_{\uparrow\downarrow}|^2, \quad (15)$$

where $K_{\mathrm{TA}} = |g|\mu_BB/\hbar c_{\mathrm{TA}}$. We assume a typical confinement length $l = 10$ nm [29,33] of the GaAs QD in an experimental setup and magnetic field up to a few teslas (see Sec. II). Since confinement in the z direction is much stronger than in the xy plane, $z_0 \ll l$, we conclude that $7K_{\mathrm{TA}}^2z_0^2/33$ is much weaker than 1. In other words, the influence of the confinement in the z direction can be neglected.

Note that $\Gamma_{\uparrow\downarrow}$ is squarely dependent on the absolute value of the dipole moment, meaning that the knowledge of the dipole moment is sufficient to fully explain the behavior of both the Rabi frequency and the spin relaxation rate.

IV. ANALYTICAL EXPRESSION FOR THE DIPOLE MOMENT

Based on the previous conclusion, we come to the main objective: to derive symmetry-allowed expression for the dipole moment. The results can be divided into three cases, according to the system's group symmetry: (1) $\mathbf{C}_{n\mathrm{v}}$ ($n \geqslant 3$) and $\mathbf{C}_{\infty\mathrm{v}}$, (2) $\mathbf{C}_{2\mathrm{v}}$ and $\mathbf{C}_{1\mathrm{v}}$, and (3) $\mathbf{C}_n$ and $\mathbf{C}_\infty$.

A. Dipole moment for systems with $\mathbf{C}_{n\mathrm{v}}$ ($n \geqslant 3$) or $\mathbf{C}_{\infty\mathrm{v}}$ symmetry

To find the SOI-induced perturbative correction of the qubit states, we first rewrite the unitarily transformed SOI

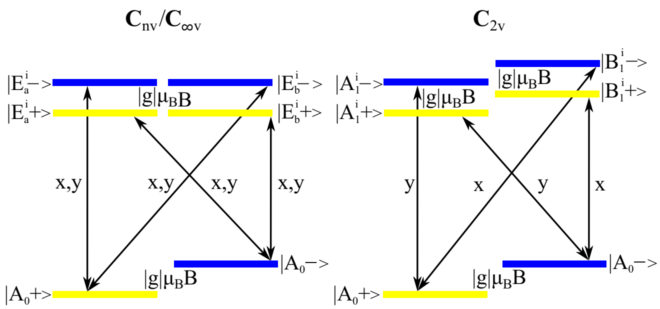


FIG. 2. A schematic view of the first-order perturbation correction of the qubit states $|A_0\pm\rangle$ in the case of $\mathbf{C}_{nv}$ ($\mathbf{C}_{\infty v}$) (left) and $\mathbf{C}_{2v}$ (right) symmetry. In the first case, states that correct the qubit states have twofold orbital degeneracy and transform according to the IR E_1. These states are split by the Zeeman energy $|g|\mu_B B$. The transition between the SOI uncorrected qubit states and the $|E^i_{a,b}\rangle$ states is enabled by the x and y terms from $H^{\text{eff}}_{\text{so}}$. In the second case, orbital states involved in the qubit states correction transform according to IRs A_1 and B_1; the transition is triggered by the terms y and x from $H^{\text{eff}}_{\text{so}}$, respectively.

Hamiltonian in the coordinate frame of the potential,

$$H^{\text{eff}}_{\text{so}} = g\mu_B B s_z(x(\sin(\nu+\varphi)\sin\alpha + \cos(\nu-\varphi)\cos\alpha) + y(\cos(\nu+\varphi)\sin\alpha + \sin(\nu-\varphi)\cos\alpha)), \quad (16)$$

and neglect the second term in Eq. (7), assuming magnetic field strengths $> \mu$T needed to appropriately define the qubit states. For simplicity, we define two factors,

$$v_x = \sin(\nu+\varphi)\sin\alpha + \cos(\nu-\varphi)\cos\alpha, \quad (17)$$

$$v_y = \cos(\nu+\varphi)\sin\alpha + \sin(\nu-\varphi)\cos\alpha, \quad (18)$$

with whose help $H^{\text{eff}}_{\text{so}}$ can be written in a more compact form.

The Hamiltonian $H^{\text{eff}}_{\text{so}}$ is in the orbital space dependent on the coordinates x and y that transform according to the IR E_1. Their symmetry behavior restricts the states that can appear in the perturbative correction of the qubit states. It is simple to check that only states transforming according to the IR E_1 are allowed. This is illustrated in the left-hand panel of Fig. 2.

We label the ground state of the orbital Hamiltonian as $|A_0\rangle$, since the ground state in quantum mechanical systems is of the maximal possible symmetry [41] and it should transform according to the A_0 IR, representing the objects invariant under all group symmetry operations (see Table I). We write two complex conjugate basis vectors of the two-dimensional IR E_1 as $|E^i_a\rangle$ and $|E^i_b\rangle$, where i labels the energy level. Also, we define the energy difference between the excited level and the ground state as $\epsilon^i = \epsilon^i_{ex} - \epsilon_{gr}$.

Due to the negative g factor, the lowest qubit state $|A_0+\rangle = |A_0\rangle \otimes |+\rangle$ is parallel to the magnetic field direction, while $|A_0-\rangle = |A_0\rangle \otimes |-\rangle$ is the qubit state with spin projection antiparallel to the magnetic field direction. The first-order perturbative correction to the qubit states is written as $|\delta A_0\pm\rangle$. Thus, we can write the SOI corrected qubit states as $|\uparrow\downarrow\rangle = |A_0\pm\rangle + |\delta A_0\pm\rangle$, where the normalization factor is omitted as the correction is small. Correspondingly, the dipole moment

TABLE I. For $\mathbf{C}_{nv}$ and $\mathbf{C}_{\infty v}$ symmetry groups, tables of matrices of the corresponding IRs are given [42], tabulated on the generators C_n (R_β) and σ_v, where C_n (R_β) represents a rotation for the angle $2\pi/n$ (β) around the z axis. In the $\mathbf{C}_{nv}$ case, two-dimensional IRs exist if $n \geqslant 3$. In both cases, two-dimensional IRs are written in a complex conjugate basis.

$\mathbf{C}_{nv}$	IR	m	C_n	σ_v
	A_0/B_0	0	1	± 1
	E_m	$(0, \frac{n}{2})$	$\begin{pmatrix} e^{i\frac{2\pi}{n}m} & 0 \\ 0 & e^{-i\frac{2\pi}{n}m} \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
	$A_{\frac{n}{2}}/B_{\frac{n}{2}}$	$\frac{n}{2}$	-1	± 1
$\mathbf{C}_{\infty v}$	IR	m	R_β	σ_v
	A_0/B_0	0	1	± 1
	E_m	$1, 2, \ldots$	$\begin{pmatrix} e^{i\beta m} & 0 \\ 0 & e^{-i\beta m} \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

is equal to

$$\mathbf{d} = \sum_{j=x,y} \langle\uparrow|\mathbf{r}\cdot\boldsymbol{e}_j|\downarrow\rangle\boldsymbol{e}_j = \sum_{j=x,y} (\langle A_0+|\mathbf{r}\cdot\boldsymbol{e}_j|\delta A_0-\rangle\boldsymbol{e}_j + \langle\delta A_0+|\mathbf{r}\cdot\boldsymbol{e}_j|A_0-\rangle\boldsymbol{e}_j). \quad (19)$$

Since $l_{\text{so}} \gg l$, we approximate the unitary operator U with I_2, where I_2 is the identity 2×2 matrix. After noticing that $\langle\pm|s_z|\mp\rangle = -1/2$, $\langle\pm|s_z|\pm\rangle = 0$, we find the SOI-induced corrections of the qubit states

$$|\delta A_0\pm\rangle = \frac{|g|\mu_B B}{2l_{\text{so}}}\sum_i \left(\frac{\langle E^i_a|xv_x + yv_y|A_0\rangle}{\epsilon^i \pm |g|\mu_B B}|E^i_a\mp\rangle + \frac{\langle E^i_b|xv_x + yv_y|A_0\rangle}{\epsilon^i \pm |g|\mu_B B}|E^i_b\mp\rangle\right). \quad (20)$$

Additionally, transition dipole matrix elements are labeled as

$$X^i = \langle E^i_a|x|A_0\rangle, \quad Y^i = \langle E^i_a|y|A_0\rangle. \quad (21)$$

Since the Zeeman splitting is much smaller than the orbital excitation energies, $|g|\mu_B B \ll \epsilon^i$, the approximation $\epsilon^i \pm |g|\mu_B B \approx \epsilon^i$ can be made. Thus, Eq. (20) is transformed into

$$|\delta A_0\pm\rangle = \frac{|g|\mu_B B}{2l_{\text{so}}}\sum_i \left(\frac{X^i v_x + Y^i v_y}{\epsilon^i}|E^i_a\mp\rangle + \frac{(X^i)^* v_x + (Y^i)^* v_y}{\epsilon^i}|E^i_b\mp\rangle\right), \quad (22)$$

where $(X^i)^*$ and $(Y^i)^*$ are the complex conjugates of X^i and Y^i, respectively. Components of the dipole moment can now be written in a more compact form:

$$d_x = \frac{2|g|\mu_B B}{l_{\text{so}}}\sum_i \frac{|X^i|^2 v_x + \text{Re}(X^i(Y^i)^*)v_y}{\epsilon^i},$$

$$d_y = \frac{2|g|\mu_B B}{l_{\text{so}}}\sum_i \frac{|Y^i|^2 v_y + \text{Re}(X^i(Y^i)^*)v_x}{\epsilon^i}, \quad (23)$$

where $\text{Re}(X^i(Y^i)^*)$ stands for the real part of $X^i(Y^i)^*$. Potential dependent parameters that enter Eq. (23) are the transition dipole matrix elements and the excitation energies. Besides them, dipole moment components are dependent on the spin-orbit angle ν, magnetic field angle α, and the angle φ between the [100] crystallographic direction and the x axis.

A further simplification of Eq. (23) stems from the existence of the vertical mirror symmetry σ_v, requiring that $\mathrm{Re}(X^i(Y^i)^*)$ must be zero. This can be proven in a few simple steps. First, we deduce from the matrix of an IR E_1, representing the vertical mirror plane, that σ_v transforms one IR vector into the other, $E_1(\sigma_v)|E^i_{a,b}\rangle = |E^i_{b,a}\rangle$. Furthermore, y remains unchanged, while x acquires a minus sign, leading to the following behavior of the transition matrix elements X^i and Y^i under vertical mirror plane symmetry:

$$X^i \xrightarrow{\sigma_v} -(X^i)^*, \quad Y^i \xrightarrow{\sigma_v} (Y^i)^*. \tag{24}$$

From the previous relations, we conclude that the term $\mathrm{Re}(X^i(Y^i)^*)$ transforms into $-\mathrm{Re}(X^i(Y^i)^*)$, meaning that this object does not obey the symmetry of a system and must vanish.

Additionally, rotational symmetry of a system imposes that matrix elements $|X^i|^2$ and $|Y^i|^2$ are equal. This can be concluded from the action of the rotation C_n for an angle $\beta_n = 2\pi/n$ around the z axis, being the element of the group symmetry. An element C_n leaves the vector $|A_0\rangle$ unchanged and adds a phase $\exp(\mathrm{i}\beta_n)$ to the vector $|E^i_a\rangle$. Also, it transforms x and y to $x\cos\beta_n + y\sin\beta_n$ and $-x\sin\beta_n + y\cos\beta_n$. Thus, X^i and Y^i are transformed into $\exp(-\mathrm{i}\beta_n)(X^i\cos\beta_n + Y^i\sin\beta_n)$ and $\exp(-\mathrm{i}\beta_n)(-X^i\sin\beta_n + Y^i\cos\beta_n)$, respectively. Correspondingly,

$$\begin{aligned}|X^i|^2 &\xrightarrow{C_n} |X^i|^2\cos^2\beta_n + |Y^i|^2\sin^2\beta_n,\\ |Y^i|^2 &\xrightarrow{C_n} |X^i|^2\sin^2\beta_n + |Y^i|^2\cos^2\beta_n,\end{aligned} \tag{25}$$

where we have neglected the $\mathrm{Re}(X^i(Y^i)^*)$ term, which was previously proven to equal to zero. Since $|X^i|^2$ and $|Y^i|^2$ must remain unchanged under the group symmetry operations, we conclude that the relation $|X^i|^2 = |Y^i|^2$ must hold. Thus, we have obtained a general relation for the dipole moment in the case of the potential symmetry $\mathbf{C}_{n\mathrm{v}}$ ($n \geqslant 3$):

$$\mathbf{d}^{\mathbf{C}_{n\mathrm{v}}}_{\uparrow\downarrow} = \frac{2|g|\mu_B B}{l_{\mathrm{so}}}\left(\sum_i \frac{|X^i|^2}{\epsilon^i}\right)(\upsilon_x\mathbf{e}_x + \upsilon_y\mathbf{e}_y). \tag{26}$$

In these situations, the absolute value of the dipole moment $|\mathbf{d}^{\mathbf{C}_{n\mathrm{v}}}_{\uparrow\downarrow}|^2 \sim (1 + \sin 2\alpha \sin 2\nu)$ is independent of the orientation of the potential with respect to the crystallographic frame.

Analogous analysis can be conducted in the $\mathbf{C}_{\infty\mathrm{v}}$ case. Since the matrix form of the IRs A_0 and E_1 (see Table I) for this symmetry group is the same as for $\mathbf{C}_{n\mathrm{v}}$, the procedure is exactly the same if the change $\beta_n \rightarrow \beta$ in the previous discussion is made.

As an example, we implement the derived formula (26) in the case of the isotropic two-dimensional harmonic confinement $V^{\mathrm{iho}}(x, y) = 1/2m^*\omega^2(x^2 + y^2)$ with $\mathbf{C}_{\infty\mathrm{v}}$ symmetry, assuming only one excited level in the perturbative correction of the qubit states. With the help of the states ψ_0 and ψ_1, corresponding to the ground and the first excited states of the one-dimensional harmonic oscillator, we can define the ground state $|A_0\rangle$ and two complex conjugate eigenstates $|E_a\rangle$ and $|E_b\rangle$ of the degenerate level: $|A_0\rangle = \psi_0(x)\psi_0(y)$, $|E_a\rangle = (\psi_0(x)\psi_1(y) + \mathrm{i}\psi_1(x)\psi_0(y))/\sqrt{2}$, and $|E_b\rangle = (\psi_0(x)\psi_1(y) - \mathrm{i}\psi_1(x)\psi_0(y))/\sqrt{2}$. In this case, the squared norm of the transition matrix element is equal to $|X|^2 = \hbar/4m^*\omega$. Using the energy difference of the ground and the first excited energy level $\epsilon = \hbar\omega$ and the confinement length $l = \sqrt{\hbar/m^*\omega}$, an expression for the dipole moment is obtained [22]:

$$\mathbf{d}^{\mathrm{iho}}_{\uparrow\downarrow} = \frac{|g|\mu_B B m^* l^4}{2l_{\mathrm{so}}\hbar^2}(\upsilon_x\mathbf{e}_x + \upsilon_y\mathbf{e}_y). \tag{27}$$

B. Dipole moment for systems with $\mathbf{C}_{2\mathrm{v}}$ or $\mathbf{C}_{1\mathrm{v}}$ symmetry

As the next step, we discuss potentials with $\mathbf{C}_{2\mathrm{v}}$ symmetry. In this case, coordinates x and y transform according to the IRs B_1 and A_1, respectively. Their symmetry behavior imposes the following: x (y) couples the ground state $|A_0\rangle$ with states transforming according to the IR B_1 (A_1) (see the right-hand panel of Fig. 2). Thus, the SOI-induced corrections of the qubit states are

$$\begin{aligned}|\delta A_0\pm\rangle = \frac{|g|\mu_B B}{2l_{\mathrm{so}}}\sum_i \Bigg(&\frac{\langle B^i_1|x\upsilon_x|A_0\rangle}{\epsilon^i_{B_1}}|B^i_1\mp\rangle\\ &+\frac{\langle A^i_1|y\upsilon_y|A_0\rangle}{\epsilon^i_{A_1}}|A^i_1\mp\rangle\Bigg),\end{aligned} \tag{28}$$

where $\epsilon^i_{B_1}$ ($\epsilon^i_{A_1}$) is the energy difference between the energy level transforming according to the IR B_1 (A_1) and the ground-state energy. We define the transition matrix elements as

$$X^i = \langle B^i_1|x|A_0\rangle,\ \ Y^i = \langle A^i_1|y|A_0\rangle, \tag{29}$$

and obtain the formula for the dipole moment,

$$\mathbf{d}^{\mathbf{C}_{2\mathrm{v}}}_{\uparrow\downarrow} = \frac{|g|\mu_B B}{l_{\mathrm{so}}}\sum_i\left(\frac{|X^i|^2}{\epsilon^i_{B_1}}\upsilon_x\mathbf{e}_x + \frac{|Y^i|^2}{\epsilon^i_{A_1}}\upsilon_y\mathbf{e}_y\right). \tag{30}$$

In this case, anisotropy of the dipole moment appears since it is not forbidden that $\sum_i |X^i|^2/\epsilon^i_{B_1}$ differs from $\sum_i |Y^i|^2/\epsilon^i_{A_1}$.

The anisotropy of the dipole moment can be illuminated using the example of the anisotropic two-dimensional harmonic potential $V^{\mathrm{aho}}(x, y) = 1/2m^*(\omega_x^2x^2 + \omega_y^2y^2)$, with different confinement lengths $l_x = \sqrt{\hbar/m^*\omega_x}$ and $l_y = \sqrt{\hbar/m^*\omega_y}$ along the x and y directions. We set $l = l_x$ and $l_y = kl$, where $k < 1$ is the measure of anisotropy. We assume two excited orbital states in the perturbative correction: one of type A_1 and one of type B_1. In this case we define the ground state $|A_0\rangle = \psi_0(x)\psi_0(y)$ and two excited orbital states $|A_1\rangle = \psi_0(x)\psi_1(y)$ and $|B_1\rangle = \psi_1(x)\psi_0(y)$, where $\psi_{0/1}(x/y)$ represents the ground or first excited state (subscript 0 or 1, respectively) of the one-dimensional harmonic oscillator problem in the x or y direction. The obtained result

$$\mathbf{d}^{\mathrm{aho}}_{\uparrow\downarrow} = \frac{|g|\mu_B B m^* l^4}{2l_{\mathrm{so}}\hbar^2}(\upsilon_x\mathbf{e}_x + k^4\upsilon_y\mathbf{e}_y) \tag{31}$$

is again consistent with Ref. [22].

In the case of the $\mathbf{C}_{1\mathrm{v}}$ symmetry, using a similar analysis as in the previous case, we obtain the expression for the dipole moment,

$$\mathbf{d}^{\mathbf{C}_{1\mathrm{v}}}_{\uparrow\downarrow} = \frac{|g|\mu_B B}{l_{\mathrm{so}}}\sum_i\left(\frac{|X^i|^2}{\epsilon^i_{B_0}}\upsilon_x\mathbf{e}_x + \frac{|Y^i|^2}{\epsilon^i_{A_0}}\upsilon_y\mathbf{e}_y\right), \tag{32}$$

where $X^i = \langle B_0^i|x|A_0\rangle$, $Y^i = \langle A_0^i|y|A_0\rangle$, and $\epsilon^i_{A_0/B_0}$ is the energy difference between the nondegenerate energy level transforming according to the IR A_0/B_0 and the ground-state energy.

C. Dipole moment for systems with $\mathbf{C}_n$ or $\mathbf{C}_\infty$ symmetry

In the case of $\mathbf{C}_n$ symmetry, all IRs A_m ($m \in (-n/2, n/2]$) are one dimensional and represent an element of symmetry C_n^s ($s = 0, 1, \ldots, n-1$) as $e^{i2\pi ms/n}$. Besides the geometric symmetry, the time-reversal symmetry Θ should be included also [43]. Time-reversal Θ changes the sign of the quantum number m labeling the IR vector $|A_m\rangle$, since it acts as a complex conjugation in the orbital space:

$$\Theta|A_m\rangle = |A_{-m}\rangle. \tag{33}$$

The eigenproblem of the Hamiltonian $H|A_m\rangle = \epsilon_m|A_m\rangle$, when combined with the commutation relation $[\Theta, H_0] = 0$, gives us

$$H|A_{-m}\rangle = \epsilon_m|A_{-m}\rangle, \tag{34}$$

stating that, for $n \geqslant 3$, vectors $|A_m\rangle$ and $|A_{-m}\rangle$ are eigenstates of the degenerate level ϵ_m. To this degenerate level corresponds the reducible representation $A_m \oplus A_{-m}$ (except for $m = n/2$). The representation $A_m \oplus A_{-m}$ is equivalent to the IR E_m of the $\mathbf{C}_{n\mathrm{v}}$ group (see Table I) if the generator σ_v is neglected. In other words, Eq. (23) for the dipole moment is valid also in this case, since it is obtained without assuming the presence of vertical mirror symmetry. In this case vectors $|E_a^i\rangle$ and $|E_b^i\rangle$ coincide with $|A_1^i\rangle$ and $|A_{-1}^i\rangle$, respectively.

A further simplification of Eq. (23) appears for systems whose symmetry element is $\pi/2$ rotation. This happens if the relation $n = 4r$ ($r \in \mathbb{N}$) is satisfied. Since $\mathrm{Re}(X^i(Y^i)^*) = 0$ and $|X^i|^2 = |Y^i|^2$ in this case, Eq. (26) is relevant. Using the same reasoning it can be concluded that Eq. (26) is valid in the $\mathbf{C}_\infty$ case also.

Finally, the dipole moment components for the $\mathbf{C}_2$ symmetry are equal to

$$\begin{aligned}(\mathbf{d}^{\mathbf{C}_2}_{\uparrow\downarrow})_x &= \frac{|g|\mu_B B}{l_{\mathrm{so}}}\sum_i \frac{|X^i|^2 v_x + \mathrm{Re}(X^i(Y^i)^*)v_y}{\epsilon^i_{A_1}},\\ (\mathbf{d}^{\mathbf{C}_2}_{\uparrow\downarrow})_y &= \frac{|g|\mu_B B}{l_{\mathrm{so}}}\sum_i \frac{|Y^i|^2 v_y + \mathrm{Re}(X^i(Y^i)^*)v_x}{\epsilon^i_{A_1}},\end{aligned} \tag{35}$$

where $X^i = \langle A_1^i|x|A_0\rangle$, $Y^i = \langle A_1^i|y|A_0\rangle$, and $\epsilon^i_{A_1}$ is the energy difference between the level transforming according to the IR A_1 and the ground-state energy.

To conclude, anisotropy of the potential orientation with respect to the crystallographic frame is present in systems without the $\pi/2$ group element ($n \neq 4r$, $r \in \mathbb{N}$); isotropic behavior is present if a rotation for $\pi/2$ is the group element, i.e., if $n = 4r$ ($r \in \mathbb{N}$) or $n = \infty$.

V. APPLICATIONS: INFINITE-WALL EQUILATERAL TRIANGLE, SQUARE, AND RECTANGULAR POTENTIAL

The results presented in the previous section fully explain the dependence of the Rabi frequency and spin relaxation rate on the spin-orbit angle, magnetic field direction, and the relative orientation of the gating potential with respect to the crystallographic frame.

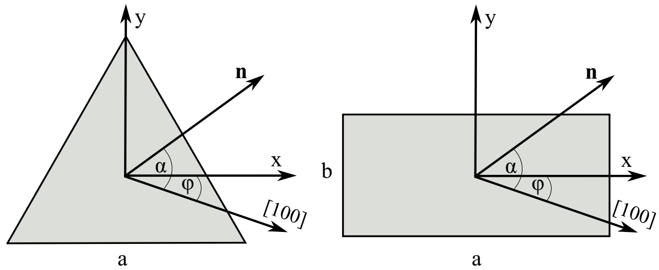


FIG. 3. Infinite-wall equilateral triangle (left) and rectangular (right) gating potential. In both cases, potential is zero inside the area of the polygon; otherwise it is ∞.

However, symmetry arguments alone cannot provide us with a qualitative estimation of the spin relaxation rate, corresponding to the phonon-allowed spin qubit lifetime. Since $\Gamma_{\uparrow\downarrow}$ is known for the harmonic gating [22], we wish to compare the phonon-induced spin relaxation rate of other confinement potentials with the known values. To this end, we analyze the spin qubit confined inside the infinite-wall equilateral triangle, square, and rectangular gating potential (see Fig. 3):

$$V^{\mathrm{tqd}} = \begin{cases} 0 & \text{for } x \in \left[\frac{y\sqrt{3}-a}{3}, \frac{a-y\sqrt{3}}{3}\right],\ y \in \left[\frac{-a\sqrt{3}}{6}, \frac{a\sqrt{3}}{3}\right] \\ \infty & \text{otherwise,}\end{cases} \tag{36}$$

$$V^{\mathrm{rqd}} = \begin{cases} 0 & \text{for } x \in \left[-\frac{a}{2}, \frac{a}{2}\right],\ y \in \left[-\frac{b}{2}, \frac{b}{2}\right] \\ \infty & \text{otherwise.}\end{cases} \tag{37}$$

In the first case, Eq. (36), the potential has $\mathbf{C}_{3\mathrm{v}}$ symmetry and the corresponding eigenvectors of the spin-independent Hamiltonian H_0 transform according to the one-dimensional IRs A_0 and B_0 and two-dimensional E_1 IR of the $\mathbf{C}_{3\mathrm{v}}$ group. The set of eigenenergies $\epsilon^{\mathrm{tqd}}_{p,q}$ and eigenvectors $\psi^{A_0}_{p,q}$, $\psi^{B_0}_{p,q}$, and $\psi^{E_1\pm}_{p,q}$ [44] are dependent on two parameters p and q that have different sets of allowed values for each IR. Their concrete form is given in Appendix A.

In the second case, Eq. (37), the symmetry of the potential is dependent on the ratio $k = b/a \in (0, 1]$: if $k = 1$, the symmetry of the problem is $\mathbf{C}_{4\mathrm{v}}$; otherwise, $\mathbf{C}_{2\mathrm{v}}$ is the symmetry of the spin-independent Hamiltonian H_0. In both situations, eigenenergies and eigenvalues can be found by using the separation of variables. The set of eigenenergies $\epsilon^{\mathrm{rqd}}_{p,q}$ and eigenvectors $\psi^{\mathrm{rqd}}_{p,q}$ in this case is

$$\epsilon^{\mathrm{rqd}}_{p,q} = \frac{\hbar^2\pi^2}{2m^*a^2}\left(p^2 + \frac{q^2}{k^2}\right), \tag{38}$$

$$\psi^{\mathrm{rqd}}_{p,q} = \frac{2}{a\sqrt{k}}\sin\left[\frac{p\pi}{a}\left(x + \frac{a}{2}\right)\right]\sin\left[\frac{q\pi}{ak}\left(y + \frac{ka}{2}\right)\right], \tag{39}$$

defined using the two independent parameters $p \geqslant 1$ and $q \geqslant 1$ that take integer values. However, these solutions do not have any definite symmetry [45]. Therefore, they need to be symmetrized to apply the general results from Sec. IV. Symmetry-adapted eigenfunctions can be found in Appendix B.

After calculating the transition dipole matrix element and the excitation energies for two excited states in the perturbative correction [46], we obtain the desired results

$$\mathbf{d}^{\mathrm{tqd}}_{\uparrow\downarrow} = \frac{3^{24}}{2^{26}35^2\pi^8}\frac{|g|\mu_B B m^* a^4}{l_{\mathrm{so}}\hbar^2}(v_x \boldsymbol{e}_x + v_y \boldsymbol{e}_y), \tag{40}$$

$$\mathbf{d}^{\mathrm{rqd}}_{\uparrow\downarrow} = \frac{2^{9}}{3^5\pi^6}\frac{|g|\mu_B B m^* a^4}{l_{\mathrm{so}}\hbar^2}(v_x \boldsymbol{e}_x + k^4 v_y \boldsymbol{e}_y), \tag{41}$$

where the first result corresponds to the infinite-wall equilateral triangle potential, while the second one is valid for both the infinite-wall square, $k = 1$, and rectangular, $k \neq 1$, potentials. Dipole moment constants $3^{24}/2^{26}35^2\pi^8 \approx 3.6 \times 10^{-4}$ and $2^9/3^5\pi^6 \approx 2.2 \times 10^{-3}$ from Eqs. (40) and (41) suggest a much weaker dipole moment when compared to the harmonic gating of the same confinement length [see Eqs. (27) and (31)].

Using the relation $\Gamma_{\uparrow\downarrow} \approx |\mathbf{d}_{\uparrow\downarrow}|^2$, we conclude that square and rectangular confined QDs have a relaxation rate that is four orders of magnitude weaker than the harmonic potential; in the equilateral triangle case, a decrease of almost six orders of magnitude is observed. Thus, our result indicates a significant influence of the gating potential on the spin qubit lifetime and a beneficial role of the equilateral triangle confinement.

VI. CONCLUSIONS

We have investigated the influence of the gating potential symmetry on the Rabi frequency and phonon-induced spin relaxation rate in a single-electron GaAs quantum dot. Our results suggest that, independently of the symmetry of the gating potential, both the Rabi frequency and spin relaxation rate are dependent on the orientation of the magnetic field and the spin-orbit angle. Additionally, in systems with $\mathbf{C}_{1v}$, $\mathbf{C}_{2v}$, and $\mathbf{C}_n$ $(n \neq 4r)$ symmetry, orientation of the quantum dot potential with respect to the crystallographic reference frame is another degree of freedom that can be used to tune the desired properties of the system. The validity of the approach is confirmed on the known results for the isotropic and anisotropic harmonic potential. Additionally, we have compared the spin qubit lifetime in the case of an infinite-wall rectangular, square, and equilateral triangle gating with the harmonic confinement. Our results indicate the enhanced lifetime of the spin qubit, reaching an almost six-order-of-magnitude increase in the case of the equilateral triangle gating. In the end, we emphasize that in the regime of strong electric field, nonlinear effects [47–49] cannot be fully explained by the symmetry of the gating potential, thus placing the conclusions of our work in the weak driving regime solely.

ACKNOWLEDGMENTS

We thank Nenad Vukmirović for fruitful discussion. This research is funded by the Serbian Ministry of Science (Project No. ON171035) and the National Scholarship Programme of the Slovak Republic (ID 28226).

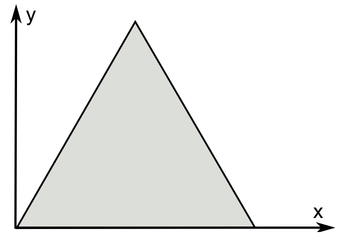


FIG. 4. Infinite-wall equilateral triangle potential with the point group symmetry $\mathbf{C}_{3v}$. Inside the equilateral triangle potential is 0, otherwise it is ∞.

APPENDIX A: PARTICLE IN THE INFINITE-WALL EQUILATERAL TRIANGLE POTENTIAL: EIGENENERGIES AND EIGENVECTORS

Here we summarize the results from Ref. [44] regarding the Schrödinger equation solution of the particle in the infinite-wall equilateral triangle potential, having $\mathbf{C}_{3v}$ symmetry. Due to the symmetry, eigenvectors transform according to the one-dimensional IRs A_0 and B_0 and the two-dimensional IR E_1. The concrete forms of eigenenergies and eigenstates,

$$\epsilon^{\mathrm{tqd}}_{p,q} = \frac{8\hbar^2\pi^2}{3m^*a^2}(p^2 + pq + q^2), \tag{A1}$$

$$\begin{aligned}\psi^{A_0}_{p,q}(x,y) = {} & \cos\left[\frac{2\pi q}{a}x\right]\sin\left[\frac{2\pi(2p+q)}{a\sqrt{3}}y\right] \\ & - \cos\left[\frac{2\pi p}{a}x\right]\sin\left[\frac{2\pi(p+2q)}{a\sqrt{3}}y\right] \\ & - \cos\left[\frac{2\pi(p+q)}{a}x\right]\sin\left[\frac{2\pi(p-q)}{a\sqrt{3}}y\right], \\ & q = 0,1,2,\ldots, \quad p = q+1, q+2, \ldots,\end{aligned} \tag{A2}$$

$$\begin{aligned}\psi^{B_0}_{p,q}(x,y) = {} & \sin\left[\frac{2\pi q}{a}x\right]\sin\left[\frac{2\pi(2p+q)}{a\sqrt{3}}y\right] \\ & - \sin\left[\frac{2\pi p}{a}x\right]\sin\left[\frac{2\pi(p+2q)}{a\sqrt{3}}y\right] \\ & + \sin\left[\frac{2\pi(p+q)}{a}x\right]\sin\left[\frac{2\pi(p-q)}{a\sqrt{3}}y\right], \\ & q = 1,2,3,\ldots, \quad p = q+1, q+2, \ldots,\end{aligned} \tag{A3}$$

$$\begin{aligned}\psi^{E_1\pm}_{p,q}(x,y) = {} & \psi^{B_0}_{p,q}(x,y) \pm \mathrm{i}\psi^{A_0}_{p,q}(x,y), \\ & q = \frac{1}{3}, \frac{2}{3}, \frac{4}{3}, \frac{5}{3}, \ldots, \quad p = q+1, q+2, \ldots,\end{aligned} \tag{A4}$$

are dependent on two parameters p and q that have different allowed values for each IR. Note that the coordinate frame used to derive the previous equations (see Fig. 4) differs from the frame used in our work (see the left-hand panel of Fig. 3). To adapt the eigenfunction from Eqs. (A2)–(A4) to our case, a suitable change of coordinates $x \rightarrow x + a/2$ and $y \rightarrow y + a\sqrt{3}/6$ should be made.

APPENDIX B: PARTICLE IN THE INFINITE-WALL SQUARE AND RECTANGULAR POTENTIAL: EIGENVECTORS

The infinite-wall square potential has $\mathbf{C}_{4\mathrm{v}}$ symmetry with the corresponding IRs A_0/B_0, A_2/B_2, and E_1. Eigenvectors that transform according to the given IRs and the set of allowed quantum numbers are

$$\psi^{A_0}_{p,q}(x,y) = \cos\left[\frac{p\pi}{a}x\right]\cos\left[\frac{q\pi}{a}y\right] + \cos\left[\frac{q\pi}{a}x\right]\cos\left[\frac{p\pi}{a}y\right],$$
$$q = 1,3,5,\ldots,\quad p = q, q+2, q+4,\ldots, \tag{B1}$$

$$\psi^{B_0}_{p,q}(x,y) = \sin\left[\frac{p\pi}{a}x\right]\sin\left[\frac{q\pi}{a}y\right] - \sin\left[\frac{q\pi}{a}x\right]\sin\left[\frac{p\pi}{a}y\right],$$
$$q = 2,4,6,\ldots,\quad p = q+2, q+4,\ldots, \tag{B2}$$

$$\psi^{A_2}_{p,q}(x,y) = \cos\left[\frac{p\pi}{a}x\right]\cos\left[\frac{q\pi}{a}y\right] - \cos\left[\frac{q\pi}{a}x\right]\cos\left[\frac{p\pi}{a}y\right],$$
$$q = 1,3,5,\ldots,\quad p = q+2, q+4,\ldots, \tag{B3}$$

$$\psi^{B_2}_{p,q}(x,y) = \sin\left[\frac{p\pi}{a}x\right]\sin\left[\frac{q\pi}{a}y\right] + \sin\left[\frac{q\pi}{a}x\right]\sin\left[\frac{p\pi}{a}y\right],$$
$$q = 2,4,6,\ldots,\quad p = q, q+2, q+4,\ldots, \tag{B4}$$

$$\psi^{E_1\pm}_{p,q}(x,y) = \cos\left[\frac{p\pi}{a}x\right]\sin\left[\frac{q\pi}{a}y\right] \pm i\sin\left[\frac{q\pi}{a}x\right]\cos\left[\frac{p\pi}{a}y\right],$$
$$p = 1,3,5,\ldots,\quad q = p+1, p+3,\ldots. \tag{B5}$$

In the case of the infinite-wall rectangular potential $\mathbf{C}_{2\mathrm{v}}$ symmetry is relevant. Eigenfunctions transforming according to the IRs A_0/B_0 and A_1/B_1 and the corresponding set of quantum numbers are

$$\psi^{A_0}_{p,q}(x,y) = \cos\left[\frac{p\pi}{a}x\right]\cos\left[\frac{q\pi}{ka}y\right],$$
$$q = 1,3,5,\ldots,\quad p = q, q+2, q+4,\ldots, \tag{B6}$$

$$\psi^{B_0}_{p,q}(x,y) = \sin\left[\frac{p\pi}{a}x\right]\sin\left[\frac{q\pi}{ka}y\right],$$
$$q = 2,4,6,\ldots,\quad p = q, q+2, q+4,\ldots, \tag{B7}$$

$$\psi^{A_1}_{p,q}(x,y) = \cos\left[\frac{p\pi}{a}x\right]\sin\left[\frac{q\pi}{ka}y\right],$$
$$p = 1,3,5,\ldots,\quad q = p+1, p+3, p+5,\ldots, \tag{B8}$$

$$\psi^{B_1}_{p,q}(x,y) = \sin\left[\frac{p\pi}{a}x\right]\cos\left[\frac{q\pi}{ka}y\right],$$
$$q = 1,3,5,\ldots,\quad p = q+1, q+3, q+5,\ldots. \tag{B9}$$

In both cases, eigenenergies are given in Eq. (38) ($k = 1$ in the $\mathbf{C}_{4\mathrm{v}}$ case and $k \neq 1$ for the $\mathbf{C}_{2\mathrm{v}}$ symmetry).

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2010).

[2] C. H. Bennett and D. P. DiVincenzo, Nature (London) **404**, 247 (2000).

[3] D. Loss and D. P. DiVincenzo, Phys. Rev. A **57**, 120 (1998).

[4] T. P. Orlando, J. E. Mooij, L. Tian, C. H. van der Wal, L. S. Levitov, S. Lloyd, and J. J. Mazo, Phys. Rev. B **60**, 15398 (1999).

[5] Y. Nakamura, Yu. A. Pashkin, and J. S. Tsai, Nature (London) **398**, 786 (1999).

[6] M. Yamamoto, S. Takada, C. Bäuerle, K. Watanabe, A. D. Wieck, and S. Tarucha, Nat. Nanotechnol. **7**, 247 (2012).

[7] E. I. Rashba, Sov. Phys. Solid State **2**, 1109 (1960).

[8] V. N. Golovach, M. Borhani, and D. Loss, Phys. Rev. B **74**, 165319 (2006).

[9] D. V. Bulaev and D. Loss, Phys. Rev. Lett. **98**, 097202 (2007).

[10] K. C. Nowack, F. H. L. Koppens, Yu. V. Nazarov, and L. M. K. Vandersypen, Science **318**, 1430 (2007).

[11] E. I. Rashba, Phys. Rev. B **78**, 195302 (2008).

[12] R. Brunner, Y.-S. Shin, T. Obata, M. Pioro-Ladrière, T. Kubo, K. Yoshida, T. Taniyama, Y. Tokura, and S. Tarucha, Phys. Rev. Lett. **107**, 146801 (2011).

[13] E. Kawakami, P. Scarlino, D. R. Ward, F. R. Braakman, D. E. Savage, M. G. Lagally, Mark Friesen, S. N. Coppersmith, M. A. Eriksson, and L. M. K. Vandersypen, Nat. Nanotechnol. **9**, 666 (2014).

[14] K. Takeda, J. Yoneda, T. Otsuka, T. Nakajima, M. R. Delbecq, G. Allison, Y. Hoshi, N. Usami, K. M. Itoh, S. Oda, T. Kodera, and S. Tarucha, npj Quantum Inf. **4**, 54 (2018).

[15] D. V. Khomitsky, E. A. Lavrukhina, and E. Y. Sherman, Phys. Rev. B **99**, 014308 (2019).

[16] S. Studenikin, M. Korkusinski, M. Takahashi, J. Ducatel, A. Padawer-Blatt, A. Bogan, D. Guy Austing, L. Gaudreau, P. Zawadzki, A. Sachrajda, Y. Hirayama, L. Tracy, J. Reno, and T. Hargett, Commun. Phys. **2**, 159 (2019).

[17] A. V. Khaetskii and Y. V. Nazarov, Phys. Rev. B **61**, 12639 (2000); **64**, 125316 (2001).

[18] P. Stano and J. Fabian, Phys. Rev. B **72**, 155410 (2005).

[19] V. N. Stavrou, J. Phys.: Condens. Matter **29**, 485301 (2017); **30**, 455301 (2018).

[20] Z.-H. Liu, R. Li, X. Hu, and J. Q. You, Sci. Rep. **8**, 2302 (2018).

[21] J. I. Climente, A. Bertoni, G. Goldoni, M. Rontani, and E. Molinari, Phys. Rev. B **75**, 081303(R) (2007).

[22] O. Malkoc, P. Stano, and D. Loss, Phys. Rev. B **93**, 235413 (2016).

[23] D. Chaney and P. A. Maksym, Phys. Rev. B **75**, 035323 (2007).

[24] V. I. Fal'ko, B. L. Altshuler, and O. Tsyplyatyev, Phys. Rev. Lett. **95**, 076603 (2005).

[25] P. Scarlino, E. Kawakami, P. Stano, M. Shafiei, C. Reichl, W. Wegscheider, and L. M. K. Vandersypen, Phys. Rev. Lett. **113**, 256802 (2014).

[26] O. Olendski and T. V. Shahbazyan, Phys. Rev. B **75**, 041306(R) (2007).

[27] S. Amasha, K. MacLean, I. P. Radu, D. M. Zumbühl, M. A. Kastner, M. P. Hanson, and A. C. Gossard, Phys. Rev. Lett. **100**, 046803 (2008).

[28] J. Planelles, F. Rajadell, and J. I. Climente, Phys. Rev. B **92**, 041302(R) (2015).

[29] M. Raith, P. Stano, and J. Fabian, Phys. Rev. B **83**, 195318 (2011).
[30] G. Dresselhaus, Phys. Rev. **100**, 580 (1955).
[31] E. I. Rashba, Fiz. Tverd. Tela (Leningrad) **2**, 1224 (1960) [Sov. Phys. Solid State **2**, 1109 (1960)].
[32] E. Ya. Sherman and D. J. Lockwood, Phys. Rev. B **72**, 125340 (2005).
[33] M. Raith, T. Pangerl, P. Stano, and J. Fabian, Phys. Status Solidi B **251**, 1924 (2014).
[34] J. L. Cheng, M. W. Wu, and C. Lü, Phys. Rev. B **69**, 115318 (2004).
[35] F. Baruffa, P. Stano, and J. Fabian, Phys. Rev. Lett. **104**, 126401 (2010).
[36] R. de Sousa and S. Das Sarma, Phys. Rev. B **68**, 155330 (2003).
[37] X. Linpeng, T. Karin, M. V. Durnev, R. Barbour, M. M. Glazov, E. Ya. Sherman, S. P. Watkins, S. Seto, and K.-M. C. Fu, Phys. Rev. B **94**, 125401 (2016).
[38] Note that both the wave vector $\mathbf{q}$ and the electron coordinate $\mathbf{r}_c$ are written in the crystallographic reference frame.
[39] J. I. Climente, A. Bertoni, G. Goldoni, and E. Molinari, Phys. Rev. B **74**, 035313 (2006).
[40] L. C. Camenzind, L. Yu, P. Stano, J. D. Zimmerman, A. C. Gossard, D. Loss, and D. M. Zumbühl, Nat. Commun. **9**, 3454 (2018).
[41] E. Lijnen, L. F. Chibotaru, and A. Ceulemans, Phys. Rev. E **77**, 016702 (2008).
[42] L. Jansen and M. Boon, *Theory of Finite Groups: Applications in Physics* (North Holland, Amsterdam, 1967); M. Damnjanović, O simetriji u kvantnoj nerelativističkoj fizici, Fizički fakultet, Beograd (2000).
[43] In the case of $\mathbf{C}_{n\text{v}}/\mathbf{C}_{\infty\text{v}}$ groups, time-reversal Θ is the symmetry of the system. However, it can be checked that Θ can be safely neglected. Time reversal in the case of vectors from one-dimensional IRs has a trivial action. In the case of the two-dimensional IRs E_m time reversal transforms one vector of the IR into the other; in other words, it has the same behavior as the vertical mirror plane symmetry σ_v and can be ignored.
[44] W.-K. Li and S. M. Blinder, J. Math. Phys. **26**, 2784 (1985).
[45] L. F. Chibotaru, A. Ceulemans, M. Morelle, G. Teniers, C. Carballeira, and V. V. Moshchalkov, J. Math. Phys. **46**, 095108 (2005).
[46] We have explicitly checked that other states appearing in the perturbative expansion can be safely ignored.
[47] D. V. Khomitsky, L. V. Gulyaev, and E. Ya. Sherman, Phys. Rev. B **85**, 125312 (2012).
[48] J. Romhányi, G. Burkard, and A. Pályi, Phys. Rev. B **92**, 054422 (2015).
[49] M. T. Veszeli and A. Pályi, Phys. Rev. B **97**, 235433 (2018).

Analytical solution for time integrals in diagrammatic expansions: Application to real-frequency diagrammatic Monte Carlo

J. Vučičević,[1] P. Stipsić,[1,2] and M. Ferrero[3,4]

[1]Scientific Computing Laboratory, Center for the Study of Complex Systems, Institute of Physics Belgrade, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia

[2]Faculty of Physics, University of Belgrade, Studentski trg 12, 11001 Belgrade, Serbia

[3]CPHT, CNRS, Ecole Polytechnique, Institut Polytechnique de Paris, Route de Saclay, 91128 Palaiseau, France

[4]Collège de France, 11 place Marcelin Berthelot, 75005 Paris, France

(Received 19 November 2020; accepted 25 March 2021; published 29 April 2021)

Recent years have seen a revived interest in the diagrammatic Monte Carlo (DiagMC) methods for interacting fermions on a lattice. A promising recent development allows one to now circumvent the analytical continuation of dynamic observables in DiagMC calculations within the Matsubara formalism. This is made possible by symbolic algebra algorithms, which can be used to analytically solve the internal Matsubara frequency summations of Feynman diagrams. In this paper, we take a different approach and show that it yields improved results. We present a closed-form analytical solution of imaginary-time integrals that appear in the time-domain formulation of Feynman diagrams. We implement and test a DiagMC algorithm based on this analytical solution and show that it has numerous significant advantages. Most importantly, the algorithm is general enough for any kind of single-time correlation function series, involving any single-particle vertex insertions. Therefore, it readily allows for the use of action-shifted schemes, aimed at improving the convergence properties of the series. By performing a frequency-resolved action-shift tuning, we are able to further improve the method and converge the self-energy in a nontrivial regime, with only 3–4 perturbation orders. Finally, we identify time integrals of the same general form in many commonly used Monte Carlo algorithms and therefore expect a broader usage of our analytical solution.

DOI: 10.1103/PhysRevResearch.3.023082

I. INTRODUCTION

Finding controlled solutions of the Hubbard model is one of the central challenges in condensed matter physics [1–4]. Many common approaches to this problem rely on the stochastic (Monte Carlo) summation of various expansions and decompositions of relevant physical quantities. However, Monte Carlo (MC) algorithms are often plagued by two notorious problems: the fermionic sign problem and the analytical continuation of frequency-dependent quantities in calculations based on the Matsubara formalism [5–8] (alternatively, the dynamical sign problem in the Kadanoff-Baym and Keldysh formalism calculations [9–23]). In diagrammatic Monte Carlo (DiagMC) methods [24–38] (as opposed to determinantal methods such as continuous-time interaction-expansion quantum Monte Carlo (CTINT) or, auxiliary-field quantum Monte Carlo (CTAUX) [39–42]), an additional problem is often the slow (or absence of) convergence of the series with respect to the perturbation order. In recent years, several works have started to address the problems of obtaining real-frequency quantities [43–51] and series convergence in DiagMC [52–57].

In Refs. [43,52], it has been shown that a convenient transformation of the interaction-expansion series can be used to significantly improve its convergence and sometimes allows one to converge the electronic self-energy with only a few perturbation orders where it would have otherwise been impossible. The method relies on a transformation of the action which affects the bare propagator at the cost of an additional expansion, i.e., more diagram topologies need to be taken into account. Alternatively, this transformation can be viewed as a Maclaurin expansion of the bare propagator with respect to a small chemical potential shift. The resulting convergence speedup comes from an increased convergence radius of the transformed series.

In a separate line of work, DiagMC methods have been proposed that are based on the Matsubara formalism that do not require an ill-defined analytical continuation [47]. Such methods have so far been implemented for the calculation of the self-energy [48,49] and the dynamical spin susceptibility [50]. The algorithms differ in some aspects, but all rely on the symbolic algebra solution of the internal Matsubara frequency summations appearing in Feynman diagrams. However, this approach has some downsides. First, numerical regulators are needed to properly evaluate Bose-Einstein distribution functions and diverging ratios that appear in the analytical expressions, and also poles on the real axis (effective broadening of the real-frequency results). In the case of finite cyclic lattice calculations, multiple precision algebra is needed in order to cancel divergences even with relatively large regulators [48]. Most importantly, in the Matsubara summation algorithm, applying the series transformation from Refs. [43,52] would require a separate analytical solution for each of the additional diagram topologies, which are very numerous, and the calculation would become rather impractical. More generally, treating any distinct diagram requires that the Matsubara frequency summations be performed algorithmically beforehand. This makes it difficult to devise MC sampling algorithms that go to indefinite perturbation orders, unless the Matsubara summation part is sufficiently optimized so that it no longer presents a prohibitive performance penalty if performed at the time of the Monte Carlo sampling.

In this paper, we show that it can be advantageous to start from the imaginary-time domain formulation of Feynman diagrams. A diagram contribution then features a multiple imaginary-time integral, rather than sums over Matsubara frequencies. The multiple integral can be solved analytically and we present a general solution. This analytical solution, although equivalent to the analytical Matsubara summation, has a simpler and more convenient form that does not feature Bose-Einstein distribution functions or diverging ratios. As a result, numerical regulators are not needed and the need for multiple precision arithmetic may arise only at very high perturbation orders. The numerical evaluation yields a sum of poles of various orders on a uniform grid on the real axis. The ability to separate contributions of poles of different orders allows one to formally extract the real-frequency result without any numerical broadening. Finally, the analytical solution is general and applies to all diagram topologies that would appear in the transformed series proposed in Refs. [43,52] or any other diagrammatic series for single-time correlation functions. This paves the way for real-frequency diagrammatic algorithms formulated in real space that are not *a priori* limited to small perturbation orders (similarly to CTINT or CTAUX [42]).

In this work, we apply the analytical time integral to the momentum-space DiagMC for the calculation of the self-energy, and implement and thoroughly test the method. We reproduce the self-energy results from Ref. [52] and supplement them with real-axis results, free of the uncontrolled systematic error that would otherwise come from the analytical continuation. Furthermore, we show that even if a full convergence is not possible with a single choice of the action-tuning parameter, one can choose the optimal tuning parameter for each frequency independently [46]. Such a frequency-resolved resummation can be used to improve the solution and in some cases systematically eliminate the nonphysical features that appear in the result due to the truncation of the series at a finite order.

The paper is organized as follows. In Sec. II, we define the model and the basic assumptions of our calculations. In Sec. III, we introduce our method in detail. First, in Sec. III A, we present the analytical solution of the general multiple-time integral that appears in the time-domain formulation of Feynman diagrams and discuss the numerical evaluation of the final expression. Then, in Sec. III B, we show the analytical solution for the Fourier transform of the Maclaurin expansion of the bare propagator, which is essential for our DiagMC algorithm. In Sec. III C, we discuss in detail how our analytical solutions can be applied in the context of DiagMC for the self-energy. In Sec. IV, we discuss our results and benchmarks and then give closing remarks in Sec. V. Additional details of the analytical derivations and further benchmarks and examples of the calculations can be found in the appendices.

II. MODEL

We solve the Hubbard model given by the Hamiltonian

$$H = -\sum_{\sigma,ij} t_{ij} c^\dagger_{\sigma,i} c_{\sigma,j} + U\sum_i n_{\uparrow,i} n_{\downarrow,i} - \mu \sum_{\sigma,i} n_{\sigma,i}, \tag{1}$$

where $\sigma \in \{\uparrow,\downarrow\}$, i, j enumerate lattice sites, t_{ij} is the hopping amplitude between the sites i and j, U is the on-site coupling constant, and μ is the chemical potential. We only consider the Hubbard model on the square lattice with the nearest-neighbor hopping t and next-nearest-neighbor hopping t'. The bare dispersion is given by

$$\varepsilon_{\mathbf{k}} = -2t(\cos k_x + \cos k_y) - 4t' \cos k_x \cos k_y. \tag{2}$$

We define $D = 4t$, which will be used as the unit of energy unless stated otherwise. We restrict to thermal equilibrium and paramagnetic phases with full lattice symmetry.

III. METHODS

The idea of DiagMC algorithms is to stochastically compute the coefficients of a perturbation series describing some physical quantity. We will focus on expansions in the coupling constant U and a shift in the chemical potential $\delta\mu$. The calculation of each coefficient involves the evaluation of many Feynman diagrams expressed in terms of the bare propagator, in our case taken as a function of momentum and two imaginary times. The evaluation of a diagram then boils down to a sum over multiple momentum variables and a multiple imaginary-time integral that is always of the same generic form. The goal of this section is to find a general analytical solution for these time integrals and reformulate the perturbation series as a function of a complex frequency z.

A. Analytical solution of time integrals

We are interested in analytically solving $(N-1)$-fold integrals over $\{\tau_{i=2...N}\}$ of the form

$$\mathcal{I}_{\mathbf{X}}(i\Omega_\eta) = \prod_{i=2}^{N} \int_0^{\tau_{i+1}} d\tau_i\ \tau_i^{l_i}\ e^{\tau_i(i\Omega_\eta \delta_{r,i} + \omega_i)}, \tag{3}$$

where the parameters of the integrand are given by

$$\mathbf{X} = (r, \{l_2...l_N\}, \{\omega_2...\omega_N\}). \tag{4}$$

The argument r is an integer and determines which of the times τ_i is multiplied by the external Matsubara frequency $i\Omega_\eta$ in the exponential. The frequency $i\Omega_\eta$ can be any Matsubara frequency, either fermionic or bosonic, depending on η; $i\Omega_{\eta=-1} \equiv i\omega \equiv i(2m+1)\pi T$ and $i\Omega_{\eta=1} \equiv i\nu \equiv 2im\pi T$, with $m \in \mathbb{Z}$. The integer powers of τ_i outside of the exponent are given by $l_i \geqslant 0$, and the parameters ω_i may be complex.

The limit of the outermost integration is the inverse temperature $\tau_{N+1} \equiv \beta$. We denote by $\delta_{x,y}$ the Kronecker delta (it will be used throughout this paper, also in the shortened version $\delta_x \equiv \delta_{x,0}$). The reason for our choice to label times starting from 2 will become clear later.

The main insight is that upon applying the innermost integral, one gets a number of terms, but each new integrand has the same general form $\sim\tau^n e^{\tau z}$. The solution therefore boils down to a recursive application of

$$\int_0^{\tau_f} \tau^n e^{\tau z} d\tau = \sum_{k=0}^{n+1} (-)^k C_{nk} \frac{\tau_f^{n+1-k-B_{nk}} e^{B_{nk} z \tau_f}}{z^{k+B_{nk}}}, \quad (5)$$

with $B_{nk} = 1 - \delta_{k,n+1}$ and $C_{nk} = \frac{n!}{(n-k+\delta_{k,n+1})!}$ (for the proof, see Appendix D), and

$$\lim_{z\to 0} \int_0^{\tau_f} \tau^n e^{\tau z} d\tau = \frac{\tau_f^{n+1}}{n+1}. \quad (6)$$

The number of terms obtained after each integration is apparently $1 + (1 - \delta_z)(n + 1)$, and we can enumerate all terms obtained after the full integration by a set of integers, $\{k_{i=2\ldots N}\}$, where $k_i \geqslant 0$ denotes the choice of the term of the integral i (over $d\tau_i$).

For a given choice of $\{k_i\}$, the propagation of exponents [n and z in Eqs. (5) and (6)] across successive integrals can be fully described by a simple set of auxiliary quantities. We denote the exponent of e in the integration i as $\tilde{z}_i$, and it is given by

$$\tilde{z}_i \equiv z_i + b_{i-1}\tilde{z}_{i-1}, \quad \tilde{z}_2 \equiv z_2, \quad (7)$$

$$z_i \equiv \delta_{i,r} i\Omega_\eta + \omega_i, \quad (8)$$

where we introduced $b_i \equiv B_{n_i,k_i}$. The meaning of b_i can be understood by looking at Eq. (5): The exponent of e that enters the integral on the left-hand side survives in all but the last term ($k = n + 1$) on the right-hand side. Therefore, $b_i = 1$ means that the exponent propagates from integration i to integration $i + 1$, while $b_i = 0$ means it does not, and the calculation of the recursive $\tilde{z}_i$ is reset with each $b_i = 0$. The auxiliary quantity n_i are the exponents of τ_i and is specified below.

We will need to obtain a more convenient expression for the exponent $\tilde{z}_i$, where $i\Omega_\eta$ appears explicitly. Straightforwardly, we can write

$$\tilde{z}_i = i\Omega_\eta h_i + \tilde{\omega}_i, \quad (9)$$

with auxiliary quantities

$$\tilde{\omega}_i \equiv \omega_i + b_{i-1}\tilde{\omega}_{i-1}, \quad \tilde{\omega}_2 \equiv \omega_2, \quad (10)$$

and

$$h_i \equiv \begin{cases} 0, & i < r \\ 1, & i = r \\ b_{i-1}h_{i-1}, & i > r. \end{cases} \quad (11)$$

To be able to determine whether the exponent in the integrand, $\tilde{z}_i$, is zero and then employ Eq. (6) if needed, we can now use

$$\delta_{\tilde{z}_i} = \begin{cases} 1, & h_i = 0 \wedge \tilde{\omega}_i = 0 \\ 0 & \text{otherwise.} \end{cases} \quad (12)$$

It is important to note that at the time of integration, $i\Omega_\eta$ is unspecified and whether $\tilde{z}_i$ is zero cannot be tested by numerical means, unless $i\Omega_\eta$ does not appear in $\tilde{z}_i$. With the convenient rewriting of Eq. (7) as Eq. (9), one can tell whether $i\Omega_\eta$ appears in $\tilde{z}_i$ by looking at h_i. If $i\Omega_\eta$ does appear in $\tilde{z}_i$ (i.e., $h_i = 1$), we cannot use Eq. (6) even if one can find such $i\Omega_\eta$ that cancels $\tilde{\omega}_i$. This is because we are working towards an analytical expression which ought to be general for all possible $i\Omega_\eta$.

The exponent of τ that will be carried over from integration i to integration $i + 1$ depends on the choice of the term from the integral i, and is given by $\mathrm{Pos}(n_i - k_i)$, where Pos denotes the positive part of the number [$\mathrm{Pos}(x) = (x + |x|)/2$]. n_i denotes the maximum exponent that can be carried over from integration i, and is obtained as

$$n_i = \begin{cases} \delta_{\tilde{z}_i} + l_i + \mathrm{Pos}(n_{i-1} - k_{i-1}), & i > 2 \\ \delta_{\tilde{z}_i} + l_i, & i = 2. \end{cases} \quad (13)$$

In the case of Eq. (5), the maximal exponent that can be carried over to the next integration coincides with the exponent that entered the integral [the integral given by Eq. (5) does not raise the power of τ], so the definition of n_i coincides with the meaning of n in Eq. (5). In the case of the integral given by Eq. (6), n_i rather denotes the exponent after the integration, i.e., $n + 1$.

After the last integration, it can happen that $i\Omega_\eta$ appears in the exponent of e (this is signaled by $h_N b_N = 1$). We can then use the property $e^{i\Omega_\eta\beta} = (-1)^{\delta_{\eta,-1}}$ to eliminate it from this exponent. Then, the solution for the integral can be continued to the whole of the complex plane $i\Omega_\eta \to z$, and can be written as (introducing the additional superscript η because the fermionic/bosonic nature of the expression can no longer be inferred from the external Matsubara frequency)

$$\begin{aligned}\mathcal{I}^\eta_{\mathbf{X}}(z) = &\sum_{\{b_i\in[\delta_{\tilde{z}_i},1]\}_{i=2\ldots N}} e^{b_N\beta\tilde{\omega}_N} \sum_{\{k_i\in[0,(1-\delta_{\tilde{z}_i})n_i]\}_{i:b_i=1}} \\ &\times \prod_{i:\delta_{\tilde{z}_i}=1} \frac{1}{n_i} \\ &\times (-1)^{b_N h_N \delta_{\eta,-1} + \sum_{i=2}^N k_i} \times \beta^{n_N+1-b_N-k_N} \\ &\times \prod_{i:h_i=0\wedge\tilde{\omega}_i\neq 0} \frac{C_{n_i,k_i}}{\tilde{\omega}_i^{k_i+b_i}} \prod_{i:h_i=1} \frac{C_{n_i,k_i}}{(z+\tilde{\omega}_i)^{k_i+b_i}}. \end{aligned} \quad (14)$$

Note that we have expressed the sum over $\{k_i\}$ as a sum over $\{b_i\}$ and a partial (inner) sum over $\{k_i\}$. This is not necessary, being that b_i is a function of k_i. Each b_i is fully determined by k_i, but not the other way around, so the inner sum over k_i in Eq. (14) goes over values that are allowed by the corresponding b_i. We present this form of Eq. (14) to emphasize that the factor $e^{b_N\beta\tilde{\omega}_N}$ depends only on $\{b_i\}$, and can thus be pulled out of the inner $\{k_i\}$ sum. The notation "$i : b_i = 1$" means that we only consider indices i such that $b_i = 1$. We therefore only sum over those k_i for which the corresponding $b_i = 1$. The remaining k_i are fixed to $n_i + 1$, which is the only possibility if $b_i = 0$. The notation is applied analogously in other products over i.

TABLE I. Illustration of the calculation of a single term in Eq. (14). Rows correspond to successive integrations over $d\tau_i$. The second to fourth columns are parameters of the integrand. The choice of the term is colored red. The remaining columns are auxiliary quantities, the integrand before and after each integration. The prefactors that are "collected" after each integration are written in blue. The full contribution is written in the last column and then simplified to the form of a term in Eq. (16).

i	$\delta_{r,i}$	l_i	ω_i	k_i	b_i	n_i	$\tilde{\omega}_i$	h_i	$\delta_{\tilde{z}_i}$	Integrand	Integral	Total
2	0	0	1	0	1	0	1	0	0	$e^{\tau_2 1}$	$\frac{1}{1}e^{\tau_3 1} - \frac{1}{1}1$	
3	0	1	2	1	1	1	3	0	0	$\tau_3 e^{\tau_3(2+1)}$	$\frac{1}{3}\tau_4 e^{\tau_4 3} - \frac{1}{3^2}e^{\tau_4 3} + \frac{1}{3^2}1$	
4	1	0	1	1	0	0	4	1	0	$e^{\tau_4(i\Omega_\eta+1+3)}$	$\frac{1}{i\Omega_\eta+4}e^{\tau_5(i\Omega_\eta+4)} - \frac{1}{i\Omega_\eta+4}1$	$\frac{1}{1}(-\frac{1}{3^2})(-\frac{1}{i\Omega_\eta+4})\frac{1}{1}\frac{1}{4}\beta e^{\beta 4}$
5	0	0	0	0	1	1	0	0	1	$e^{\tau_5 0}$	$\frac{1}{1}\tau_6^1$	$\to \frac{\beta e^{4\beta}/36}{[z-(-4)]^1}$
6	0	0	4	0	1	1	4	0	0	$\tau_6 e^{\tau_6 4}$	$\frac{1}{4}\beta e^{\beta 4} - \frac{1}{4^2}e^{\beta 4} + \frac{1}{4^2}1$	

The only remaining step is to expand the product of poles in Eq. (14) into a sum of poles (see Ref. [48] for more details),

$$\prod_\gamma \frac{1}{(z-z_\gamma)^{m_\gamma}} = \sum_\gamma \sum_{r=1}^{m_\gamma} \frac{1}{(z-z_\gamma)^r} \times (-1)^{m_\gamma - r} \sum_{\mathcal{C}\{p_{\gamma'\neq\gamma}\in\mathbb{N}_0\}:\sum_{\gamma'\neq\gamma} p_{\gamma'} = m_\gamma - r} \times \prod_{\gamma'\neq\gamma} \frac{(m_{\gamma'}+p_{\gamma'}-1)!}{p_{\gamma'}!(m_{\gamma'}-1)!} \frac{1}{(z_\gamma - z_{\gamma'})^{m_{\gamma'}+p_{\gamma'}}}, \tag{15}$$

and the final expression has the form

$$\mathcal{I}^\eta_{\mathbf{X}}(z) = \sum_{j,p\in\mathbb{N}} \frac{\mathcal{A}_{j,p}}{(z-\mathcal{Z}_j)^p}. \tag{16}$$

In order to illustrate our solution, we present in tabular form (Table I) a summary of all intermediate steps, integrand parameters, and auxiliary quantities that are used in calculating the contribution for a single choice of $\{k_i\}$, in an example with $N = 6$ and $r = 4$.

Also note that if $r \notin [2, N]$ (no Matsubara frequency appearing in any exponent), the result of the integral is a number, rather than a frequency-dependent quantity. In that case, the integral can be straightforwardly generalized to the case of real time, where integrations go to some externally given time t (instead of β), and the resulting expression is a function of that time. The step given by Eq. (15) is then not needed. See Appendix A for details.

Numerical evaluation of the analytical expression and relation to other algorithms

The implementation of Eq. (14) is rather straightforward and much simpler than the algorithmic Matsubara summations in our previous work [48]. Indeed, most of the calculations just require the numerical evaluation of an analytical expression and it is not necessary to implement a dedicated symbolic algebra to manipulate the expressions. The only exception is the last step, Eq. (15). This transformation was the centerpiece of the algorithm in Ref. [48] and was applied recursively many times, leading to complex bookkeeping and data structures. Ultimately, the result was a symbolic expression that was stored, and a separate implementation was needed for the comprehension and numerical evaluation of such a general symbolic expression. In the present context, however, Eq. (15) is applied only once to produce numbers, and is simple to implement.

The other important point is that we analytically treat cases with $\delta_{\tilde{z}_i} = 1$ by employing Eq. (6). With the frequency-summation algorithms [48,49], one cannot take into account possible cancellations of the ω_i terms in Eq. (10) without computing a large number of separate analytical solutions. When untreated, these cancellations yield diverging ratios in the final expressions, which need to be regularized. On the contrary, in Eq. (14), the ratio $1/\tilde{\omega}_i^{k_i+b_i}$ cannot have a vanishing denominator and its size will, in practice, be limited by the energy resolution. This will also allow us to have the final result in the form of a sum of poles on an equidistant grid on the real axis, and extract the real-axis results without any numerical pole broadening (see Sec. III C 2 and Appendix B).

It is interesting to compare the computational effort for the numerical evaluation of our analytical solution to the straightforward numerical integration. In the most straightforward integration algorithm, one would discretize the imaginary-time interval $[0, \beta]$ with N_τ times, and then perform the summation which has the complexity $O(N_\tau^{N-1})$ for each external τ, so that overall $O(N_\tau^N)$. With our algorithm, we do not have to go through all of the configurations of internal times, but we do need to go through all of the possible permutations of the internal times, and for each permutation there is at least 2^{N-1} terms to be summed over. So the number of terms one has to sum grows at least as $O[(N-1)!2^{N-1}]$. At sufficiently high N, this number is bound to outgrow the exponential N_τ^N, whatever the N_τ. This will happen, however, only at very large N. For example, if $N_\tau = 30$, the analytical solution becomes slower at around $N = 40$. Moreover, one actually needs a much larger N_τ, especially at low temperature. In any case, the additional computational effort can be understood as coming from the difference in the information content of the result, which is a lot more substantial in the case of the analytical solution.

At orders $N < 6$ (within context of DiagMC), we find that the implementation of our algorithm is significantly more efficient than our current implementation of the Matsubara summations from Ref. [48], and at $N = 6$, they are about equally efficient. However, we anticipate that further optimizations will be possible at the level of Eq. (14).

B. Expansion of the bare propagator

The central quantity is the Green's function defined in Matsubara formalism as

$$G_{\sigma\mathbf{k}}(\tau-\tau') = -\langle T_\tau c_{\sigma\mathbf{k}}(\tau)c^\dagger_{\sigma\mathbf{k}}(\tau')\rangle = \begin{cases} -\langle c_{\sigma\mathbf{k}}(\tau)c^\dagger_{\sigma\mathbf{k}}(\tau')\rangle, & \tau > \tau' \\ \langle c^\dagger_{\sigma\mathbf{k}}(\tau')c_{\sigma\mathbf{k}}(\tau)\rangle, & \tau' > \tau, \end{cases} \quad (17)$$

where $\tau, \tau' \in [0, \beta]$. The noninteracting Green's function (or the bare propagator) in the eigenbasis of the noninteracting Hamiltonian has a very simple general form,

$$G_0(\varepsilon, i\omega) \equiv \frac{1}{i\omega - \varepsilon}, \quad (18)$$

and for the plane wave $\mathbf{k}$, the propagator is $G_{0,\mathbf{k}}(i\omega) = G_0(\varepsilon_\mathbf{k} - \mu, i\omega)$.

As we will discuss below, the diagrammatic series for the self-energy will, in general, be constructed from different powers of the bare propagator,

$$G_0^l(\varepsilon, i\omega) \equiv \frac{1}{(i\omega - \varepsilon)^l}. \quad (19)$$

Indeed, these powers naturally arise after expanding the bare propagator in a Maclaurin series, $\frac{1}{z+x} = \sum_{n=0}^{\infty} \frac{(-x)^n}{z^{n+1}}$, around a small chemical potential shift,

$$G_0(\varepsilon, i\omega) = \sum_{l=1}^{\infty}(-\delta\mu)^{l-1}G_0^l(\varepsilon + \delta\mu, i\omega). \quad (20)$$

This series converges (for all $i\omega$) if $\delta\mu$ is smaller in amplitude than the first Matsubara frequency: $|\delta\mu| < \pi T$. Nevertheless, this expression will become a part of a larger series with additional expansion parameters, which may result in a modified convergence radius of the overall series with respect to $\delta\mu$.

We anticipate that the Feynman diagrams will be formulated in the imaginary-time domain, so it is essential to work out the Fourier transform of $G_0^l(\varepsilon, i\omega)$. We present the full derivation in Appendix E and here only write the final solution,

$$G_0^l(\varepsilon, \tau - \tau') = s_{\tau,\tau'} e^{-\varepsilon(\tau-\tau')} n_\mathrm{F}(s_{\tau,\tau'}\varepsilon) \sum_{\zeta=0}^{l-1} \sum_{\varsigma=0}^{l-\zeta-1} c_{l\zeta\varsigma}^{s_{\tau,\tau'}}(\varepsilon)\, \tau^\zeta \tau'^\varsigma, \quad (21)$$

with $s_{\tau,\tau'} = \mathrm{sgn}(\tau' - \tau)$. In our notation, l in G_0^l is a superscript index, rather than the power of G_0 [although these meanings coincide in the case of $G_0^l(\varepsilon, i\omega)$]. The Fermi function is defined as $n_\mathrm{F}(\varepsilon) = 1/(e^{\beta\varepsilon} + 1)$ and the coefficients that go with the $\tau^\zeta\tau'^\varsigma$ terms are

$$c^-_{l,\zeta,\varsigma}(\varepsilon) = \sum_{n=0}^{l-\varsigma-\zeta-1} \frac{n!(-1)^{l+\varsigma-1}[-n_\mathrm{F}(\varepsilon)]^n \beta^{l-\varsigma-\zeta-1}}{(l-\varsigma-\zeta-1)!(\varsigma+\zeta)!} \times \left\{ {l-\varsigma-\zeta-1 \atop n} \right\} \binom{\varsigma+\zeta}{\zeta}, \quad (22)$$

and $c^+_{l,\zeta,\varsigma}(\varepsilon) = (-1)^{l-1} c^-_{l,\varsigma,\zeta}(-\varepsilon)$. Here we make use of binomial coefficients $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ and the Stirling number of the second kind, $\left\{ {n \atop k} \right\} = \sum_{i=0}^{k} \frac{(-1)^i}{k!}\binom{k}{i}(k-i)^n$.

C. Application to DiagMC

In the following, we apply the analytic time integral and the expansion of the bare propagator in the context of DiagMC. We discuss two kinds of self-energy series (Hartree shifted and bare) and the corresponding implementation details. Note that some symbols will be redefined with respect to previous sections.

1. Hartree-shifted series

In this section, we discuss the construction of the self-energy series, where all tadpolelike insertions are omitted in the topologies of the diagrams. Rather, the full Hartree shift is absorbed in the bare propagator. The diagrams are therefore expressed in terms of the Hartree-shifted bare propagator,

$$G^\mathrm{HF}_{0,\mathbf{k}}(i\omega) = G_0(\tilde\varepsilon_\mathbf{k}, i\omega), \quad (23)$$

with the Hartree-shifted dispersion defined as

$$\tilde\varepsilon_\mathbf{k} = \varepsilon_\mathbf{k} - \mu + U\langle n_\sigma\rangle, \quad (24)$$

where $\langle n_{\bar\sigma}\rangle$ is the average site occupation per spin.

After constructing the tadpoleless topologies, we are free to expand all propagators that appear in the diagrams according to Eq. (20):

$$G^\mathrm{HF}_{0,\mathbf{k}}(i\omega) = \sum_{l=1}^{\infty}(-\delta\mu)^{l-1}G_0^l(\tilde\varepsilon_\mathbf{k} + \delta\mu, i\omega). \quad (25)$$

In the frequency domain, this step can be viewed as introducing new topologies: we now have diagrams with any number of single-particle-vertex ($\delta\mu$) insertions on any of the propagator lines. Each arrangement of these additional single-particle vertices on the diagram does require a separate solution by the symbolic algebra algorithm, as presented in Refs. [48,49]. Nevertheless, as a $\delta\mu$ vertex cannot carry any momentum or energy, the formal effect of it is that it just raises the power l of the propagator that passes through it. In the imaginary-time domain, it turns out that the contribution of the $\delta\mu$-dressed diagrams is readily treatable by the analytical expression (14) and we no longer have to view the $\delta\mu$ insertions as changes to topology, but rather as additional internal degrees of freedom to be summed over. This is illustrated in Fig. 1.

Up to the Hartree shift, the self-energy expansion can now be made in powers of the interaction U and the small chemical-potential shift $\delta\mu$,

$$\Sigma^\mathrm{(HF)}_\mathbf{k}(\tau) = \sum_N (-U)^N \times \sum_{l_1,\dots,l_{2N-1}=1}^{\infty} (-\delta\mu)^{\sum_j (l_j-1)} \sum_{\Upsilon_N} D_{\Upsilon_N,\mathbf{k},\{l_j\},\delta\mu}(\tau), \quad (26)$$

where j enumerates the propagators, of which there are $N_\mathrm{prop} = 2N-1$, N is the perturbation order in U, each l_j goes from 1 to ∞, Υ_N enumerates distinct topologies of the diagram at order N (without any $\delta\mu$ or Hartree insertions), and D is the contribution of the diagram. The general form of the

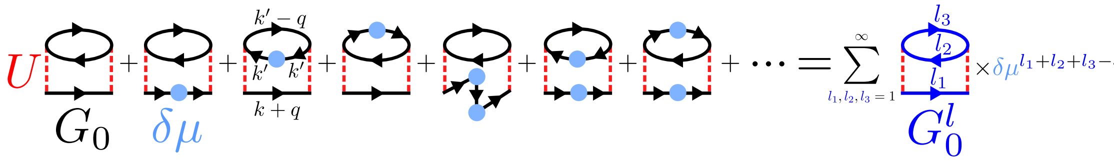


FIG. 1. Illustration of the use of the $G_0^l(\varepsilon, \tau - \tau')$ propagator. The entire series of diagrams with all possible $\delta\mu$ insertions can be captured by a single diagram with additional degrees of freedom.

diagram contribution is

$$D_{\Upsilon_N,\mathbf{k},\{l_j\},\delta\mu}(\tau) = (-1)^{N_{\text{bub}}} \prod_{i=2}^{N-1} \int_0^\beta d\tau_i \sum_{\mathbf{k}_1\ldots\mathbf{k}_N} \prod_{j=1}^{2N-1} G_0^{l_j}\left(\bar{\varepsilon}_{\tilde{\mathbf{k}}_j}, \tilde{\tau}_j - \tilde{\tau}'_j\right), \quad (27)$$

with $\bar{\varepsilon}_{\mathbf{k}} \equiv \tilde{\varepsilon}_{\mathbf{k}} + \delta\mu$. We denote N_{bub} as the number of closed fermion loops in the diagram; $\tau_1 \ldots \tau_{N-1}$ are internal times, and we fix $\tau_{i=1} = 0$; τ is the external time, $\mathbf{k}$ is the external momentum, $\mathbf{k}_1 \ldots \mathbf{k}_N$ are the independent internal momenta, j indexes the propagator lines, and $\tilde{\mathbf{k}}$ are the corresponding linear combinations of the momenta $\tilde{\mathbf{k}}_j \equiv \sum_{\lambda=0}^{N} \tilde{s}_{j\lambda}\mathbf{k}_\lambda$, where $\tilde{s}_{j\lambda} \in \{-1, 0, 1\}$, and we index with 0 the external momentum $\mathbf{k}_0 \equiv \mathbf{k}$. $\tilde{\tau}_j$ and $\tilde{\tau}'_j$ are the outgoing and incoming times for the propagator j, and take values in $\{\tau_1 \ldots \tau_N\}$, where we denote with index N the external time $\tau_N \equiv \tau$. The coefficients $\tilde{s}_{j\lambda}$, times $\tilde{\tau}_j$, $\tilde{\tau}'_j$, and the number N_{bub} are implicit functions of the topology Υ_N. Throughout the paper, we assume normalized $\mathbf{k}$ sums, $\sum_{\mathbf{k}} \equiv \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}}$, where $N_{\mathbf{k}}$ is the number of lattice sites.

We can perform the Fourier transform of the external time to obtain the contribution of the diagram in the Matsubara-frequency domain,

$$D_{\Upsilon_N,\mathbf{k},\{l_j\},\delta\mu}(i\omega) = (-1)^{N_{\text{bub}}} \prod_{i=2}^{N} \int_0^\beta d\tau_i e^{i\omega\tau_N} \sum_{\mathbf{k}_1\ldots\mathbf{k}_N} \prod_{j=1}^{2N-1} G_0^{l_j}\left(\bar{\varepsilon}_{\mathbf{k}_j}, \tilde{\tau}_j - \tilde{\tau}'_j\right). \quad (28)$$

The Green's function $G_0^l(\varepsilon, \tau - \tau')$ is discontinuous at $\tau = \tau'$, so to be able to perform the τ integrations analytically, we first need to split the integrals into ordered parts,

$$\int_0^\beta d\tau_2 \ldots \int_0^\beta d\tau_N = \sum_{(\tau_{p_2}\ldots\tau_{p_N}) \in \mathcal{P}(\{\tau_2\ldots\tau_N\})} \int_0^\beta d\tau_{p_N} \int_0^{\tau_{p_N}} d\tau_{p_{N-1}} \ldots \int_0^{\tau_{p_4}} d\tau_{p_3} \int_0^{\tau_{p_3}} d\tau_{p_2}, \quad (29)$$

where $\mathcal{P}$ denotes all $(N-1)!$ permutations of the time indices. p labels the permutation and p_i is the permuted index of vertex i.

Let us rewrite the contribution of the diagram, with propagators written explicitly using the expression (21),

$$D_{\Upsilon_N,\mathbf{k},\{l_j\},\delta\mu}(i\omega) = (-1)^{N_{\text{bub}}} \sum_{\mathbf{k}_1\ldots\mathbf{k}_N} \sum_{(\tau_{p_2}\ldots\tau_{p_N}) \in \mathcal{P}(\{\tau_2\ldots\tau_N\})} (-1)^{N_{\text{fwd}}(p)} \prod_j n_{\text{F}}\left(s_j \bar{\varepsilon}_{\tilde{\mathbf{k}}_j}\right) \sum_{\zeta_j=0}^{l_j-1} \sum_{\varsigma_j=0}^{l_j-\zeta_j-1} c^{s_j}_{l_j,\zeta_j,\varsigma_j}\left(\bar{\varepsilon}_{\tilde{\mathbf{k}}_j}\right) \prod_{j\in\mathcal{J}_{\text{i}}(i=1)} \delta_{\zeta_j} \prod_{j\in\mathcal{J}_{\text{o}}(i=1)} \delta_{\varsigma_j}$$

$$\times \int_0^\beta d\tau_{p_N} \int_0^{\tau_{p_N}} d\tau_{p_{N-1}} \ldots \int_0^{\tau_{p_4}} d\tau_{p_3} \int_0^{\tau_{p_3}} d\tau_{p_2} e^{i\omega\tau_N} \prod_{i=2}^{N} \tau_i^{\sum_{j\in\mathcal{J}_{\text{i}}(i)} \zeta_j + \sum_{j\in\mathcal{J}_{\text{o}}(i)} \varsigma_j} e^{\tau_i\left(\sum_{j\in\mathcal{J}_{\text{o}}(i)} \bar{\varepsilon}_{\tilde{\mathbf{k}}_j} - \sum_{j\in\mathcal{J}_{\text{i}}(i)} \bar{\varepsilon}_{\tilde{\mathbf{k}}_j}\right)}, \quad (30)$$

where $\mathcal{J}_{\text{i/o}}(i)$ is the set of incoming/outgoing propagators j of the vertex i, which depends on the topology Υ_N. We also introduced shorthand notation $s_j = s_{\tilde{\tau}_j,\tilde{\tau}'_j}$. Practically, s_j depends on whether $p(i(j)) > p(i'(j))$ or the other way around, where $i(j)/i'(j)$ is the outgoing/incoming vertex of propagator j in the given permutation p. The total number of forward-facing propagators is $N_{\text{fwd}}(p) = \sum_j \delta_{-1,s_j}$, which depends on the permutation and the topology. The products of δ_{ζ_j} and δ_{ς_j} are there to ensure that the time $\tau_1 = 0$ is not raised to any power other than 0, as such terms do not contribute.

Now we can apply the analytic solution for the time integrals [Eq. (14)] to arrive at the final expression:

$$D_{\Upsilon_N,\mathbf{k},L,\delta\mu}(z) = (-1)^{N_{\text{bub}}} \sum_{\{\tilde{l}_j \geqslant 0\}: \sum_j \tilde{l}_j = L} \sum_{\mathbf{k}_1\ldots\mathbf{k}_N} \sum_{(\tau_{p_2}\ldots\tau_{p_N}) \in \mathcal{P}(\{\tau_2\ldots\tau_N\})} (-1)^{N_{\text{fwd}}(p)}$$

$$\times \prod_j n_{\text{F}}\left(s_j \bar{\varepsilon}_{\tilde{\mathbf{k}}_j}\right) \sum_{\zeta_j=0}^{\tilde{l}_j} \sum_{\varsigma_j=0}^{\tilde{l}_j-\zeta_j} c^{s_j}_{\tilde{l}_j+1,\zeta_j,\varsigma_j}\left(\bar{\varepsilon}_{\tilde{\mathbf{k}}_j}\right) \prod_{j\in\mathcal{J}_{\text{i}}(i=1)} \delta_{\zeta_j} \prod_{j\in\mathcal{J}_{\text{o}}(i=1)} \delta_{\varsigma_j} \mathcal{I}^{\eta=-1}_{\mathbf{X}}(z),$$

$$\mathbf{X} = \left(p(N), \left\{ \sum_{j\in\mathcal{J}_{\text{i}}(i(p_{i'}))} \zeta_j + \sum_{j\in\mathcal{J}_{\text{o}}(i(p_{i'}))} \varsigma_j \right\}_{i'=2\ldots N}, \left\{ \sum_{j\in\mathcal{J}_{\text{o}}(i(p_{i'}))} \bar{\varepsilon}_{\tilde{\mathbf{k}}_j} - \sum_{j\in\mathcal{J}_{\text{i}}(i(p_{i'}))} \bar{\varepsilon}_{\tilde{\mathbf{k}}_j} \right\}_{i'=2\ldots N} \right), \quad (31)$$

where $i(p_{i'})$ is the vertex index i of the permuted index $p_{i'}$ and we have introduced a new expansion variable $L = \sum_j (l_j - 1)$ and a convenient variable $\tilde{l}_j = l_j - 1$, so that

$$\Sigma_{\mathbf{k}}^{(\mathrm{HF})}(z) = \sum_{K=2}^{\infty} \sum_{N=2}^{K} \sum_{L=0}^{K-N} (-U)^N (-\delta\mu)^L \sum_{\Upsilon_N} D_{\Upsilon_N,\mathbf{k},L,\delta\mu}(z), \tag{32}$$

which is the series that we implement and use in practice. The meaning of K is the number of all independent (internal and external) times in the diagram. Note that in $\mathcal{D}$, we perform only $N-1$ integrations over time. Those are the times associated with N interaction vertices, minus the one that is fixed to zero. The integrations of the times associated with $\delta\mu$ insertions have already been performed in Eq. (21), and there are L such integrals. Overall, the number of independent times is $K = N + L$. Ultimately, we group contributions by the expansion order K and look for convergence with respect to this parameter.

2. Numerical implementation of DiagMC and relation to other algorithms

The expression (31) is very convenient for numerical evaluation. First, we restrict the values of $\bar{\varepsilon}_{\mathbf{k}}$ to a uniform grid on the real axis with the step $\Delta\omega$ ($\bar{\varepsilon}_{\mathbf{k}} = j\Delta\omega$). These appear in $\omega_2, ..., \omega_K$ as terms with integer coefficients, which means that $\{\omega_i\}$ entering $\mathcal{I}_{\mathbf{X}}$ will also be restricted to the same uniform grid. The final result therefore has the form

$$D_{\Upsilon_N,\mathbf{k},L,\delta\mu}(z) = \sum_{j\in\mathbb{Z},p\in\mathbb{N}} \frac{\mathcal{A}_{j,p}}{(z - j\Delta\omega)^p}. \tag{33}$$

This form allows us to reinterpret the finite-lattice results as that of the thermodynamic limit and extract $D_{\Upsilon_N,\mathbf{k},L,\delta\mu}(\omega + i0^+)$ without any numerical broadening (see Appendix B for details).

In our present implementation, we perform a flat-weight (uniform) MC sampling over internal momenta $\{\mathbf{k}_i\}$, do a full summation of all the other sums, and accumulate the amplitudes $\mathcal{A}_{j,p}$. There are, however, other options. For example, one may sample $\{\mathbf{k}_i\}, \{p_i\}, \{b_i\}$ and use $P \equiv \prod_j n_{\mathrm{F}}(s_j \bar{\varepsilon}_{\tilde{\mathbf{k}}_j}) e^{b_N \beta \tilde{\omega}_N}$ as the weighting function. We have thoroughly checked that the factor P closely correlates with the contribution to $\mathcal{A}_{j,p}$ coming from a given choice of the $\{\mathbf{k}_i\}, \{p_i\}, \{b_i\}$ variables (with other variables summed over), and thus P could be a good choice for a weighting function. However, this requires additional operations related to move proposals and trials, and we have not yet been able to make such an algorithm more efficient than the flat-weight MC. Nevertheless, it is apparent that our approach offers more flexibility than the algorithmic Matsubara summations (AMS). In AMS, no convenient weighting function can be defined for the Monte Carlo, so one either does the flat-weight summation [48] or uses the whole contribution to the result as the weight, which comes at the price of having to repeat the calculation for each frequency of interest [49] (on the contrary, in Ref. [48], as well as in this paper, the entire frequency dependence of the self-energy is obtained in a single MC run). At present, it is unclear which scheme is best—whether one should evaluate $D(z)$ one z at a time or capture all z at once as we do here. This choice, as well as the choice of the weighting function, likely needs to be made on a case-by-case basis, as it is probable that in different regimes, different approaches will be optimal. In that sense, the added flexibility of our time-integration approach in terms of the choice of the weighting function may prove valuable in the future.

Concerning floating-point arithmetic, it is important that the factor $e^{b_N \beta \tilde{\omega}_N}$ stemming from $\mathcal{I}_{\mathbf{X}}$ can always be absorbed into the product of n_{F} functions in the second row of Eq. (31). This can be understood as follows. A given $\bar{\varepsilon}_{\tilde{\mathbf{k}}_j}$ can, at most, appear twice as a term in $\tilde{\omega}_N$, once with sign $+1$ and once with sign -1, corresponding to the incoming $\tilde{\tau}'_j$ and outgoing $\tilde{\tau}_j$ ends of the propagator j. In that case, the exponent cancels. The other possibility is that it appears only once, in which case it must correspond to the later time in the given permutation. If the later time is the outgoing end of the propagator, then the propagator is forward facing and the sign in front is $s = -1$; if it is the incoming end, then the propagator is backward facing and the sign in front is $s = 1$. In both cases, we can make use of

$$e^{s\beta\varepsilon} n_{\mathrm{F}}(s\varepsilon) = n_{\mathrm{F}}(-s\varepsilon). \tag{34}$$

Therefore, no exponentials will appear in the final expression. A product of n_{F} functions is, at most, 1 and the coefficients c are not particularly big. Then, the size of the pole amplitudes that come out of Eq. (14) is determined by the energy resolution ($1/\Delta\omega$) and temperature ($\beta^{n_N+1-b_N-k_N}$). In our calculations so far, the amplitudes remain relatively small. Our approach ensures that we do not have very large canceling terms, such as we had in Ref. [48]. Indeed, we have successfully implemented Eq. (31) without the need for multiple-precision floating-point types.

Compared to the Matsubara-frequency summation algorithm [47–49], Eq. (31) presents an improved generality. Equation (31) is valid for any number and arrangement of instantaneous (i.e., frequency-independent) insertions, i.e., any choice of $\{\tilde{l}_j\}$. In contrast, the algorithmic Matsubara summation has to be performed for each choice of $\{\tilde{l}_j\}$ independently, and the resulting symbolic expressions need to be stored. For example, at $N = 4$, we have 12 Υ_N topologies. Therefore, at $L = 0$, the number of analytical solutions to prepare is 12. However, at $L = 2$, this number is 336, i.e., 28-fold bigger (we can place $L = 2$ insertions on $2N - 1 = 7$ fermionic lines in $7 \times 6/2 + 7 = 28$ ways, times 12 Υ_N topologies, i.e., 336).

3. Bare series

We are also interested in constructing a bare series where tadpole insertions are present in diagram topologies. Tadpole (or Hartree) insertions are instantaneous and an evaluation of their amplitudes can be done relatively simply by various means. At the level of the Hubbard model, the Hartree insertions factor out: For each Hartree diagram, the internal momentum summations and time integrations can be performed beforehand and only once, leading to a significant speedup.

In the expression (31), there is no difference between a Hartree insertion and a chemical-potential vertex insertion. Therefore, the inclusion of the Hartree insertions can be en-

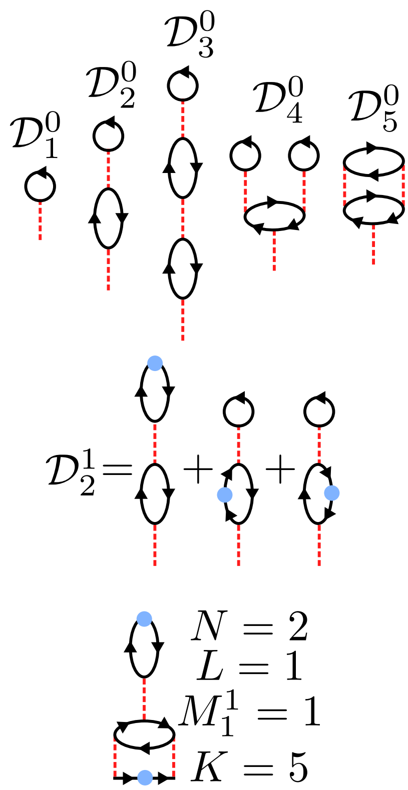


FIG. 2. Top: Illustration of possible Hartree diagrams, without any $\delta\mu$ insertions. Middle: Amplitude of a Hartree diagram with a single $\delta\mu$ insertion. Bottom: An example of a diagram dressed with both Hartree and $\delta\mu$ insertions, and the values of the parameters $N, L, \{M_i^{L'}\}, K$ that it falls under (with $M_{i\neq 1}^{L'\neq 1} = 0$).

tirely accounted for in the resummation of the $D_{\Upsilon_N,\mathbf{k},L,\delta\mu}(z)$ contributions from the previous section, with the replacement

$$\bar{\varepsilon}_{\mathbf{k}} \equiv \varepsilon_{\mathbf{k}} - \mu + \delta\mu \tag{35}$$

(i.e., full Hartree shift excluded).

Note that the expansion of the propagators in $\delta\mu$ is performed in Hartree insertions as well, so we need to account for possible additional $\delta\mu$ insertions inside the Hartree diagrams. As before, our expansion order will be K, which is the total number of independent times, with each time associated to a single interaction or a $\delta\mu$ vertex, including those within Hartree insertions.

We will for now focus on the series up to $K = 5$. As the number of interactions in Υ_N is at least two, we can have, at most, three interaction vertices in a Hartree insertion. There are only five such Hartree diagrams (Fig. 2). We can evaluate these five amplitudes with very little effort by making use of spatial and temporal Fourier transforms.

Before we proceed with the calculation of the amplitudes $\mathcal{D}$ of possible Hartree insertions relevant for the series up to $K = 5$, we define some auxiliary quantities. We first define the bare density,

$$n_0^{\tilde{l}} = \sum_{\mathbf{k}} G_0^{l=1+\tilde{l}}(\bar{\varepsilon}_{\mathbf{k}}, \tau = 0^-), \tag{36}$$

and the real-space propagator,

$$G_{0,\mathbf{r}}^{l=1+\tilde{l}} = \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} G_0^{l=1+\tilde{l}}(\bar{\varepsilon}_{\mathbf{k}}, \tau = 0^-). \tag{37}$$

We will also need the polarization bubble diagram,

$$\chi_{0,\mathbf{r}}^{\tilde{l}_1,\tilde{l}_2}(\tau) = G_{0,\mathbf{r}}^{l=1+\tilde{l}_1}(\tau) G_{0,-\mathbf{r}}^{l=1+\tilde{l}_2}(-\tau), \tag{38}$$

$$\chi_{0,\mathbf{q}=0}^{\tilde{l}_1,\tilde{l}_2}(i\nu = 0) = \sum_{\mathbf{r}} \int d\tau\, \chi_{0,\mathbf{r}}^{\tilde{l}_1,\tilde{l}_2}(\tau), \tag{39}$$

and the second-order self-energy diagram (up to the constant prefactor),

$$\Sigma_{2,\mathbf{r}}^{\tilde{l}_1,\tilde{l}_2,\tilde{l}_3}(\tau) = G_{0,\mathbf{r}}^{l=1+\tilde{l}_1}(\tau) \chi_{0,\mathbf{r}}^{\tilde{l}_2,\tilde{l}_3}(\tau), \tag{40}$$

which can be Fourier transformed to yield $\Sigma_{2,\mathbf{k}}^{\tilde{l}_1,\tilde{l}_2,\tilde{l}_3}(i\omega)$.

We can now calculate the amplitudes of the possible Hartree insertions with a number L of $\delta\mu$ insertions on them, in any arrangement

$$\mathcal{D}_1^L = (-)n_0^L, \tag{41}$$

$$\mathcal{D}_2^L = (-)^2 \sum_{\substack{\tilde{l}_1,\tilde{l}_2,\tilde{l}_3 \\ \tilde{l}_1+\tilde{l}_2+\tilde{l}_3=L}} n_0^{\tilde{l}_1} \chi_{0,\mathbf{q}=0}^{\tilde{l}_2,\tilde{l}_3}(i\nu = 0), \tag{42}$$

$$\mathcal{D}_3^L = (-)^3 \sum_{\substack{\tilde{l}_1,\ldots,\tilde{l}_5 \\ \sum_i \tilde{l}_i = L}} n_0^{\tilde{l}_1} \chi_{0,\mathbf{q}=0}^{\tilde{l}_2,\tilde{l}_3}(i\nu = 0) \chi_{0,\mathbf{q}=0}^{\tilde{l}_4,\tilde{l}_5}(i\nu = 0), \tag{43}$$

$$\mathcal{D}_4^L = (-)^3 \sum_{\substack{\tilde{l}_1,\ldots,\tilde{l}_3 \\ \sum_i \tilde{l}_i = L}} \binom{2+\tilde{l}_3}{2} n_0^{\tilde{l}_1} n_0^{\tilde{l}_2} n_0^{2+\tilde{l}_3}, \tag{44}$$

$$\mathcal{D}_5^L = (-)^2 \sum_{\substack{\tilde{l}_1,\ldots,\tilde{l}_5 \\ \sum_i \tilde{l}_i = L}} T \sum_{i\omega} e^{-i\omega 0^-} \sum_{\mathbf{k}} G_{0,\mathbf{k}}^{l=1+\tilde{l}_1}(i\omega) \Sigma_{2,\mathbf{k}}^{\tilde{l}_2,\tilde{l}_3,\tilde{l}_4}(i\omega) G_{0,\mathbf{k}}^{l=1+\tilde{l}_5}(i\omega). \tag{45}$$

As we are restricting to $K \leqslant 5$ calculations, the $\mathcal{D}_{3\ldots5}^L$ insertions can only be added once, and only with $L = 0$. We now define M_i^L as the number of insertions of $\mathcal{D}_i^L$ tadpoles, and we define N_i as the number of interaction vertices contained in the tadpole $\mathcal{D}_i$ (regardless of L, we have $N_1 = 1, N_2 = 2, N_3 = N_4 = N_5 = 3$).

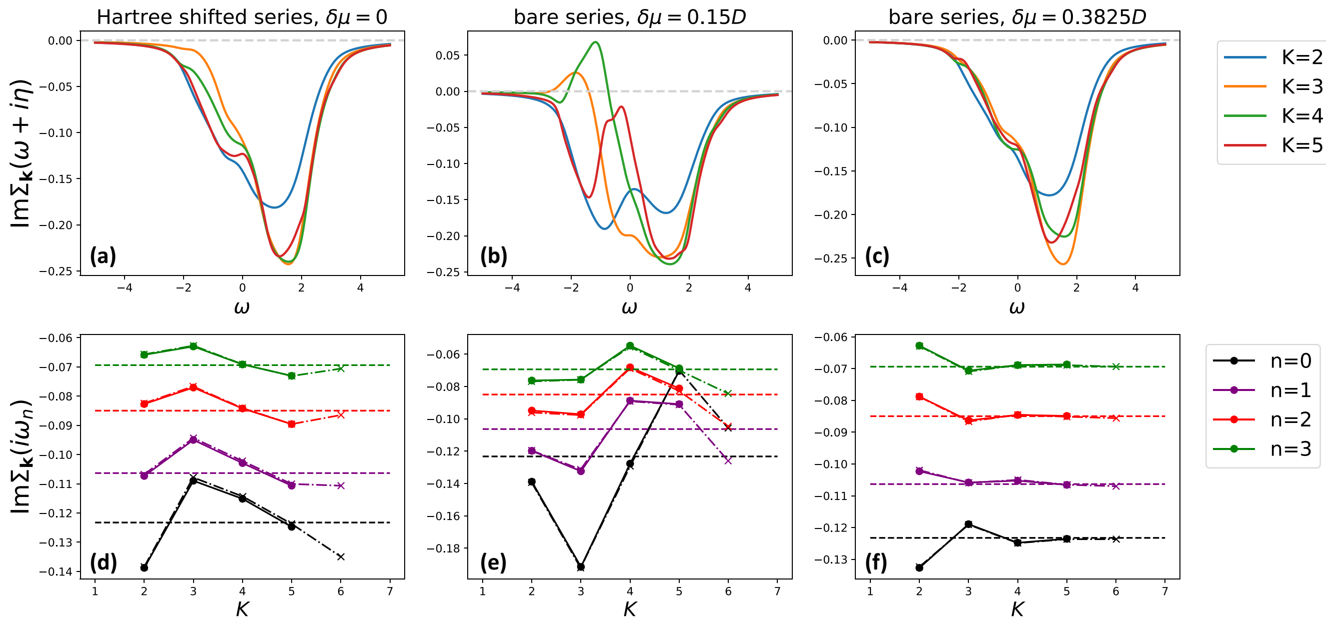


FIG. 3. DiagMC solution for the Hubbard model on a square lattice. Parameters of the model are $t' = -0.3t$, $\mu = 0$, $U = 1D$, and $T = 0.125D$, which corresponds to $\langle n_\sigma \rangle \approx 0.3625$. Top row: Imaginary part of self-energy at $\mathbf{k} = (\pi/4, \pi)$ on the real axis (with broadening $\eta = 0.3D$) obtained with three different series, up to perturbation order K. Bottom row: Illustration of convergence with respect to perturbation order K, using values of the imaginary part of the self-energy at the lowest four Matsubara frequencies, $i\omega_{n=0...3}$. Full lines are our result, dash-dotted lines with crosses are the analogous result with a numerical τ-integration algorithm from Ref. [52], and horizontal dashed lines are the determinantal QMC result on a 16 × 16 lattice from Ref. [52].

The series can now be resummed as

$$\Sigma_{\mathbf{k}}^{(\mathrm{HF})}(z) = \sum_{K=2}^{\infty}\sum_{N=2}^{K}\sum_{L=0}^{K-N} \sum_{\substack{\{M_i^{L'}\}=0 \\ N+L+\sum_{i,L'} M_i^{L'}(N_i+L')=K}}^{K-N-L} (-U)^{N+\sum_{i,L'} M_i^{L'} N_i} (-\delta\mu)^{L+\sum_{i,L'} M_i^{L'} L'} \prod_{i,L'} \left(\mathcal{D}_i^{L'}\right)^{M_i^{L'}} \Omega\left(L, \left\{M_i^{L'}\right\}\right)$$
$$\times \sum_{\Upsilon_N} D_{\Upsilon_N,\mathbf{k},L+\sum_{i,L'} M_i^{L'}}(z), \qquad (46)$$

where $\Omega(L, \{M_i^{L'}\})$ is the combinatorial prefactor which counts all the possible ways the selected single-particle vertices $\delta\mu, \{\mathcal{D}_i\}$ can be arranged. This corresponds to the number of permutations of the multisets,

$$\Omega\left(L, \left\{M_i^{L'}\right\}\right) = \frac{\left(L + \sum_{i,L'} M_i^{L'}\right)!}{L! \prod_{i,L'} M_i^{L'}!}. \qquad (47)$$

We emphasize that Eq. (46) is fully general, but at orders $K \geqslant 5$, additional Hartree insertions $\mathcal{D}$ [compared to Eqs. (41)–(45)] need to be considered.

Finally, we stress that our analytical time-integral solution and action-shift tuning scheme in DiagMC are not restricted to the treatment of the Hubbard Hamiltonian. See Appendix F for a discussion of DiagMC in the case of a general Hamiltonian with two-body interactions.

IV. RESULTS

A. Convergence speedup with $\delta\mu$ expansion in the bare series

Here we focus on supplementing the results from Ref. [52] with real-frequency self-energies calculated without any numerically ill-defined analytical continuation.

The model parameters are $t' = -0.3t$, $\mu = 0$, $U = 1.0D$, $T = 0.125D$, and $\langle n_\sigma \rangle = 0.3625$. In Ref. [52], the calculation was performed with the Hartree-shifted series with $\delta\mu = 0$, as well as with the bare series, with two values of $\delta\mu$, namely, $0.15D$ and $0.3825D$. We repeat these calculations with our method. We use lattice size 32 × 32, and project the dispersion onto a uniform energy grid, as described in Ref. [48] and discussed in Sec. III C 2. In Fig. 3, we show our results and compare them with the results of Ref. [52].

In the upper row of Fig. 3 are the real-frequency self-energies calculated up to order $K \leqslant 5$. We are keeping a finite broadening $\eta = 0.3D$ to smoothen the curves. As discussed in Appendix B, in our method, numerical pole broadening is not a formal necessity. However, there is still a significant amount of statistical noise in our real-frequency result (although the imaginary-frequency result is already very well converged). It is important to note that some of the noisy features in our real-frequency result may be artifacts of the finite-lattice size that would not vanish with increasing number of MC steps. However, by comparing the result with a 256 × 256 lattice calculation (Appendix C), we check that already at $\eta = 0.2D$, no such artifact should be visible. It appears that for the given external **k** and broadening $\eta = 0.2D$, increasing the lattice size further from 32×32 brings no new information, but it also does not present an additional cost: at $\eta = 0.2D$, our 256×256 lattice calculation appears equally well converged as the 32×32 lattice calculation, with the equal number of MC steps and a similar runtime, and yields a result that is on top of the 32×32 calculation.

In the bottom row of Fig. 3, we show the change in the imaginary part of the self-energy at the first four Matsubara frequencies, as a function of the maximal order K. Full-line and dots are the result of our calculations. The dash-dotted lines with crosses are data points taken from Ref. [52]. The horizontal dashed lines are the 16×16-lattice determinantal QMC result, also from Ref. [52].

The excellent agreement with the results from Ref. [52] serves as a stringent test of our implementation. In the $\delta\mu = 0.3825D$ calculation, even on the real axis, the self-energy does appear well converged by order $K = 5$, although there is some discrepancy between $K = 4$ and $K = 5$ at around $\omega = 1.5D$.

B. ω-resolved resummation

We can now go one step further by resumming the series presented in Figs. 3(a) and 3(c) for each ω individually, using an ω-dependent optimal shift $\delta\mu^*(\omega)$. The results are shown in Figs. 4 and 5.

We determine the optimal $\delta\mu^*(\omega)$ by minimizing the spread of the $\mathrm{Im}\Sigma(\omega + i\eta)$ results between orders $K = 3$ and $K = 5$. This spread as a function of ω and $\delta\mu$ is color plotted in Figs. 4 and 5. We have results for a discrete set of $\delta\mu \in \{\delta\mu_i\}$, so the optimal $\delta\mu^*(\omega)$ is *a priori* a discontinuous curve. As this is clearly nonsatisfactory, we smoothen the curve (shown with the blue line on the top panels in Figs. 4 and 5). However, we do not have results for each precise value of this optimal $\delta\mu^*(\omega)$. One could take, for each ω, the available $\delta\mu_i$ that is closest to $\delta\mu^*(\omega)$, but this would, again, result in a discontinuous curve. To avoid this, we average the available results as

$$\Sigma(\omega) = \frac{\sum_i \Delta\delta\mu_i \Sigma(\omega; \delta\mu_i) w(\delta\mu^*(\omega), \delta\mu_i)}{\sum_i \Delta\delta\mu_i w(\delta\mu^*(\omega), \delta\mu_i)}, \quad (48)$$

where $\Delta\delta\mu_i$ is the size of the $\delta\mu$ step in the available results at the ith value (allows for nonuniform grids). We use a narrow Gaussian weighting kernel,

$$w(\delta\mu^*(\omega), \delta\mu_i) = e^{-(\delta\mu_i - \delta\mu^*(\omega))^2/W^2}. \quad (49)$$

The width of the kernel W is chosen such that it is as narrow as possible, while still encompassing at least 3–4 $\delta\mu_i$ points, so that the final result is reasonably smooth as a function of ω; W is therefore determined according to the resolution in $\delta\mu$. We use $W = 0.05$ and $\Delta\delta\mu_i \approx 0.02$ and have checked that the results are insensitive to the precise choice of this numerical parameter.

The results of the averaging around the optimal $\delta\mu^*(\omega)$ are shown in the middle and bottom panels of Figs. 4 and 5. In both cases, the ω-resolved resummation helps to converge the result. In the case of the bare series, the convergence is now almost perfect, and already order $K = 3$ is on top of the exact result. In the case of the Hartree-shifted series, the results are

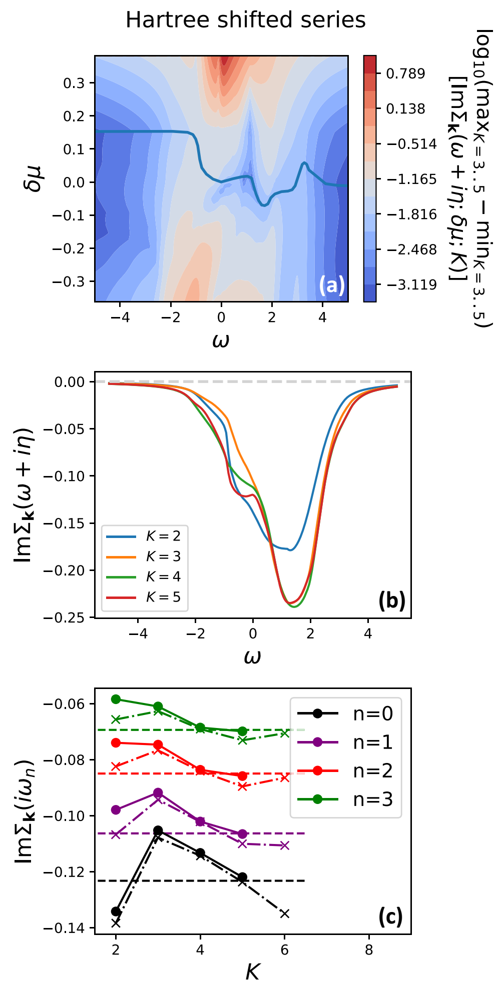


FIG. 4. Results of the Hartree-shifted series with ω-resolved resummation, to be compared to Figs. 3(a) and 3(d) (all parameters are the same). Top panel: Color plot of the spread of the imaginary part of the self-energy at a given $\omega + i\eta$ between orders $K = 3$ and 5, in a calculation with a given $\delta\mu$. The blue line smoothly connects the minima of the spread (at each ω), and defines the ω-dependent optimal shift $\delta\mu^*(\omega)$ used in the resummation. Middle and bottom panels are analogous to Figs. 3(a) and 3(d). In the bottom panel, the dash-dotted and dashed lines are the same as in Fig. 3(d).

not perfectly converged at $\omega < 0$, yet the $K = 5$ calculation is practically on top of the exact result on the imaginary axis, and presents an improvement to the $\delta\mu = 0$ series in Fig. 3(a). Note that the improvement in convergence is seen on the imaginary axis, as well.

C. Removing nonphysical features

In this section, we focus on the parameters case discussed in Ref. [48]. We calculate the Hartree-shifted series with parameters of the model $t' = 0$, $\mu - U\langle n_\sigma \rangle = -0.1D$, $T = 0.1D$, and employ various $\delta\mu$ shifts. The lattice size is again 32×32 and we focus on the self-energy at $\mathbf{k} = (0, \pi)$. Note that in Hartree-shifted series, the quantity that enters the calculation is $\mu - U\langle n_\sigma \rangle$, rather than μ. If $\langle n \rangle$ is calculated, μ can be estimated *a posteriori*. In our calculation, we fix

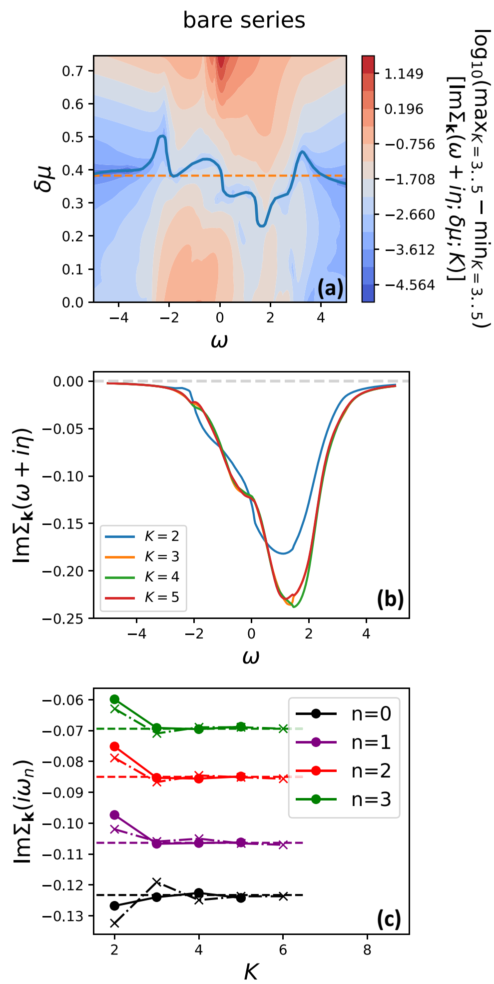


FIG. 5. Results of the bare series with ω-resolved resummation, to be compared to Figs. 3(c) and 3(f) (all parameters are the same). The top panel is analogous to Fig. 4(a). The horizontal orange dashed line denotes the value of $\delta\mu$ used in Figs. 3(c) and 3(f) to best converge the imaginary-axis result. The middle and bottom panels are analogous to Figs. 3(c) and 3(f). In the bottom panel, the dash-dotted and dashed lines are the same as in Fig. 3(f).

$\mu - U\langle n_\sigma\rangle$, and $\langle n_\sigma\rangle$ is then U dependent. Roughly, as given in Ref. [48], at $U = 1$, we have $\langle n_\sigma\rangle \approx 0.455$.

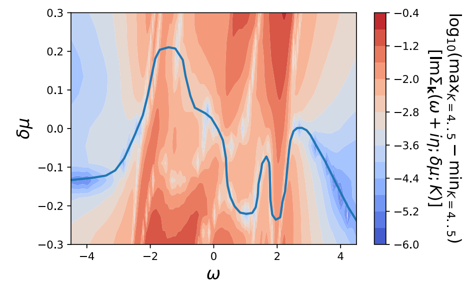


FIG. 7. Analogous to Fig. 4(a), for the parameters of the model corresponding to Fig. 6. The blue line is the optimal $\delta\mu^*$, to be used in Fig. 8.

The results are presented in Fig. 6 for three values of U. At low U, the series is well converged by $K = 5$, and the result is entirely insensitive to the choice of $\delta\mu$, as expected. At intermediate and high U, the result can be strongly $\delta\mu$ sensitive. The $\delta\mu$ dependence of the result, however, strongly varies with ω. It appears that for a given ω, there are ranges of the $\delta\mu$ value where the result (at fixed order K) is insensitive to the precise choice of $\delta\mu$. This presents an alternative way of choosing an optimal $\delta\mu$ (a similar idea was employed in a different context in Ref. [58]).

The striking feature at large U is the causality violations at $|\omega| \approx 2$ that were previously discussed in Ref. [48] (note that the broadening somewhat masks the extent of the problem). The dips in the self-energy spectrum appear to happen only at certain values of $\delta\mu$: at $\omega = -2$, the problem is present at $\delta\mu$ large and negative, and at $\omega = 2$, at $\delta\mu$ large and positive. In particular, at $\omega = 2$, the result appears to vary uniformly with $\delta\mu$, and one cannot select an optimal $\delta\mu$ based on the sensitivity of the result to the $\delta\mu$ value. We therefore repeat the procedure from the previous section and select the optimal $\delta\mu^*(\omega)$ based on the level of convergence between orders $K = 4$ and $K = 5$. The spread of the results and a smooth choice of $\delta\mu^*(\omega)$ are presented in Fig. 7.

In Fig. 8, the results of the averaging are shown and compared to the $\delta\mu = 0$ results at the highest available orders $K = 4$ and $K = 5$, at three values of U. The convergence is visibly better around our $\delta\mu^*$ than with $\delta\mu = 0$ at problematic frequencies $|\omega| \approx 2$. More importantly, the nonphysical features are clearly absent. At $U = 1$, in the $\delta\mu = 0$

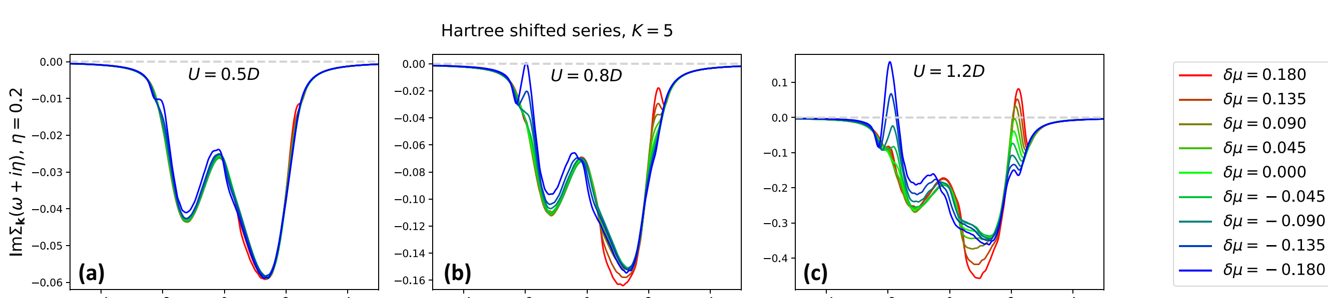


FIG. 6. Imaginary part of the self-energy on the real axis (with broadening η), at different values of coupling constant U, obtained with our method at $K = 5$ using different chemical-potential shifts $\delta\mu$. The parameters of the calculation are the same as in Ref. [48], i.e., $t' = 0$, $\mu - U\langle n_\sigma\rangle = -0.1D$, $T = 0.1D$. The self-energy is calculated at $\mathbf{k} = (0, \pi)$. Passing of the curves above the gray dashed line indicates breaking of causality.

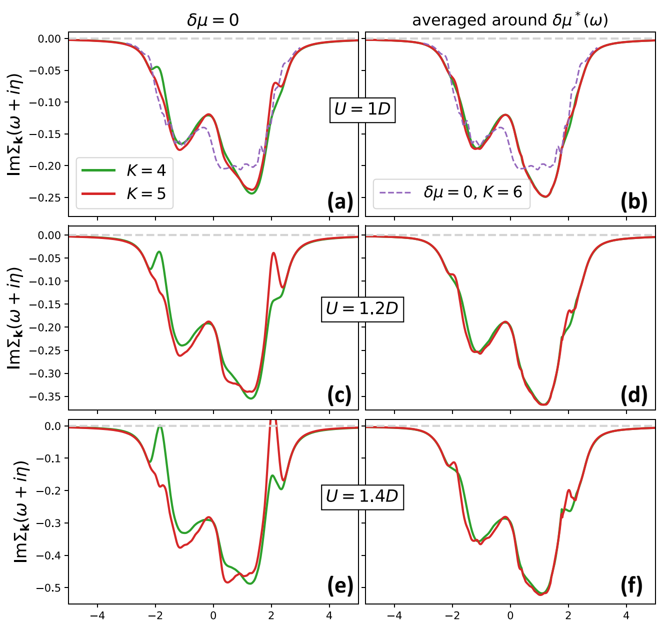


FIG. 8. Imaginary part of self-energy, real-frequency results (with broadening η). Right column: obtained with the ω-resolved resummation for the model parameters from Fig. 6, using the optimal $\delta\mu^*(\omega)$ from Fig. 7; to be compared to the standard $\delta\mu = 0$ calculation in the left column. Purple dashed lines in the top row are the $K = 6$ calculation with $\delta\mu = 0$.

calculation, the causality is not yet violated, but the dip at $\omega = 2$ is already starting to appear, which is clearly an artifact of the series truncation which should be removed systematically. It is important that the intermediate frequency behavior that we obtained by averaging results around the optimal $\delta\mu$ is indeed the correct one, and it will not change much further with increasing orders. We show in the top panels the $K = 6$, the $\delta\mu = 0$ result of which has been benchmarked against a fully converged imaginary-axis result in Fig. 9 (the converged result was obtained with the ΣDet method [59,60] at order 8). Clearly, the improved convergence between orders 4 and 5 that we have achieved by choosing $\delta\mu$ appropriately does indeed mean an improved final result. However, our procedure does not improve the result at around $\omega = 0$, where the optimal $\delta\mu$ does appear to be close to 0. The $K = 6$, $\delta\mu = 0$ result shown in the upper panels of Fig. 8 is still a bit different from the $K = 5$, $\delta\mu \approx \delta\mu^*(\omega)$ results around $\omega = 0$.

In the case of $U = 1D$, it is interesting that a large negative $\delta\mu$ does bring the $\omega \approx 0$ result at order $K = 5$ much closer to the exact value. This can be anticipated from Fig. 6, where we show the corresponding results for $U = 0.8D$ and $U = 1.2D$. Also, by looking at the color plot in Fig. 7, we see that at $\omega = 0$, there is indeed a local minimum in the spread at around $\delta\mu = -0.2$, which could be used as the optimal $\delta\mu^*$. This minimum, however, cannot be continuously connected with the other minima that we observe at $\omega < 0$, so we chose a different trajectory in the $(\omega, \delta\mu)$ space. It would be interesting for future work to inspect the behavior at even more negative $\delta\mu$, where another continuous trajectory $\delta\mu^*(\omega)$ might be found.

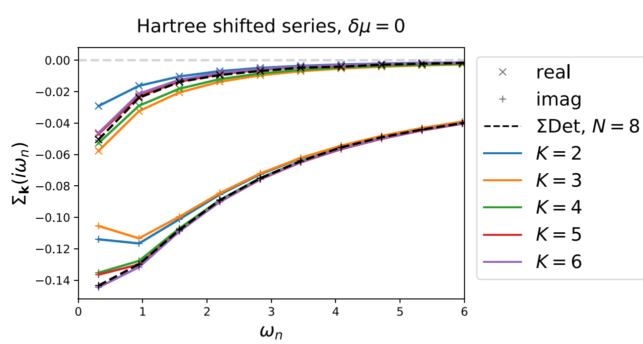


FIG. 9. Matsubara-frequency self-energy result, with model parameters as in Fig. 6. Crosses are the real part, pluses are the imaginary part, and lines are eye guides. Solid lines are the Hartree-shifted series with $\delta\mu = 0$ at different maximal K. The same result was obtained with both the algorithm presented in this work and the algorithmic Matsubara summation method from Ref. [48] (the two methods were compared diagram by diagram). Black dashed lines are the ΣDet result at maximal order $N = 8$.

V. DISCUSSION, CONCLUSIONS, AND PROSPECTS

In this paper, we have derived an analytical solution for the multiple-time integral that appears in the imaginary-time Feynman diagrams of an interaction series expansion. The solution is general for any diagram with a single external time or no external times. We find this generality to be a great advantage compared to the recently proposed algorithmic solutions of the corresponding Matsubara-frequency summations. Our analytical solution allowed us to develop a very flexible DiagMC algorithm that can make use of the possibility to optimize the series with shifted actions. As a result, we were able to almost perfectly converge a real-frequency self-energy in just 3–4 orders of perturbation, in a nontrivial regime and practically in the thermodynamic limit.

More importantly, the fact that one does not have to prepare a solution for each diagram topology individually opens the possibility to develop algorithms more akin to CTINT and allow the MC sampling to go to indefinite perturbation orders. In fact, upon a simple inspection of CTINT and continuous-time hybridization-expansion quantum Monte Carlo in the segment picture (segment-CTHYB) equations [42], it becomes clear that our solution can, in principle, be applied there, so as to reformulate these methods in real frequency. This would, however, come at the price of having to break into individual terms the determinant that captures all the contributions to the partition function at a given perturbation order. In turn, this may lead to a more significant sign problem, and an effective cap on the perturbation orders that can be handled in practice. On the other hand, it is not entirely clear how much of the sign problem comes from summing the individual terms and how much from the integration of the internal times, and we leave such considerations for future work. In any case, DiagMC algorithms based on hybridization expansion have been proposed before (see Refs. [23,28,61]), where our analytical solution may be applied.

Our solution also trivially generalizes to real-time integrals and may have use in Keldysh and Kadanoff-Baym [9] calculations, where the infamous dynamical sign problem arises precisely due to oscillating time integrands. There have been recent works [62,63] with imaginary-time propagation of randomized walkers where our solution may also find application.

Finally, we emphasize that avoiding analytical continuation could be beneficial at high temperature where the Matsubara frequencies become distant from the real axis, and thus noisy imaginary-axis correlators contain little information [64,65]. The high-temperature regime is particularly relevant for optical lattice simulations of the Hubbard model [66]. In that context, we anticipate our method will find application in the calculation of conductivity and other response functions.

ACKNOWLEDGMENTS

We thank Fedor Šimkovic for useful discussions and for sharing his diagram topology data. Computations were performed on the PARADOX supercomputing facility (Scientific Computing Laboratory, Center for the Study of Complex Systems, Institute of Physics Belgrade) and the ALPHA cluster (Collège de France). This work was also granted access to the HPC resources of TGCC and IDRIS under the Allocations No. A0090510609 and No. A0070510609 attributed by GENCI (Grand Equipement National de Calcul Intensif). J.V. acknowledges funding provided by the Institute of Physics Belgrade, through the grant by the Ministry of Education, Science, and Technological Development of the Republic of Serbia, as well as by the Science Fund of the Republic of Serbia, under the Key2SM project (PROMIS program, Grant No. 6066160). M.F. acknowledges support by the Simons Foundation. We also acknowledge support by the European Research Council for the European Union Seventh Framework Program (FP7/2007-2013) with ERC Grant No. 319286 (QMAC).

APPENDIX A: REAL-TIME INTEGRATION

Let us consider the following special case of the integral given by Eq. (3), which is relevant for real-time integrations featuring integrands of the form e^{itE}:

$$\tilde{\mathcal{I}}_{\{l_2\dots l_N\},\{E_2\dots E_N\}}(t) = \prod_{i=2}^{N} \int_0^{t_{i+1}} dt_i\, t_i^{l_i}\, e^{it_i E_i}, \tag{A1}$$

with $t_{N+1} \equiv t$. This corresponds to the case $r \notin [2, N]$ in Eq. (3), and $\omega_i = iE_i$, and we will define $\tilde{E}_i$ analogously to $\tilde{\omega}_i$. The result is then obtained straightforwardly from Eq. (14),

$$\begin{aligned}\tilde{\mathcal{I}}_{\{l_2\dots l_N\},\{E_2\dots E_N\}}(t) = &\sum_{\{b_i\in[\delta_{\tilde{z}_i},1]\}_{i=2\dots N}} e^{it\tilde{E}_N b_N} \sum_{\{k_i\in[0,(1-\delta_{\tilde{z}_i})n_i]\}_{i:b_i=1}} \\ &\times (-1)^{\sum_{i=2}^N k_i} \prod_{i:\delta_{\tilde{z}_i}=1} \frac{1}{n_i} \\ &\times t^{n_N+1-b_N-k_N} \prod_{i:\tilde{E}_i\neq 0} \frac{C_{n_i,k_i}}{(i\tilde{E}_i)^{k_i+b_i}},\end{aligned} \tag{A2}$$

which has the following general form:

$$\tilde{\mathcal{I}}(t) = \sum_{j;p\in\mathbb{N}_0} \mathcal{Z}_{p,j} t^p e^{it\mathcal{E}_j}. \tag{A3}$$

APPENDIX B: EXTRACTING REAL-AXIS RESULTS WITHOUT POLE BROADENING

In this section, we show how the results on the real axis can be extracted without any numerical broadening of the poles. Rather, we make use of the pole amplitudes by interpreting the result as being representative of the thermodynamic limit, where poles on the real axis merge into a branch cut, and thus we consider that the pole amplitude is a continuous function of the real frequency. We extract the imaginary part of the contribution [Im$D(\omega)$], and then the Hilbert transform can be used to reconstruct the real part.

The procedure relies on the following construction: A function $f(z)$ which is analytic everywhere in the upper half of the complex plane ($z^+ = x + iy$ with $y > 0$) and decays to zero with $|z^+|$ satisfies the relation

$$f(z^+) = -\frac{1}{\pi} \int dx' \frac{\mathrm{Im} f(x' + i0^+)}{z^+ - x'}. \tag{B1}$$

After applying the pth derivative with respect to x (i.e., the real part of z^+) on both sides of the equation, one obtains

$$\begin{aligned}\partial_x^p f(z^+) &= -\frac{1}{\pi} \int dx' \partial_x^p \frac{\mathrm{Im} f(x' + i0^+)}{z^+ - x'} \\ &= -\frac{1}{\pi} \int dx' (-1)^p (p+1)! \frac{\mathrm{Im} f(x' + i0^+)}{(z^+ - x')^{p+1}}.\end{aligned} \tag{B2}$$

We can now move the constant prefactors to the left-hand side and rename $p + 1 \to p$. Just above the real axis, we have

$$\frac{(-1)^p\pi}{p!} \partial_x^{p-1} f(x + i0^+) = \int dx' \frac{\mathrm{Im} f(x' + i0^+)}{(x - x' + i0^+)^p}. \tag{B3}$$

We can now discretize the expression on a uniform x grid with the step Δx, say, $x_j = j\Delta x$, and we see that the right-hand side has the form of a sum of poles of order p, equidistant along the real axis, and with amplitudes $\mathcal{A}_j = \mathrm{Im} f(x_j + i0^+)$,

$$\frac{(-1)^p\pi}{p!} \tilde{\partial}_j^{p-1} \mathcal{A}_j \approx \mathrm{Im} \sum_{j'} \Delta x \frac{\mathcal{A}_{j'}}{(x_j - x_{j'} + i0^+)^p}, \tag{B4}$$

where $\tilde{\partial}$ is the finite-difference approximation for the derivative along the x axis. Clearly, the imaginary part of the entire sum of p-order poles at a certain point x_j can be estimated by looking only at the $(p-1)$th derivative of the amplitudes of these poles at x_j, as given in the above expression.

The expression (B4) can be readily applied in our case [Eq. (33)] where the real axis is the frequency axis ω, with step $\Delta\omega$ and $\omega_j = j\Delta\omega$, and the sum of the poles determines our diagram contribution D. In general we have poles of various orders, but we can group the poles by order and treat their contributions separately. We therefore have

$$\mathrm{Im}D(\omega_j + i0^+) \approx \frac{\pi}{\Delta\omega} \sum_p \frac{(-1)^p}{p!} \tilde{\partial}_j^{p-1} \mathcal{A}_{j,p}. \tag{B5}$$

In the case of simple poles only, the contribution at any ω_j is simply proportional to the amplitude of the pole $\mathcal{A}_{j,1}$. Other-

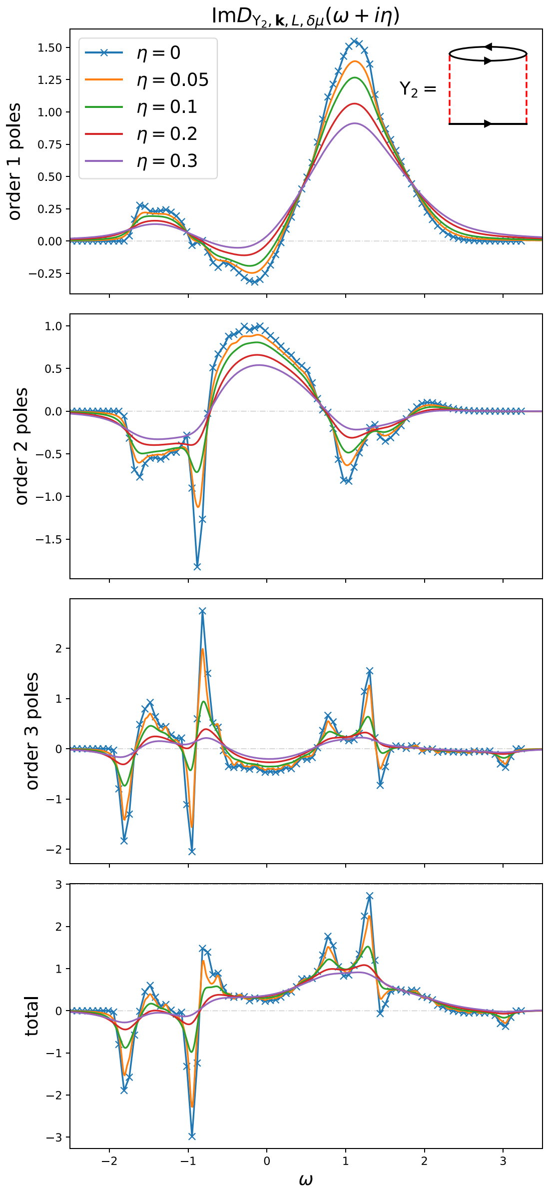


FIG. 10. Illustration of a $\eta = 0^+$ result obtained from Eq. (31) without any numerical broadening, based only on pole amplitudes. The diagram used is the second-order diagram (illustrated in the top panel), with $L = 2$. In the propagators, we take $\delta\mu = 0$. The rest of the parameters are $\mu - U\langle n_\sigma\rangle = -0.1D$, $T = 0.1D$, and the external momentum is $\mathbf{k} = (0, \pi)$. The top three panels are contributions from first-, second-, and third-order poles, respectively. The bottom panel is the total result. Lines with $\eta > 0$ are obtained with numerical broadening. The crosses on the $\eta = 0$ result denote the available frequencies (in between, we assume linear interpolation).

wise, the procedure requires that the pole amplitudes form a reasonably smooth function of the real frequency. Additionally, the energy resolution is a measure of the systematic error made in this procedure.

To avoid statistical noise and noisy features coming from the finite size of the lattice (see next section), we test our method on the example of a $N = 2$, $L = 2$ diagram, which we can solve with the full summation of Eq. (31), on a lattice

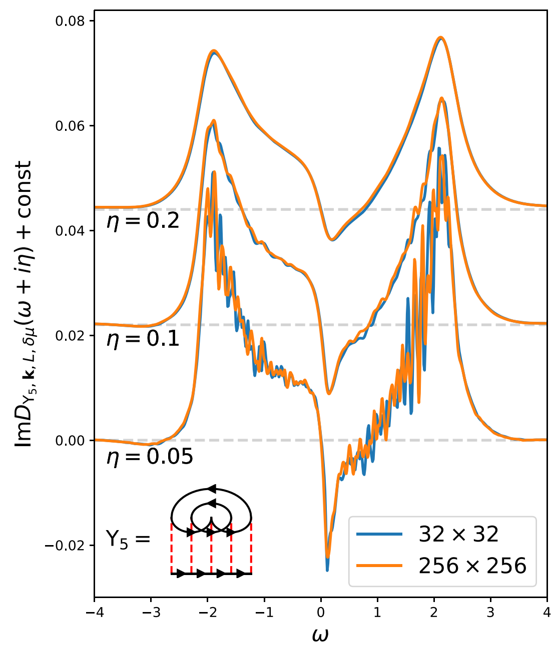


FIG. 11. Comparison of the real-frequency imaginary self-energy result for a single fifth-order diagram (illustrated in the bottom-left corner), for the lattice sizes 32×32 and 256×256, at three different levels of broadening. The calculation is in both cases performed with the same number of MC steps and took similar time. The parameters are $L = 0$, $\delta\mu = 0$, $\mu - U\langle n_\sigma\rangle = -0.1D$, $T = 0.1D$, and the external momentum is $\mathbf{k} = (0, \pi)$.

of the size 96×96. This diagram produces poles up to order 3. The result is shown in Fig. 10. In the first three panels, we show the contribution from the poles of each order, and in the bottom panel, we show the total result.

APPENDIX C: CONVERGENCE WITH LATTICE SIZE

In this section we discuss the convergence of the result with respect to the lattice size. In Fig. 11, we compare the results for a single $N = 5$, $L = 0$ diagram on the lattices of size 32×32 and 256×256. We observe that the result is almost exactly the same at broadening level $\eta = 0.2$, which brings further confidence in the results in the main part of the paper.

In Fig. 12, we illustrate how the size of the lattice determines the highest energy resolution that one can have, under requirement that the results form a continuous curve on the real axis and are, therefore, representative of the thermodynamic limit. We perform the full summation for the second-order diagram with $L = 0$, with various sizes of the lattice and various resolutions. Clearly, the bigger the lattice, the higher the energy resolution one can set without affecting the smoothness of the results.

The numerical parameters of the calculation are therefore the size of the lattice, the energy resolution, and the broadening (the resolution and the broadening can be tuned *a posteriori*), and one can tune them to get the optimal ratio between performance and the error bar. If the pole amplitudes $\mathcal{A}_{jp}$ are a relatively smooth function of j, no broadening is then needed at all.

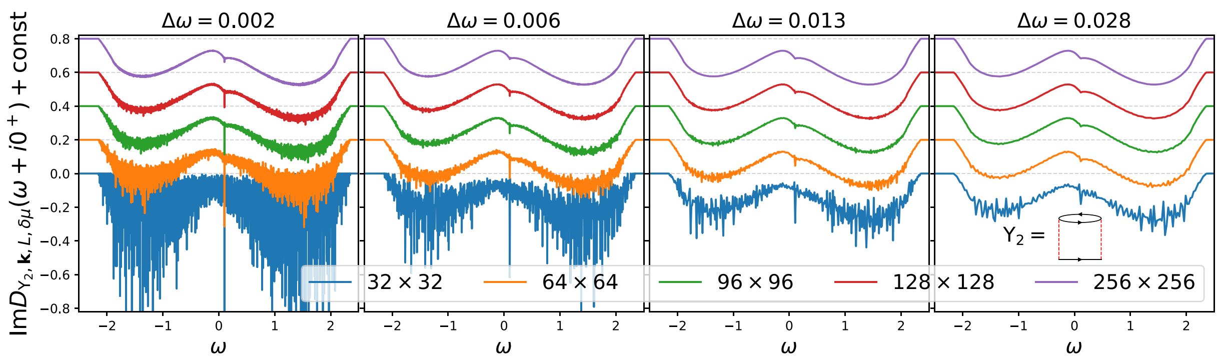


FIG. 12. Real-frequency result ($\eta = 0^+$) for the contribution of the lowest-order diagram (illustrated in the rightmost panel) at various lattice sizes and frequency resolutions, obtained with full summation (gray code). The step of the uniform energy grid is denoted $\Delta\omega$. The parameters are $L = 0$, $\delta\mu = 0$, $\mu - U\langle n_\sigma\rangle = -0.1D$, $T = 0.1D$, and the external momentum is $\mathbf{k} = (0, \pi)$.

APPENDIX D: DERIVATION OF EQ. (5)

After applying n times the partial integration over the integral from the left-hand side of Eq. (5), we get

$$\begin{aligned}\int_0^{\tau_f} \tau^n e^{\tau z} d\tau &= \frac{1}{z^{n+1}} \int_0^{z\tau_f} \tau^n e^\tau d\tau \\ &= \frac{1}{z^{n+1}}\left[e^{z\tau_f}(z\tau_f)^n - ne^{z\tau_f}(z\tau_f)^{n-1} + \cdots + (-1)^n n! \int_0^{z\tau_f} \tau^0 e^\tau d\tau\right] \\ &= \frac{1}{z^{n+1}}\left[\frac{n!}{(n-0)!}(-1)^0 e^{z\tau_f}(z\tau_f)^{n-0} + (-1)^1\frac{n!}{(n-1)!}e^{z\tau_f}(z\tau_f)^{n-1} + \cdots + (-1)^n\frac{n!}{(n-n)!}\int_0^{z\tau_f}\tau^0 e^\tau d\tau\right] \\ &= \frac{1}{z^{n+1}}\left[\frac{n!}{(n-0)!}(-1)^0 e^{z\tau_f}(z\tau_f)^{n-0} + (-1)^1\frac{n!}{(n-1)!}e^{z\tau_f}(z\tau_f)^{n-1} + \cdots + (-1)^n\frac{n!}{(n-n)!}(z\tau_f)^0(e^{z\tau_f}-1)\right] \\ &= \frac{1}{z^{n+1}} e^{z\tau_f}\sum_{k=0}^{n}(-1)^k (z\tau_f)^{n-k}\frac{n!}{(n-k)!} - (-1)^n\frac{n!}{z^{n+1}},\end{aligned} \quad \text{(D1)}$$

which can be readily identified with the right-hand side of Eq. (5).

APPENDIX E: DERIVATION OF EQ. (21)

We are looking for a solution of the Fourier transform

$$G_0^l(\varepsilon,\tau) = \frac{1}{\beta}\sum_{i\Omega_\eta}\frac{e^{-i\Omega_\eta\tau}}{(i\Omega_\eta - \varepsilon)^l}. \quad \text{(E1)}$$

For any τ, we can express the sum above as a contour integral, and we find

$$\begin{aligned}G_0^l(\varepsilon,\tau) &= -\mathrm{Res}_{z=\varepsilon}\frac{e^{-z\tau}}{(z-\varepsilon)^l}\frac{\eta^{\lfloor\frac{\tau}{\beta}\rfloor}e^{\lfloor\frac{\tau}{\beta}\rfloor\beta z}}{1-\eta e^{-\beta z}}dz \\ &= -\frac{\eta^{\lfloor\frac{\tau}{\beta}\rfloor}}{(l-1)!}\frac{d^{l-1}}{dz^{l-1}}\frac{e^{-\beta z\{\frac{\tau}{\beta}\}}}{1-\eta e^{-\beta z}}\bigg|_{z=\varepsilon},\end{aligned} \quad \text{(E2)}$$

where $\lfloor\ldots\rfloor$ denotes the integer part (floor function), and $\{x\} \equiv x - \lfloor x\rfloor$ denotes the fractional part.

We see that it will be useful to have an expression for derivatives of $(1-\eta e^z)^{-1}$. They have the general form

$$\frac{d^k}{dz^k}\frac{1}{1-\eta e^z} = \sum_{n=0}^{k} C_n^k \frac{(e^z)^n}{(1-\eta e^z)^{n+1}}. \quad \text{(E3)}$$

By deriving this expression on both sides, one obtains a recursion for the coefficients C_n^k,

$$C_n^{k+1} = nC_n^k + \eta n C_{n-1}^k, \quad \text{(E4)}$$

with holds for $k > -1$ and $n > 0$ with $C_0^0 = 1$. That can be rewritten

$$\frac{\eta^n}{n!}C_n^{k+1} = n\frac{\eta^n}{n!}C_n^k + \frac{\eta^{n-1}}{(n-1)!}C_{n-1}^k. \quad \text{(E5)}$$

If we define $S_n^k = \frac{\eta^n}{n!}C_n^k$, we have the recursion $S_n^{k+1} = nS_n^k + S_{n-1}^k$, which is the recursion for the Stirling numbers of the second kind. This allows one to have the following important result:

$$\begin{aligned}\frac{d^k}{dz^k}\frac{1}{1-\eta e^z} &= \sum_{n=0}^{k}\eta^n n!\left\{ {k \atop n} \right\}\frac{(e^z)^n}{(1-\eta e^z)^{n+1}} \\ &= \sum_{n=0}^{k}\eta^n n!\left\{ {k \atop n} \right\}\frac{e^{-z}}{(e^{-z}-\eta)^{n+1}}.\end{aligned} \quad \text{(E6)}$$

With this, one obtains the following expression:

$$G_0^l(\varepsilon,\tau) = -e^{\varepsilon\beta(1-\{\frac{\tau}{\beta}\})}\eta^{\lfloor\frac{\tau}{\beta}\rfloor+1}(-\beta)^{l-1} \times \sum_{m=0}^{l-1}\sum_{n=0}^{l-m-1}\frac{n!}{(l-m-1)!m!}\left\{\begin{matrix} l-m-1 \\ n\end{matrix}\right\} \times \left(\frac{1}{\eta e^{\varepsilon\beta}-1}\right)^{n+1}\left\{\frac{\tau}{\beta}\right\}^m, \quad \text{(E7)}$$

which already satisfies the (anti)periodicity properties of the Green's function.

To make use of the result given by Eq. (E7), we need to express $G_0^l(\varepsilon,\tau)$ as a function of two times $G_0^l(\varepsilon,\tau,\tau') \equiv G_0^l(\varepsilon,\tau-\tau')$, with $\tau,\tau' \in [0,\beta]$. We first consider $\tau \geqslant \tau'$. By substituting $(\tau-\tau')^m = \sum_{\zeta=0}^m(-1)^{m-\zeta}\binom{m}{\zeta}\tau^\zeta\tau'^{m-\zeta}$ into Eq. (E7) and substituting $m-\zeta$ with ς, we get

$$G_0^l(\varepsilon,\tau-\tau') = \eta e^{\varepsilon(\tau'-\tau)}n_\eta(-\varepsilon)\sum_{\zeta=0}^{l-1}\sum_{\varsigma=0}^{l-\zeta-1}c^-_{l,\zeta,\varsigma}(\varepsilon)\tau^\zeta\tau'^\varsigma, \quad \text{(E8)}$$

with $c^-_{l,\zeta,\varsigma}(\varepsilon)$ as defined in Eq. (22). The result for $\tau < \tau'$ can then be easily obtained by proving the property $G_0^l(\varepsilon,\tau) = (-1)^l G_0^l(-\varepsilon,-\tau)$,

$$\begin{aligned} G_0^l(\varepsilon,-\tau) &= \frac{1}{\beta}\sum_{n=-\infty}^{\infty}\frac{e^{i\Omega_\eta\tau}}{(i\Omega_\eta-\varepsilon)^l} \\ &= \frac{1}{\beta}\sum_{n=-\infty}^{\infty}\frac{e^{-i\Omega_\eta\tau}}{(-i\Omega_\eta-\varepsilon)^l} \\ &= (-1)^l\frac{1}{\beta}\sum_{n=-\infty}^{\infty}\frac{e^{-i\Omega_\eta\tau}}{(i\Omega_\eta+\varepsilon)^l} \\ &= (-1)^l G_0^l(-\varepsilon,\tau), \end{aligned}$$

which implies that in the definition (21), we must have

$$c^+_{l,\zeta,\varsigma}(\varepsilon) = (-1)^{l-1}c^-_{l,\varsigma,\zeta}(-\varepsilon). \quad \text{(E9)}$$

APPENDIX F: GENERAL HAMILTONIAN CASE

It is important to show that our method is not restricted to a specific choice of Hamiltonian. The local density-density interaction and the single band of the Hubbard Hamiltonian bring many simplifications, but none of them are necessary for our imaginary-time integral solution or the chemical-potential tuning scheme.

Consider the general Hamiltonian

$$H = \sum_\alpha(\varepsilon_\alpha-\mu) + \sum_{\alpha_1\alpha_2\alpha_3\alpha_4}U_{\alpha_1\alpha_2\alpha_3\alpha_4}c^\dagger_{\alpha_1}c_{\alpha_2}c^\dagger_{\alpha_3}c_{\alpha_4}. \quad \text{(F1)}$$

The α are the eigenstates of the noninteracting Hamiltonian, e.g., a combined momentum, band, and spin index. The self-energy can be now expressed as a series,

$$\Sigma^{(\mathrm{HF})}_{\alpha,\alpha'}(\tau) = \sum_N\sum_{\Upsilon_N}\prod_{j=1}^{2N-1}\sum_{l_j=1}^{\infty}\sum_{\alpha_{j,1}\ldots\alpha_{j,l_j}}\prod_{n=1}^{l_j-1}\sum_{\mathbf{V}_{j,n}} \times [\mathbf{V}_{j,n}]_{\alpha_{j,n}\alpha_{j,n+1}}\prod_{i=1}^{N}U_{\alpha_{j_1}(i)\alpha_{j_2}(i)\alpha_{j_3}(i)\alpha_{j_4}(i)} \times \prod_{m=1}^{N-1+\sum_j(l_j-1)}\int_0^\beta d\tau_m\, G_0\big(\bar{\varepsilon}_{\alpha_{j,n}},\tilde{\tau}_{j,n}-\tilde{\tau}'_{j,n}\big). \quad \text{(F2)}$$

Similarly as before, Υ_N enumerates topologies without any instantaneous insertions (Hartree or chemical potential) at perturbation order N (the number of interaction vertices). The fermionic lines in the Υ_N topology are enumerated with j. On each fermionic line, we make any number l_j-1 of instantaneous insertions with amplitudes $\mathbf{V}_{j,n}$ (interaction amplitudes in Hartree insertions are included in $\mathbf{V}$; n enumerates the insertions at the fermionic line j). In general, Hartree insertions may contain off-diagonal terms in the α basis and are therefore a matrix in the α space. However, it is necessary that chemical-potential shifts are diagonal in this basis, as we want to have the bare propagator diagonal in this basis as well. Otherwise, the form of G_0 from Eq. (18) would no longer hold. Nevertheless, one may still have a separate chemical-potential shift for each state, $\delta\mu_\alpha$. After making insertions, the number of fermionic lines increases to $\sum_j l_j$. The fermionic lines are now enumerated with j,n, and the corresponding states are $\alpha_{j,n}$. The index i enumerates the interaction vertices outside of any Hartree insertions. We denote $\alpha_{j_{1\ldots4}}(i)$ as the single-particle states at four terminals of each interaction vertex. The interaction vertices at incoming $(i=1)$ and outgoing $(i=N)$ terminals of the self-energy diagram are $\alpha_{j_1}(i=N)=\alpha$, $\alpha_{j_2}(i=1)=\alpha'$. With m, we enumerate all times to be integrated over. With each interaction vertex $i>1$, we associate one time, and there is a time associated to each instantaneous insertion of which there are $\sum_j(l_j-1)$. We assume that the incoming time corresponding to the vertex $i=1$ is 0. The times on the terminals of each bare propagator j,n are $\tilde{\tau}_{j,n}$ and $\tilde{\tau}'_{j,n}$ and they take on values from the set $\{\tau_m\}_{m=0\ldots N-1+\sum_j(l_j-1)}$, with the external incoming time fixed, $\tau_0\equiv 0$. $\tilde{\tau}_{j,n}$, $\tilde{\tau}'_{j,n}$, and $\alpha_{j_{1\ldots4}}(i)$ are implicit functions of topology Υ_N. Finally, $\bar{\varepsilon}_{\alpha_{j,n}} \equiv \varepsilon_{\alpha_{j,n}}-\mu+\delta\mu_{\alpha_{j,n}}$. We can now focus only on the time-integral part and proceed completely analogously to Eqs. (27)–(31).

It is worth noting that with general interactions, pulling the coupling constant in front of the diagram contribution is impossible, as the frequency dependence of the contribution of each diagram will depend on the precise form of $U_{\alpha_1\alpha_2\alpha_3\alpha_4}$. In the most general case, one must set specific values for $U_{\alpha_1\alpha_2\alpha_3\alpha_4}$ and $\delta\mu_\alpha$ before performing the Monte Carlo summation. One can then choose the variables that will be sampled stochastically and the ones that will be fully summed over. In the end, the contributions can be easily grouped by total number of independent times (K), including those in Hartree insertions. The integration of times in Hartree insertions can always be performed beforehand. Therefore, in the fully general case, the number of integrations to be performed at the time of Monte Carlo sampling is $N - 1 + \sum_j (l_j - 1)$. In the case of purely density-density interactions (as is the case in the Hubbard model) or spin-spin interactions in the absence of external magnetic fields, this simplifies further because instantaneous insertions lead to expressions of the type $\frac{1}{(i\omega-\varepsilon)^l}$ for which we can work out the temporal Fourier transform analytically [Eq. (21)] and the remaining number of integrations to perform is $N - 1$ [as we do in Eq. (31)]. In the general case, when Hartree insertions are not diagonal in the α basis, one has expressions of the type $\frac{1}{i\omega-\varepsilon_1}\frac{1}{i\omega-\varepsilon_2}\cdots\frac{1}{i\omega-\varepsilon_l}$. In principle, one could prepare the analytical Fourier transforms for a general function of this form, but it might be increasingly involved at large l, so we assume one would do these integrations at the level of the Monte Carlo, when $\varepsilon_{1\ldots l}$ are already specified.

We finally emphasize that even more general constructions are possible, even in bases other than the noninteracting eigenbasis. In such cases, the G_0's are nondiagonal and may have a continuous real-frequency dependence, instead of being a single pole. We leave such considerations for future work.

[1] The Hubbard model at half a century, Nat. Phys. **9**, 523 (2013).
[2] J. P. F. LeBlanc, A. E. Antipov, F. Becca, I. W. Bulik, Garnet Kin-Lic Chan, Chia-Min Chung, Y. Deng, M. Ferrero, T. M. Henderson, C. A. Jiménez-Hoyos, E. Kozik, Xuan-Wen Liu, A. J. Millis, N. V. Prokof'ev, M. Qin, G. E. Scuseria, Hao Shi, B. V. Svistunov, L. F. Tocchio, I. S. Tupitsyn, S. R. White, S. Zhang, Bo-Xiao Zheng, Z. Zhu, and E. Gull (Simons Collaboration on the Many-Electron Problem), Solutions of the Two-Dimensional Hubbard Model: Benchmarks and Results from a Wide Range of Numerical Algorithms, Phys. Rev. X **5**, 041041 (2015).
[3] G. Rohringer, H. Hafermann, A. Toschi, A. A. Katanin, A. E. Antipov, M. I. Katsnelson, A. I. Lichtenstein, A. N. Rubtsov, and K. Held, Diagrammatic routes to nonlocal correlations beyond dynamical mean field theory, Rev. Mod. Phys. **90**, 025003 (2018).
[4] T. Schäfer, N. Wentzell, F. Šimkovic, Y.-Y. He, C. Hille, M. Klett, C. J. Eckhardt, B. Arzhang, V. Harkov, F.-M. Le Régent, A. Kirsch, Y. Wang, A. J. Kim, E. Kozik, E. A. Stepanov, A. Kauch, S. Andergassen, P. Hansmann, D. Rohe, Y. M. Vilk, J. P. F. LeBlanc, S. Zhang, A.-M. S. Tremblay, M. Ferrero, O. Parcollet, and A. Georges, Tracking the Footprints of Spin Fluctuations: A MultiMethod, MultiMessenger Study of the Two-Dimensional Hubbard Model, Phys. Rev. X **11**, 011058 (2021).
[5] A. W. Sandvik, Stochastic method for analytic continuation of quantum Monte Carlo data, Phys. Rev. B **57**, 10287 (1998).
[6] O. F. Syljuåsen, Using the average spectrum method to extract dynamics from quantum Monte Carlo simulations, Phys. Rev. B **78**, 174429 (2008).
[7] S. Fuchs, T. Pruschke, and M. Jarrell, Analytic continuation of quantum Monte Carlo data by stochastic analytical inference, Phys. Rev. E **81**, 056701 (2010).
[8] K. Ghanem and E. Koch, Extending the average spectrum method: Grid point sampling and density averaging, Phys. Rev. B **102**, 035114 (2020).
[9] H. Aoki, N. Tsuji, M. Eckstein, M. Kollar, T. Oka, and P. Werner, Nonequilibrium dynamical mean-field theory and its applications, Rev. Mod. Phys. **86**, 779 (2014).
[10] J. K. Freericks, Quenching Bloch oscillations in a strongly correlated material: Nonequilibrium dynamical mean-field theory, Phys. Rev. B **77**, 075109 (2008).
[11] M. Eckstein and M. Kollar, Nonthermal Steady States after an Interaction Quench in the Falicov-Kimball Model, Phys. Rev. Lett. **100**, 120404 (2008).
[12] M. Schiró and M. Fabrizio, Real-time diagrammatic Monte Carlo for nonequilibrium quantum transport, Phys. Rev. B **79**, 153302 (2009).
[13] P. Werner, T. Oka, and A. J. Millis, Diagrammatic Monte Carlo simulation of nonequilibrium systems, Phys. Rev. B **79**, 035320 (2009).
[14] M. Eckstein, M. Kollar, and P. Werner, Thermalization after an Interaction Quench in the Hubbard Model, Phys. Rev. Lett. **103**, 056403 (2009).
[15] M. Schiró and M. Fabrizio, Time-Dependent Mean Field Theory for Quench Dynamics in Correlated Electron Systems, Phys. Rev. Lett. **105**, 076401 (2010).
[16] P. Werner, T. Oka, M. Eckstein, and A. J. Millis, Weak-coupling quantum Monte Carlo calculations on the Keldysh contour: Theory and application to the current-voltage characteristics of the Anderson model, Phys. Rev. B **81**, 035108 (2010).
[17] M. Eckstein and P. Werner, Nonequilibrium dynamical mean-field calculations based on the noncrossing approximation and its generalizations, Phys. Rev. B **82**, 115115 (2010).
[18] M. Schiró and M. Fabrizio, Quantum quenches in the Hubbard model: Time-dependent mean-field theory and the role of quantum fluctuations, Phys. Rev. B **83**, 165105 (2011).
[19] P. Werner and M. Eckstein, Phonon-enhanced relaxation and excitation in the Holstein-Hubbard model, Phys. Rev. B **88**, 165108 (2013).
[20] M. Eckstein and P. Werner, Nonequilibrium dynamical mean-field simulation of inhomogeneous systems, Phys. Rev. B **88**, 075135 (2013).
[21] G. Cohen, D. R. Reichman, A. J. Millis, and E. Gull, Green's functions from real-time bold-line Monte Carlo, Phys. Rev. B **89**, 115139 (2014).
[22] G. Cohen, E. Gull, D. R. Reichman, and A. J. Millis, Green's Functions from Real-Time Bold-Line Monte Carlo Calculations: Spectral Properties of the Nonequilibrium Anderson Impurity Model, Phys. Rev. Lett. **112**, 146802 (2014).
[23] G. Cohen, E. Gull, D. R. Reichman, and A. J. Millis, Taming the Dynamical Sign Problem in Real-Time Evolution of Quantum Many-Body Problems, Phys. Rev. Lett. **115**, 266802 (2015).
[24] N. V. Prokof'ev and B. V. Svistunov, Polaron Problem by Diagrammatic Quantum Monte Carlo, Phys. Rev. Lett. **81**, 2514 (1998).
[25] A. S. Mishchenko, N. V. Prokof'ev, A. Sakamoto, and B. V. Svistunov, Diagrammatic quantum Monte Carlo study of the Fröhlich polaron, Phys. Rev. B **62**, 6317 (2000).

[26] N. Prokof'ev and B. Svistunov, Bold Diagrammatic Monte Carlo Technique: When the Sign Problem Is Welcome, Phys. Rev. Lett. **99**, 250201 (2007).
[27] N. V. Prokof'ev and B. V. Svistunov, Bold diagrammatic Monte Carlo: A generic sign-problem tolerant technique for polaron models and possibly interacting many-body problems, Phys. Rev. B **77**, 125101 (2008).
[28] E. Gull, D. R. Reichman, and A. J. Millis, Bold-line diagrammatic Monte Carlo method: General formulation and application to expansion around the noncrossing approximation, Phys. Rev. B **82**, 075109 (2010).
[29] E. Kozik, K. Van Houcke, E. Gull, L. Pollet, N. Prokof'ev, B. Svistunov, and M. Troyer, Diagrammatic Monte Carlo for correlated fermions, Europhys. Lett. **90**, 10004 (2010).
[30] K. V. Houcke, E. Kozik, N. Prokof'ev, and B. Svistunov, Diagrammatic Monte Carlo, Phys. Proc. **6**, 95 (2010).
[31] L. Pollet, Recent developments in quantum Monte Carlo simulations with applications for cold gases, Rep. Prog. Phys. **75**, 094501 (2012).
[32] K. Van Houcke, F. Werner, E. Kozik, N. Prokof'ev, B. Svistunov, M. J. H. Ku, A. T. Sommer, L. W. Cheuk, A. Schirotzek, and M. W. Zwierlein, Feynman diagrams versus Fermi-gas Feynman emulator, Nat. Phys. **8**, 366 (2012).
[33] S. A. Kulagin, N. Prokof'ev, O. A. Starykh, B. Svistunov, and C. N. Varney, Bold Diagrammatic Monte Carlo Method Applied to Fermionized Frustrated Spins, Phys. Rev. Lett. **110**, 070601 (2013).
[34] S. A. Kulagin, N. Prokof'ev, O. A. Starykh, B. Svistunov, and C. N. Varney, Bold diagrammatic Monte Carlo technique for frustrated spin systems, Phys. Rev. B **87**, 024407 (2013).
[35] R. Rossi, T. Ohgoe, E. Kozik, N. Prokof'ev, B. Svistunov, K. Van Houcke, and F. Werner, Contact and Momentum Distribution of the Unitary Fermi Gas, Phys. Rev. Lett. **121**, 130406 (2018).
[36] R. Rossi, Determinant Diagrammatic Monte Carlo Algorithm in the Thermodynamic Limit, Phys. Rev. Lett. **119**, 045701 (2017).
[37] K. Van Houcke, F. Werner, T. Ohgoe, N. V. Prokof'ev, and B. V. Svistunov, Diagrammatic Monte Carlo algorithm for the resonant Fermi gas, Phys. Rev. B **99**, 035140 (2019).
[38] K. Chen and K. Haule, A combined variational and diagrammatic quantum Monte Carlo approach to the many-electron problem, Nat. Commun. **10**, 3725 (2019).
[39] A. N. Rubtsov and A. I. Lichtenstein, Continuous-time quantum Monte Carlo method for fermions: Beyond auxiliary field framework, J. Expt. Theor. Phys. Lett. **80**, 61 (2004).
[40] A. N. Rubtsov, V. V. Savkin, and A. I. Lichtenstein, Continuous-time quantum Monte Carlo method for fermions, Phys. Rev. B **72**, 035122 (2005).
[41] E. Gull, P. Werner, O. Parcollet, and M. Troyer, Continuous-time auxiliary-field Monte Carlo for quantum impurity models, Europhys. Lett. **82**, 57003 (2008).
[42] E. Gull, A. J. Millis, A. I. Lichtenstein, A. N. Rubtsov, M. Troyer, and P. Werner, Continuous-time Monte Carlo methods for quantum impurity models, Rev. Mod. Phys. **83**, 349 (2011).
[43] R. E. V. Profumo, C. Groth, L. Messio, O. Parcollet, and X. Waintal, Quantum Monte Carlo for correlated out-of-equilibrium nanoelectronic devices, Phys. Rev. B **91**, 245154 (2015).
[44] A. Moutenet, P. Seth, M. Ferrero, and O. Parcollet, Cancellation of vacuum diagrams and the long-time limit in out-of-equilibrium diagrammatic quantum Monte Carlo, Phys. Rev. B **100**, 085125 (2019).
[45] C. Bertrand, O. Parcollet, A. Maillard, and X. Waintal, Quantum Monte Carlo algorithm for out-of-equilibrium Green's functions at long times, Phys. Rev. B **100**, 125129 (2019).
[46] C. Bertrand, S. Florens, O. Parcollet, and X. Waintal, Reconstructing Nonequilibrium Regimes of Quantum Many-Body Systems from the Analytical Structure of Perturbative Expansions, Phys. Rev. X **9**, 041008 (2019).
[47] A. Taheridehkordi, S. H. Curnoe, and J. P. F. LeBlanc, Algorithmic Matsubara integration for Hubbard-like models, Phys. Rev. B **99**, 035120 (2019).
[48] J. Vučičević and M. Ferrero, Real-frequency diagrammatic Monte Carlo at finite temperature, Phys. Rev. B **101**, 075113 (2020).
[49] A. Taheridehkordi, S. H. Curnoe, and J. P. F. LeBlanc, Optimal grouping of arbitrary diagrammatic expansions via analytic pole structure, Phys. Rev. B **101**, 125109 (2020).
[50] A. Taheridehkordi, S. H. Curnoe, and J. P. F. LeBlanc, Algorithmic approach to diagrammatic expansions for real-frequency evaluation of susceptibility functions, Phys. Rev. B **102**, 045115 (2020).
[51] M. Maček, P. T. Dumitrescu, C. Bertrand, B. Triggs, O. Parcollet, and X. Waintal, Quantum Quasi-Monte Carlo Technique for Many-Body Perturbative Expansions, Phys. Rev. Lett. **125**, 047702 (2020).
[52] W. Wu, M. Ferrero, A. Georges, and E. Kozik, Controlling Feynman diagrammatic expansions: Physical nature of the pseudogap in the two-dimensional Hubbard model, Phys. Rev. B **96**, 041105(R) (2017).
[53] R. Rossi, F. Šimkovic IV, and M. Ferrero, Renormalized perturbation theory at large expansion orders, Europhys. Lett. **132**, 11001 (2020).
[54] F. Šimkovic IV, R. Rossi, and M. Ferrero, Efficient one-loop-renormalized vertex expansions with connected determinant diagrammatic Monte Carlo, Phys. Rev. B **102**, 195122 (2020).
[55] F. Šimkovic IV, J. P. F. LeBlanc, A. J. Kim, Y. Deng, N. V. Prokof'ev, B. V. Svistunov, and E. Kozik, Extended Crossover from a Fermi Liquid to a Quasiantiferromagnet in the Half-Filled 2D Hubbard Model, Phys. Rev. Lett. **124**, 017003 (2020).
[56] A. J. Kim, F. Šimkovic IV, and E. Kozik, Spin and Charge Correlations across the Metal-to-Insulator Crossover in the Half-Filled 2D Hubbard Model, Phys. Rev. Lett. **124**, 117602 (2020).
[57] C. Lenihan, A. J. Kim, F. Šimkovic IV., and E. Kozik, Entropy in the Non-Fermi-Liquid Regime of the Doped 2D Hubbard Model, Phys. Rev. Lett. **126**, 105701 (2021).
[58] T. Ayral, J. Vučičević, and O. Parcollet, Fierz Convergence Criterion: A Controlled Approach to Strongly Interacting Systems with Small Embedded Clusters, Phys. Rev. Lett. **119**, 166401 (2017).
[59] A. Moutenet, W. Wu, and M. Ferrero, Determinant Monte Carlo algorithms for dynamical quantities in fermionic systems, Phys. Rev. B **97**, 085117 (2018).
[60] F. Šimkovic IV and E. Kozik, Determinant Monte Carlo for irreducible Feynman diagrams in the strongly correlated regime, Phys. Rev. B **100**, 121102(R) (2019).

[61] E. Eidelstein, E. Gull, and G. Cohen, Multiorbital Quantum Impurity Solver for General Interactions and Hybridizations, Phys. Rev. Lett. **124**, 206405 (2020).

[62] A. A. Kunitsa and So Hirata, Grid-based diffusion Monte Carlo for fermions without the fixed-node approximation, Phys. Rev. E **101**, 013311 (2020).

[63] M. Hutcheon, Stochastic nodal surfaces in quantum Monte Carlo calculations, Phys. Rev. E **102**, 042105 (2020).

[64] J. Vučičević, J. Kokalj, R. Žitko, N. Wentzell, D. Tanasković, and J. Mravlje, Conductivity in the Square Lattice Hubbard Model at High Temperatures: Importance of Vertex Corrections, Phys. Rev. Lett. **123**, 036601 (2019).

[65] E. W. Huang, R. Sheppard, B. Moritz, and T. P. Devereaux, Strange metallicity in the doped Hubbard model, Science **366**, 987 (2019).

[66] P. T. Brown, D. Mitra, E. Guardado-Sanchez, R. Nourafkan, A. Reymbaut, Charles-David Hébert, S. Bergeron, A.-M. S. Tremblay, J. Kokalj, D. A. Huse, P. Schauß, and W. S. Bakr, Bad metallic transport in a cold atom Fermi-Hubbard system, Science **363**, 379 (2018).

IOP Publishing Journal of Physics A: Mathematical and Theoretical
J. Phys. A: Math. Theor. **58** (2025) 215201 (41pp) https://doi.org/10.1088/1751-8121/add7a5

Symmetry breaking mechanisms of the 3*BF* action for the Standard Model coupled to gravity

Pavle Stipsić and Marko Vojinović*

Institute of Physics, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia

E-mail: vmarko@ipb.ac.rs and pstipsic@ipb.ac.rs

Received 12 February 2025; revised 21 April 2025
Accepted for publication 12 May 2025
Published 21 May 2025

Abstract
We study the details of the explicit and spontaneous symmetry breaking of the constrained 3*BF* action representing the Standard Model coupled to Einstein–Cartan gravity. First we discuss how each particular constraint breaks the original symmetry of the topological 3*BF* action. Then we investigate the spontaneous symmetry breaking and the Higgs mechanism for the electroweak theory in the constrained 3*BF* form, in order to demonstrate that they can indeed be performed in the framework of higher gauge theory. A formulation of the Proca action as a constrained 3*BF* theory is also studied in detail.

Keywords: quantum gravity, higher gauge theory, 3-group, 3*BF* action, symmetry breaking, Higgs mechanism

1. Introduction

The formulation of a quantum theory of the gravitational field represents one of the main open problems in the modern fundamental theoretical physics. Over the years, there have been many attempts to tackle this problem, and several major approaches have been developed, including string theory [1, 2], loop quantum gravity (LQG) [3, 4], and various other frameworks. Each of these approaches has its own set of advantages and disadvantages. In particular, the covariant version of the LQG approach [5] focuses on the rigorous definition of the gravitational path integral, which is used as a key ingredient in defining the corresponding quantum theory. One of the main advantages of the covariant LQG lies in the fact that such a rigorous definition can in fact be formulated, using the so-called *spinfoam quantization procedure*, and the gravitational field can be quantized successfully. On the other hand, the main disadvantage lies in the fact that the spinfoam quantization procedure works well for the pure gravitational field, but is not compatible with all other fields in nature (collectively called matter fields) [6–8]. In

* Author to whom any correspondence should be addressed.

other words, while it is possible to quantize the gravitational field, it is not straightforward to quantize the gravitational field with matter.

In recent years, progress has been made to circumvent this disadvantage. One of the promising avenues is based on the so-called *higher gauge theory* [9, 10], which represents a framework that generalizes the notion of symmetry using mathematical techniques of higher category theory. Attention is mostly focused on the categorical structures called n-groups, which are a certain type of generalization of the algebraic notion of a group, and are used instead of groups to describe the gauge symmetry of the theory [11] (for various other applications of n-groups to physics see for example [12–24]). In particular, the structure of 3-groups appears to be suitable to give an algebraic description of all relevant fields in nature—the gravitational field, the Yang–Mills field, the scalar field and the Dirac field [25]. On the other hand, the structure of the 3-group lends itself nicely to a generalization of the spinfoam quantization procedure [26], which opens the door to study the quantization of the gravitational field with matter within a unified mathematical description.

One of the central elements in the above construction is a notion of the BF theory and its higher gauge theory generalizations called nBF theories. Historically, one of the natural approaches to these theories relies on the Batalin–Vilkovisky formalism [27–31], and gives rise to the so-called Alexandrov–Kontsevich–Schwarz–Zaboronsky construction, see [32] and further developments in [33–38]. The classical formulation of general relativity and other theories of gravity based on the BF theory have initial results in the work of Plebanski [39], see also [40–42] for a comprehensive review of various models. The 2-group formulation, called $2BF$ or $BFCG$ model, was first introduced in [43, 44] and further studied in [45–50]. The classical $3BF$ and $4BF$ theories were formulated in [25] and [51], respectively. At the quantum level, nBF theories give rise to a class of topological quantum field theories, first introduced in the works of Porter [52], see also [26, 53].

The higher gauge theory programme based on 3-groups has recently given some promising concrete results. First, it was understood how to construct a classical action that describes the full Standard Model (SM) naturally coupled to Einstein–Cartan gravity (our naming convention follows the textbook [54]), so that it has a form compatible with the generalized spinfoam quantization procedure [25]. This amounts to the reformulation of the classical theory into a form of the so-called constrained $3BF$ action. Such an action consists of two main parts—the topological $3BF$ part, specified by the postulated structure of a 3-group, and the constraint part, which deforms the topological theory into a non-topological one, with nontrivial dynamics. Next, the quantization procedure for the topological sector has been sucessfully implemented, leading to a formulation of the path integral corresponding to the topological quantum field theory based on a given 3-group [26]. In addition to these results, the symmetries of the topological $3BF$ theory have been studied in full detail [55, 56], leading to deeper understanding of the various properties of the models. Some important mathematical results have also been established [17, 57, 58]. Nevertheless, the symmetries of the constrained $3BF$ action, which represents a realistic classical theory, have not been studied so far. The main purpose of this paper is to fill this missing step, and study the symmetries of the constrained $3BF$ action.

From the structure of the constrained $3BF$ action, it is straightforward to see that the topological sector has a certain (large) symmetry, while the constraints mainly break this symmetry to one of its subgroups. Therefore, our work focuses on the study of various symmetry breaking mechanisms and the role played by each individual constraint. In particular, we examine precisely how each constraint individually breaks the $3BF$ symmetry group and down to which subgroup. As it turns out, some constraints have bigger influence and break the symmetry down to a smaller subgroup, while other constraints have smaller influence and break the symmetry only slightly. We shall also find out that one part of the symmetry group remains unbroken even in the presence of all constraints. All these results are then organized and presented in a form of a table. After the analysis of the explicit symmetry breaking, our attention turns to the details of the spontaneous symmetry breaking, and the Higgs mechanism. This is very important, since the Higgs mechanism is crucial for the SM, and it is not clear whether the constrained $3BF$ formulation of the SM action admits the implementation of the Higgs mechanism with the same outcome as the ordinary textbook formulation of the theory. It turns out that the Higgs mechanism does indeed yield the expected outcome, but the details of the implementation of spontaneous symmetry breaking are very far from the ordinary textbook approach. Instead, a set of completely new calculational techniques has been developed, including one theorem, and they represent our second main result. These new techniques are necessary, due to the fact that the constrained $3BF$ formulation of the SM action is based on a very different set of variables, compared to the textbook SM Lagrangian. Finally, as one of the major steps in this analysis, we also provide a formulation of the Proca action within the framework of higher gauge theory, and discuss explicitly three different versions of the constrained $3BF$ action for the Proca theory coupled to gravity. This also represents a novel result, not present in the previous literature.

The layout of the paper is as follows. In section 2, we present a short overview of the higher gauge theory construction of the SM action coupled to gravity. We introduce the notion of a 3-group and the corresponding topological $3BF$ action, and then we demonstrate how these should be chosen and deformed with constraint terms in order to reproduce the SM coupled to Einstein–Cartan gravity. Section 3 contains the short review of the gauge symmetry group of the topological $3BF$ action, and the analysis how this group is being broken by each individual constraint term in the action. This represents the study of the explicit symmetry breaking as a consequence of the constraints present in the theory. Section 4 is devoted to the 3-group formulation of the Proca action. The Proca action is a necessary ingredient one needs to understand, in order to compare it to the action obtained via spontaneous symmetry breaking. The analysis of the spontaneous symmetry breaking and the Higgs mechanism is presented in full detail in section 5, discussing the most convenient example of electroweak theory. Despite the fact that it is conceptually the same as the ordinary Higgs mechanism described in textbooks, the specific properties of the $3BF$ formulation of the action renders the technical details of the procedure highly nontrivial, and represents one of the main results of the paper. Our concluding remarks are given in section 6, with a summary and a discussion of the results. The Appendices contain some additional technical material.

Our notation and conventions are as follows. Spacetime indices, denoted by the mid-alphabet Greek letters $\mu,\nu,\dots$, are raised and lowered by the spacetime metric $g_{\mu\nu}$. The Lorentz metric is denoted as $\eta_{ab} = \mathrm{diag}(-1,+1,+1,+1)$. The indices that are counting the generators of Lie groups G, H, and L are denoted with initial Greek letters $\alpha,\beta,\dots$, lowercase initial Latin letters $a,b,c,\dots$, and uppercase Latin indices $A,B,C,\dots$, respectively. The generators themselves are typically denoted as τ_α, t_a and T_A, respectively. We work in the natural system of units, defined by $c = \hbar = 1$ and $G = l_\mathrm{p}^2$, where l_p is the Planck length. All additional notation and conventions used throughout the paper are explicitly defined in the text where they first appear.

2. Review of the 3*BF* theory with constraints

The main idea of the so-called higher gauge theory approach [9] is to describe the symmetry of a physical theory with a mathematical structure that is different from an ordinary Lie group. In particular, in the context of category theory description of groups, the natural generalizations are the structures called n-groups. A reader interested in the mathematical details of higher category theory, n-groups and L_∞ algebras is referred to corresponing literature [9, 10, 12, 14, 24, 43, 44, 52, 57–66]. For the purpose of this work, the attention focuses on the notion of a strict Lie 3-group structure as a symmetry of the theory.

A strict Lie 3-group is defined as a 3-category over a single object with invertibile 1-, 2- and 3-morphisms, and it is equivalent to a Lie 2-crossed module $(L \xrightarrow{\delta} H \xrightarrow{\partial} G, \triangleright, \{\_,\_\}_{\mathrm{pf}})$. A 2-crossed module consists of three Lie groups L, H, and G, with the homomorphism $\partial : H \to G$, the homomorphism $\delta : L \to H$, the action $\triangleright : G \times X \to X$ (where $X \in \{G, H, L\}$), and Peiffer lifting $\{\_,\_\}_{\mathrm{pf}} : H \times H \to L$. In order to give rise to a Lie 2-crossed module, all these maps must satisfy a set of appropriate axioms [25].

One of the main benefits of the categorical generalization of the notion of a Lie group lies in the corresponding generalization of certain notions in differential geometry. In particular, for the case of a Lie 3-group, one can generalize the notion of parallel transport along a curve to the notions of parallel transport across a surface and through a volume. These operations are described by the so-called 2- and 3-holonomies, which are in turn constructed from the $\mathfrak{h}$-valued connection two-form β and the $\mathfrak{l}$-valued connection three-form γ, in addition to the ordinary $\mathfrak{g}$-valued connection one-form α. The triple (α, β, γ) is called a 3-connection, and $\mathfrak{g}$, $\mathfrak{h}$, $\mathfrak{l}$ are Lie algebras of the Lie groups G, H, and L, respectively.

The mathematical structure of a 3-group gives rise to a natural choice of an action, called $3BF$ action, that can be constructed from the 3-connection. The $3BF$ action is purely topological, and defined as:

$$S_{3BF} = \int_{\mathcal{M}_4} \langle B \wedge \mathcal{F} \rangle_{\mathfrak{g}} + \langle C \wedge \mathcal{G} \rangle_{\mathfrak{h}} + \langle D \wedge \mathcal{H} \rangle_{\mathfrak{l}}. \tag{1}$$

Quantities B, C and D are two-, one- and zero-forms, which play the role of the Lagrange multipliers, and they are elements of algebras $\mathfrak{g}$, $\mathfrak{h}$ and $\mathfrak{l}$, respectively. The field strengths $\mathcal{F}$, $\mathcal{G}$ and $\mathcal{H}$ are defined as

$$\mathcal{F} = \mathrm{d}\alpha + \alpha \wedge \alpha - \partial\beta\,, \qquad \mathcal{G} = \mathrm{d}\beta + \alpha \wedge^{\triangleright} \beta - \delta\gamma\,, \qquad \mathcal{H} = \mathrm{d}\gamma + \alpha \wedge^{\triangleright} \gamma + \{\beta \wedge \beta\}_{\mathrm{pf}}\,, \tag{2}$$

and they are called fake curvatures for the connection one-form α, two-form β and three-form γ. Bilinear forms $\langle\_,\_\rangle_{\mathfrak{g}}$, $\langle\_,\_\rangle_{\mathfrak{h}}$ and $\langle\_,\_\rangle_{\mathfrak{l}}$ are assumed to be symmetric, nondegenerate and G-invariant, and they map a pair of algebra elements into a real number. Evaluated on the corresponding basis vectors, the bilinear forms are written in components as follows:

$$\langle \tau_\alpha\,, \tau_\beta \rangle_{\mathfrak{g}} = g_{\alpha\beta}\,, \qquad \langle t_a\,, t_b \rangle_{\mathfrak{h}} = g_{ab}\,, \qquad \langle T_A\,, T_B \rangle_{\mathfrak{l}} = g_{AB}\,. \tag{3}$$

Since the bilinear forms are assumed to be nondegenerate, the inverses of the above components also exist, denoted as $g^{\alpha\beta}$, g^{ab} and g^{AB}. They are collectively used to raise and lower all group indices as necessary.

Varying the action (1) with respect to Lagrange multipliers, one obtains the equations of motion for fake curvatures:

$$\mathcal{F} = 0\,, \qquad \mathcal{G} = 0\,, \qquad \mathcal{H} = 0\,. \tag{4}$$

Also, varying with respect to the connections α, β and γ, we get the remaining three equations of motion:

$$\nabla B_\alpha - \triangleright_{\alpha a}{}^{b} C_b \wedge \beta^a + \triangleright_{\alpha B}{}^{A} D_A \wedge \gamma^B = 0\,, \tag{5}$$

$$\nabla C_a - \partial_a{}^\alpha B_\alpha + 2X_{(ab)}{}^A D_A \wedge \beta^b = 0\,, \tag{6}$$

$$\nabla D_A + \delta_A{}^a C_a = 0\,. \tag{7}$$

Here the quantities $X_{ab}{}^A$ are components of the Peiffer lifting evaluated on the basis, $\{t_a, t_b\}_{\mathrm{pf}} \equiv X_{ab}{}^A T_A$. An analogous notation is used for the homomorphisms ∂ and δ, and the action $\triangleright$:

$$\partial t_a = \partial_a{}^\alpha \tau_\alpha\,, \qquad \delta T_A = \delta_A{}^a t_a\,, \qquad \tau_\alpha \triangleright t_a = \triangleright_{\alpha a}{}^b t_b\,, \qquad \tau_\alpha \triangleright T_A = \triangleright_{\alpha A}{}^B T_B\,. \tag{8}$$

At this point it is important to note one feature of the relationship between a 3-group and the three bilinear forms. Specifically, the requirement that the bilinear forms must be G-invariant places a restriction on the allowed choice of the action $\triangleright$. This is specified in the following theorem.

Theorem. *Given a 2-crossed module* $(L \xrightarrow{\delta} H \xrightarrow{\partial} G,\ \triangleright,\ \{\_,\_\}_{\mathrm{pf}})$ *and symmetric, nondegenerate bilinear forms* $\langle\_,\_\rangle_{\mathfrak{g}}$, $\langle\_,\_\rangle_{\mathfrak{h}}$ *and* $\langle\_,\_\rangle_{\mathfrak{l}}$, *if the bilinear forms are G-invariant then the components of the action* $\triangleright_{\alpha\beta\gamma}$, $\triangleright_{\alpha ab}$ *and* $\triangleright_{\alpha AB}$ *are antisymmetric with respect to the second and third index. In addition, there exists a choice of basis in Lie algebras* $\mathfrak{g}$, $\mathfrak{h}$ *and* $\mathfrak{l}$ *such that* $\triangleright_{\alpha\beta}{}^\gamma$, $\triangleright_{\alpha a}{}^b$ *and* $\triangleright_{\alpha A}{}^B$ *have vanishing diagonal elements with respect to the second and third index, and in this basis the bilinear form is also diagonal.*

For the proof of the theorem, see appendix A. It is important to stress that these restrictions on the action $\triangleright$ arise only due to assumed G-invariance of the bilinear forms, and do not hold otherwise. Furthermore, these restrictions will play an important role later in section 5, in the discussion of the Higgs mechanism.

As noted in the Introduction section, one can apply *BF* and more generally *nBF* actions to construct physically interesting models and even realistic theories. This is typically performed by adding an additional term to the topological action, called a constraint term, which deforms the theory and may give rise to physical degrees of freedom. In the context of ordinary *BF* theory, relevant physical models include Plebanski formulation of general relativity, Husain–Kuchař model, MacDowell–Mansouri model, JT model, and general relativity in $2+1$ dimensions and more than 4 dimensions. A comprehensive reivew of these models is given in [41]. Regarding the 2*BF* theory, relevant models include Einstein–Cartan gravity and Yang–Mills theory, see [11, 25]. In the context of 3*BF* theory, one can construct models with matter fields in addition to gravity and Yang–Mills. In particular, models with scalar, Dirac, Weyl and Majorana fields coupled to Einstein–Cartan gravity and Yang–Mills theory have been constructed, including the full SM coupled to Einstein–Cartan gravity. For a review of these models see [25]. Finally, there is also a similar construction based on 4*BF* model, see [51].

In this work, we will focus our attention on a constrained 3*BF* action giving rise to the Einstein–Cartan gravity coupled to the full SM, and the electroweak model, as well as a formulation of the Proca theory.

In order to construct a physical theory based on a 3*BF* action, we need to introduce some constraints between the fields and choose an appropriate 3-group as a gauge symmetry structure. The constraints are discussed below, while the choice of the 3-group is as follows. A simple 3-group which corresponds to the SM of elementary particles coupled to Einstein–Cartan gravity in the usual way is called the SM 3-group [25], and is defined by the following choice:

$$G = SO(3,1)\times SU(3)\times SU(2)\times U(1)\,, \qquad H = \mathbb{R}^4\,, \qquad L = \mathbb{C}^4 \times \mathbb{G}^{64} \times \mathbb{G}^{64} \times \mathbb{G}^{64}\,. \tag{9}$$

The group G is a product of the Lorentz group and the usual internal gauge symmetry groups of the SM. The group H represents spacetime translations. The choice of the group L corresponds to the Higgs and fermion sector of the SM, where $\mathbb{G}$ denotes the Grassmann algebra. In addition to this choice of groups, we choose the maps ∂, δ and $\{\_\,,\_\}_{\mathrm{pf}}$ to be trivial:

$$\partial h = \mathbb{1}_G\,, \qquad \delta l = \mathbb{1}_H\,, \qquad \{h_1,h_2\}_{\mathrm{pf}} = \mathbb{1}_L\,, \tag{10}$$

for every $h, h_1, h_2 \in H$ and $l \in L$.

In order to complete the definition of the 3-group, we also choose the map $\triangleright$ as follows. Because of the specific structure of the group G, it is natural to distinguish $\mathfrak{g}$-indices for the Lorentz part from the internal part, and we will write the former in pairs of small alphabet indices in angular brackets, $[ab]$, while the latter will remain denoted with Greek letters from the beginning of the alphabet. Denoting the structure constants for the internal subgroup $SU(3)\times SU(2)\times U(1)$ as $f_{\alpha\beta}{}^{\gamma}$, the action of the group G on itself is defined as:

$$\triangleright_{[ab][cd]}{}^{[ef]} \equiv f_{[ab][cd]}{}^{[ef]} = \frac{1}{2}\left(\eta_{[a|c}\delta^{[f]}_{|b]}\delta^{|e]}_{d} - \eta_{[a|d}\delta^{[f]}_{|b]}\delta^{|e]}_{c}\right)\,, \qquad \triangleright_{[ab]\beta}{}^{\gamma} = 0\,, \tag{11}$$

$$\triangleright_{\alpha\beta}{}^{\gamma} = f_{\alpha\beta}{}^{\gamma}\,, \qquad \triangleright_{\alpha[ab]}{}^{[cd]} = 0\,. \tag{12}$$

Equation (11) define the action of the Lorentz subgroup on G, while equations (12) define the action of the internal subgroup on G. The action of the Lorentz and internal subgroups of G on the group H is defined as:

$$\triangleright_{[cd]a}{}^{b} = \frac{1}{2}\eta_{a[d|}\delta^{b}_{|c]}\,, \qquad \triangleright_{\alpha a}{}^{b} = 0\,. \tag{13}$$

Finally, the action of the Lorentz and internal subgroups of G on L is given in a natural way, in accordance with the transformation properties of various fermions and the Higgs scalar. For example, the action of G on left-isospin fermions is given as:

$$\triangleright_{[cd]A}{}^{B} = \frac{1}{2}\left(\sigma_{[cd]}\right)_A{}^{B}\,, \qquad \triangleright_{\alpha A}{}^{B} = \frac{1}{2}\left(\sigma_{\alpha}\right)_A{}^{B}\,. \tag{14}$$

Here the matrices $(\sigma_\alpha)_A{}^B$ are Pauli matrices, and $\left(\sigma_{[ab]}\right)_A{}^B = \frac{1}{4}[\gamma_a,\gamma_b]_A{}^B$, where γ_a are the standard Dirac matrices satisfying the anticommutation rule $\gamma_a\gamma_b + \gamma_b\gamma_a = -2\eta_{ab}$. Here we also introduce $\gamma_5 \equiv -\gamma_0\gamma_1\gamma_2\gamma_3$. In a similar way, one defines the action of group G for all other fermions and scalars in the group L, depending on their precise transformation properties (see [25] for details).

In addition to the specification of the 3-group, the action (1) also depends on the choice of bilinear forms. For the non-Abelian groups one can naturally choose the Cartan–Killing form, while for the Abelian groups there is no natural choice, and one is mostly restricted by the property that the bilinear form must be G-invariant. Taking this into account, for the SM 3-group and the action (1) we choose the bilinear forms as follows. For the algebra $\mathfrak{g}$, we have

$$g_{[ab][cd]} = \frac{1}{2}\eta_{d[a|}\eta_{|b]c}\,, \qquad g_{\alpha\beta} = \delta_{\alpha\beta}\,, \qquad g_{\alpha[ab]} = 0\,. \tag{15}$$

For the algebra $\mathfrak{h}$, due to the G-invariance, we have

$$g_{ab} = \eta_{ab}\,. \tag{16}$$

Finally, for the algebra $\mathfrak{l}$ the situation is more complicated, since the Grassmann numbers anticommute. Namely, note that for general $A, B \in \mathfrak{l}$, we can write

$$\langle A,B\rangle_{\mathfrak{l}} = A^I B^J g_{IJ}\,, \qquad \langle B,A\rangle_{\mathfrak{l}} = B^J A^I g_{JI}\,. \tag{17}$$

Since the bilinear form must be symmetric, the two expressions must be equal. However, depending on whether the coefficients A^I and B^J are Grassmann numbers or ordinary real numbers, they will either anticommute or commute, and consequently the component matrix g_{IJ} of the bilinear form must be antisymmetric or symmetric, respectively. In our case, the generators T_A of the algebra $\mathfrak{l}$ can be grouped into three classes: $T_{\tilde{A}}$ which belong to the Higgs sector, and a pair $T_{\hat{A}}$, $T^{\hat{A}}$ which belong to the fermion sector. Then, the components of the bilinear form can be written as

$$g_{AB} = \left[\begin{array}{c|cc} \delta_{\tilde{A}\tilde{B}} & 0 & 0 \\ \hline 0 & 0 & \delta_{\hat{A}}^{\hat{B}} \\ 0 & -\delta_{\hat{A}}^{\hat{B}} & 0 \end{array}\right]. \tag{18}$$

The upper-left block corresponds to the algebra $\mathbb{C}^4$, while the bottom-right block corresponds to the algebra $\mathbb{G}^{64} \times \mathbb{G}^{64} \times \mathbb{G}^{64}$.

Once we have specified both the 3-group and the bilinear forms, we can introduce the full classical action corresponding to the SM coupled to Einstein–Cartan gravity. This action is written as the 3BF action (1) plus the constraint terms that give rise to the desired dynamics, and has the following form:

$$S = S_{3BF} + S_{\text{grav}} + S_{\text{scal}} + S_{\text{Dirac}} + S_{\text{Yang–Mills}} + S_{\text{Higgs}} + S_{\text{Yukawa}} + S_{\text{spin}} + S_{\text{CC}}\,. \tag{19}$$

Here we have:

$$S_{3BF} = \int B_\alpha \wedge F^\alpha + B^{[ab]} \wedge R_{[ab]} + e_a \wedge \nabla\beta^a + \phi^A \left(\nabla\tilde{\gamma}\right)_A + \bar{\psi}_A \left(\overrightarrow{\nabla}\gamma\right)^A - \left(\bar{\gamma}\overleftarrow{\nabla}\right)_A \psi^A\,, \tag{20}$$

$$S_{\text{grav}} = -\int \lambda_{[ab]} \wedge \left(B^{[ab]} - \frac{1}{8\pi l_p^2}\varepsilon^{[ab]cd} e_c \wedge e_d\right), \tag{21}$$

$$\begin{aligned} S_{\text{scal}} = \int & \tilde{\lambda}^A \wedge \left(\tilde{\gamma}_A - H_{abcA} e^a \wedge e^b \wedge e^c\right) \\ & + \Lambda^{abA} \wedge \left(H_{abcA}\varepsilon^{cdef} e_d \wedge e_e \wedge e_f - (\nabla\phi)_A \wedge e_a \wedge e_b\right), \end{aligned} \tag{22}$$

$$\begin{aligned} S_{\text{Dirac}} = \int & \bar{\lambda}_A \wedge \left(\gamma^A + \frac{\mathrm{i}}{6}\varepsilon_{abcd} e^a \wedge e^b \wedge e^c \left(\gamma^d\psi\right)^A\right) \\ & - \lambda^A \wedge \left(\bar{\gamma}_A - \frac{\mathrm{i}}{6}\varepsilon_{abcd} e^a \wedge e^b \wedge e^c \left(\bar{\psi}\gamma^d\right)_A\right), \end{aligned} \tag{23}$$

$$\begin{aligned} S_{\text{Yang–Mills}} = \int & \lambda^\alpha \wedge \left(B_\alpha - 12 C_{\alpha\beta} M^\beta{}_{ab} e^a \wedge e^b\right) \\ & + \zeta_\alpha{}^{ab}\left(M^\alpha{}_{ab}\varepsilon_{cdef} e^c \wedge e^d \wedge e^e \wedge e^f - F^\alpha \wedge e_a \wedge e_b\right), \end{aligned} \tag{24}$$

$$S_{\text{Higgs}} = -\int \frac{2}{4!}\chi\left(\phi^A\phi_A - v^2\right)^2 \varepsilon_{abcd} e^a\wedge e^b\wedge e^c\wedge e^d\,, \tag{25}$$

$$S_{\text{Yukawa}} = -\int \frac{2}{4!} Y_{ABC}\bar{\psi}^A\psi^B\phi^C \varepsilon_{abcd} e^a\wedge e^b\wedge e^c\wedge e^d\,, \tag{26}$$

$$S_{\text{spin}} = \int 2\pi\,\mathrm{i} l_p^2 \bar{\psi}_A\gamma_5\gamma^a\psi^A \varepsilon_{abcd} e^b\wedge e^c\wedge \beta^d\,, \tag{27}$$

$$S_{\text{CC}} = -\int \frac{1}{96\pi l_p^2}\Lambda\varepsilon_{abcd} e^a\wedge e^b\wedge e^c\wedge e^d\,. \tag{28}$$

In addition to the topological $3BF$ term (20), one can recognize:

- the gravitational constraint term (21), giving rise to the gravitational degrees of freedom,
- the scalar constraint (22), giving rise to the massless scalar degrees of freedom,
- the Dirac constriant (23), giving rise to the massless fermions,
- the Yang–Mills constraint (24), giving rise to the massless gauge bosons,
- the Higgs potential constraint (25), containing the self-interactions and the mass of the Higgs field,
- the Yukawa constraint (26), containing the interactions between the Higgs field and fermions, as well as fermion mixing matrices,
- the spin constraint (27), necessary for the appropriate coupling between fermion spins and torsion, and
- the CC constraint (28), introducing the cosmological constant.

The following free parameters are present in the action:

- l_{p} is the Planck length, featuring in S_{grav}, S_{spin} and S_{CC},
- $C_{\alpha\beta}$ represents the gauge coupling constant bilinear form, featuring in $S_{\text{Yang–Mills}}$,
- χ is the coupling constant for the quartic self-interaction of the Higgs field, featuring in S_{Higgs},
- v is the vacuum expectation value of the Higgs field, also featuring in S_{Higgs},
- Y_{ABC} represent the Yukawa couplings and fermion mixing matrices, featuring in S_{Yukawa}, and
- Λ is the cosmological constant, featuring in S_{CC}.

The topological part S_{3BF} and the constraints S_{scal} and S_{Dirac} do not contain any free parameters. Finally, note that the topological part S_{3BF} is now rewritten in new notation. Specifically, $\mathcal{F}$ is split into the internal symmetry field strength F^α (which is a function of the internal symmetry connection α^α) and the Riemann curvature two-form $R_{[ab]}$ (which is a function of the spin connection $\omega^{[ab]}$). The Lagrange multiplier C is rewritten as the tetrad field one-form e_a, and the Lagrange multiplier D is rewritten as a tuple of scalar and fermion fields $(\phi^A,\psi^A,\bar{\psi}_A)$. This change of notation also suggests the physical interpretation of the fields in (1).

Let us discuss the equations of motion for this action. After a certain amount of calculation, we obtain the equations solved for all Lagrange multiplier fields, in terms of the dynamical fields and their derivatives:

$$
\begin{aligned}
M_{\alpha ab} &= -\frac{1}{48}\varepsilon_{abcd}F_{\alpha}{}^{\mu\nu}e^{c}{}_{\mu}e^{d}{}_{\nu}\,, & \zeta^{\alpha ab} &= \frac{1}{4}C_{\beta}{}^{\alpha}\varepsilon^{abcd}F^{\beta}{}_{\mu\nu}e_{c}{}^{\mu}e_{d}{}^{\nu}\,,\\
\lambda_{\alpha\mu\nu} &= -F_{\alpha\mu\nu}\,, & B_{\alpha\mu\nu} &= -\frac{e}{2}C_{\alpha}{}^{\beta}\varepsilon_{\mu\nu\rho\sigma}F_{\beta}{}^{\rho\sigma}\,,\\
\lambda_{[ab]\mu\nu} &= R_{[ab]\mu\nu}\,, & B_{[ab]\mu\nu} &= \frac{1}{8\pi l_p^2}\varepsilon_{[ab]cd}e^{c}{}_{\mu}e^{d}{}_{\nu}\,,\\
\tilde{\lambda}^{A}{}_{\mu} &= \left(\nabla_{\mu}\phi\right)^{A}\,, & \tilde{\gamma}^{A}{}_{\mu\nu\rho} &= -e\varepsilon_{\mu\nu\rho\sigma}\left(\nabla^{\sigma}\phi\right)^{A}\,,\\
H^{abcA} &= \frac{1}{6e}\varepsilon^{\mu\nu\rho\sigma}\left(\nabla_{\mu}\phi\right)^{A}e^{a}{}_{\nu}e^{b}{}_{\rho}e^{c}{}_{\sigma}\,, & \Lambda^{abA}{}_{\mu} &= \frac{1}{6e}g_{\mu\lambda}\varepsilon^{\lambda\nu\rho\sigma}\left(\nabla_{\nu}\phi\right)^{A}e^{a}{}_{\rho}e^{b}{}_{\sigma}\,,\\
\gamma^{A}{}_{\mu\nu\rho} &= -\mathrm{i}\varepsilon_{abcd}e^{a}{}_{\mu}e^{b}{}_{\nu}e^{c}{}_{\rho}\left(\gamma^{d}\psi\right)^{A}\,, & \bar{\gamma}_{A\mu\nu\rho} &= \mathrm{i}\varepsilon_{abcd}e^{a}{}_{\mu}e^{b}{}_{\nu}e^{c}{}_{\rho}\left(\bar{\psi}\gamma^{d}\right)_{A}\,,\\
\lambda^{A}{}_{\mu} &= \left(\overrightarrow{\nabla}_{\mu}\psi\right)^{A}\,, & \bar{\lambda}_{A\mu} &= \left(\bar{\psi}\overleftarrow{\nabla}_{\mu}\right)_{A}\,,\\
\beta^{a}{}_{\mu\nu} &= 0\,.
\end{aligned}
\tag{29}
$$

Next we look at the equations of motion for the dynamical fields. The spin connection $\omega_{[ab]\mu}$ is not equivalent to the Levi-Civita connection, since fermionic fields give rise to nonzero torsion T_a:

$$
\omega_{[ab]\mu} = \Delta_{[ab]\mu} + K_{[ab]\mu}\,, \tag{30}
$$

$$
T_a \equiv \nabla e_a = 2\pi\,\mathrm{i}l_p^2\,\varepsilon_{abcd}\,e^b \wedge e^c\,\bar{\psi}_A\gamma_5\gamma^d\psi^A = 2\pi\,\mathrm{i}l_p^2\,s_a\,. \tag{31}
$$

Spin connection is represented as sum of Ricci rotation coefficients $\Delta_{[ab]\mu}$ and contorsion tensor $K_{[ab]\mu}$. Torsion 2-form is proportional to spin 2-form s_a, as usual in the Einstein–Cartan gravity.

The Einstein field equation has the usual form:

$$
R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu} = 8\pi\,l_p^2\,T_{\mu\nu}\,, \tag{32}
$$

where the stress–energy tensor is given as:

$$
\begin{aligned}
T_{\mu\nu} = {} & F^{\alpha}{}_{\mu\rho}C_{\alpha}{}^{\beta}F_{\beta\nu}{}^{\rho} - \frac{1}{4}g_{\mu\nu}F^{\alpha}{}_{\rho\sigma}C_{\alpha}{}^{\beta}F_{\beta}{}^{\rho\sigma}\\
& + \nabla_{\mu}\phi^{A}\nabla_{\nu}\phi_{A} - \frac{1}{2}g_{\mu\nu}\left(\nabla_{\rho}\phi^{A}\nabla^{\rho}\phi_{A} + 2\chi\left(\phi^{A}\phi_{A} - v^{2}\right)^{2}\right)\\
& + \frac{\mathrm{i}}{2}\left(\bar{\psi}_{A}\overleftrightarrow{\nabla}_{\mu}\gamma_{d}\psi^{A}\right)e^{d}_{\nu} - \frac{1}{2}g_{\mu\nu}\left(\mathrm{i}\left(\bar{\psi}_{A}\overleftrightarrow{\nabla}_{\rho}\gamma^{d}\psi^{A}\right)e_{d}{}^{\rho} - 2Y_{ABC}\bar{\psi}^{A}\psi^{B}\phi^{C}\right)\,.
\end{aligned}
\tag{33}
$$

It features three parts, describing the Yang–Mills, scalar and fermion stress–energy, respectively.

Equations of motion for fermion and scalar fields are

$$
\left(\mathrm{i}\gamma^{\mu}\overrightarrow{\nabla}_{\mu}\delta^{A}_{B} - Y^{A}{}_{BC}\phi^{C}\right)\psi^{B} = 0\,, \tag{34}
$$

$$
\bar{\psi}_{B}\left(\delta^{B}_{A}\mathrm{i}\overleftarrow{\nabla}_{\mu}\gamma^{\mu} + Y_{BAC}\phi^{C}\right) = 0\,, \tag{35}
$$

$$
\nabla_{\mu}\nabla^{\mu}\phi^{A} - 4\chi\left(\phi^{B}\phi_{B} - v^{2}\right)\phi^{A} = 0\,, \tag{36}
$$

while the equation of motion for Yang–Mills fields is:

$$\nabla_\mu F_\alpha{}^{\mu\nu} + \frac{1}{2} C^{-1}{}_\alpha{}^\beta \left(\rhd_{\beta AB} \left(\phi^A \nabla^\nu \phi^B - \phi^B \nabla^\nu \phi^A \right) + \mathrm{i} \bar{\psi}_A \psi_B \left(\rhd_{\beta C}{}^A \gamma^{\nu CB} - \gamma^{\nu AC} \rhd_{\beta C}{}^B \right) \right) = 0 . \tag{37}$$

One can observe that all these equations of motion correspond precisely to the SM coupled to Einstein–Cartan gravity, along with the cosmological constant term.

This completes the review of the realistic classical theory based on the constrained $3BF$ action and the 3-group approach. In the next section, we turn to the discussion of the symmetries of this model.

3. Constrained 3*BF* action and explicit symmetry breaking

By adding simplicity constraints to the topological $3BF$ action, we reproduce the equations of motions for all known fields. But adding constraints also leads to breaking of the initial symmetry. In what follows, we will study each of the constraints separately, in order to determine which constraint breaks which symmetry group.

The total symmetry group of the topological $3BF$ action has been studied in [56, 67], and has been shown to have the form $\mathcal{G}_{3BF} = (\tilde{G} \ltimes (\tilde{H}_L \ltimes (\tilde{N} \times \tilde{M}))) \ltimes \mathrm{HT}$. The semidirect and direct products of groups $\tilde{G}$, $\tilde{H}_L$, $\tilde{M}$, $\tilde{N}$ correspond to the ordinary gauge symmetry of the action, while the Henneaux–Teitelboim (HT) group corresponds to the so-called HT symmetry, which is trivial on-shell (for a review, see [67]).

The choice of the SM 3-group, introduced in the previous section, implies however a more specific structure for the gauge group $\mathcal{G}_{3BF}$. Namely, in the general case, the generators of the $\tilde{H}_L$ group can be naturally divided into the $\hat{H}$-generators and $\hat{L}$-generators, satisfying the Lie algebra commutation relations of the form

$$\left[\hat{H},\hat{H}\right] \sim \hat{L} , \qquad \left[\hat{H},\hat{L}\right] \sim 0 , \qquad \left[\hat{L},\hat{L}\right] \sim 0 , \tag{38}$$

where the structure constants in the first commutator are proportional to the components of the Peiffer lifting map in the given 3-group (see [56] for a detailed analysis). Nevertheless, for our specific choice of the 3-group the Peiffer lifting is trivial, implying that

$$\left[\hat{H},\hat{H}\right] \sim 0 , \qquad \left[\hat{H},\hat{L}\right] \sim 0 , \qquad \left[\hat{L},\hat{L}\right] \sim 0 . \tag{39}$$

This means that the group $\tilde{H}_L$ can be rewritten as a direct product

$$\tilde{H}_L = \tilde{H} \times \tilde{L} \tag{40}$$

of two Abelian normal subgroups $\tilde{H}$ and $\tilde{L}$. Additionally, since in general the $\hat{H}$-generators are responsible for the semidirect product $\tilde{H}_L \ltimes (\tilde{N} \times \tilde{M})$ within $\mathcal{G}_{3BF}$, with commutation relations of the form

$$\left[\hat{H},\hat{N}\right] \sim \hat{M} , \qquad \left[\hat{H},\hat{M}\right] \sim 0 , \qquad \left[\hat{L},\hat{M}\right] \sim 0 , \qquad \left[\hat{L},\hat{N}\right] \sim 0 , \tag{41}$$

it is straightforward to conclude that in a case of any 3-group with trivial Peiffer lifting, the symmetry group $\mathcal{G}_{3BF}$ takes on a more specific form:

$$\mathcal{G}_{3BF} = \left(\tilde{G} \ltimes \left(\tilde{L} \times \left(\tilde{H} \ltimes \left(\tilde{N} \times \tilde{M}\right)\right)\right)\right) \ltimes \mathrm{HT} . \tag{42}$$

Thus, (42) represents the gauge group of the 3BF theory based on the SM 3-group.

Let us now introduce the action of this group on the fields present in the 3BF action. For a general 3-group, the infinitesimal transformations of the ordinary gauge part are derived in [56] and listed as form-variations, while the infinitesimal transformations of the HT-part are defined in [67]. For the special case of the SM 3-group, the ordinary gauge transformations are given explicitly as follows:

$$
\begin{aligned}
\delta_0^g \alpha^\alpha &= -\nabla \epsilon_{\mathfrak{g}}{}^\alpha \,, \\
\delta_0^g \omega^{[ab]} &= -\nabla \epsilon_{\mathfrak{g}}{}^{[ab]} \,, \\
\delta_0^g \beta^a &= \rhd_{\alpha b}{}^a \epsilon_{\mathfrak{g}}{}^\alpha \beta^b - \nabla \epsilon_{\mathfrak{h}}{}^a \,, \\
\delta_0^g \gamma^A &= \rhd_{\alpha B}{}^A \epsilon_{\mathfrak{g}}{}^\alpha \gamma^B + \nabla \epsilon_{\mathfrak{l}}{}^A \,, \\
\delta_0^g B^\alpha &= f_{\beta\gamma}{}^\alpha \epsilon_{\mathfrak{g}}{}^\beta B^\gamma + e_a \wedge \epsilon_{\mathfrak{h}}{}^b \rhd^\alpha{}_b{}^a - D_A \rhd^\alpha{}_B{}^A \epsilon_{\mathfrak{l}}{}^B - \nabla \epsilon_{\mathfrak{m}}{}^\alpha + \beta_b \rhd^\alpha{}_a{}^b \epsilon_{\mathfrak{n}}{}^a \,, \\
\delta_0^g B^{[ab]} &= f_{[gh][ij]}{}^{[ab]} B^{[ij]} \epsilon_{\mathfrak{g}}{}^{[gh]} - \nabla \epsilon_{\mathfrak{m}}{}^{[ab]} + e_c \wedge \epsilon_{\mathfrak{h}}{}^d \rhd^{[ab]}{}_d{}^c + \beta_d \rhd^{[ab]}{}_c{}^d \epsilon_{\mathfrak{n}}{}^c - \epsilon_{\mathfrak{l}}{}^A \rhd^{[ab]}{}_A{}^B D_B \,, \\
\delta_0^g e^a &= -\nabla \epsilon_{\mathfrak{n}}{}^a + \epsilon_{\mathfrak{g}}{}^\alpha \rhd_{\alpha b}{}^a e^b \,, \\
\delta_0^g D^A &= \rhd_{\alpha B}{}^A \epsilon_{\mathfrak{g}}{}^\alpha D^B \,.
\end{aligned}
\tag{43}
$$

The transformations are specified by the five free parameters, corresponding to their Lie algebras—the parameters $\epsilon_{\mathfrak{g}}{}^\alpha$ and $\epsilon_{\mathfrak{n}}{}^a$ are zero-forms, $\epsilon_{\mathfrak{h}}{}^a$ and $\epsilon_{\mathfrak{m}}{}^\alpha$ are one-forms, and $\epsilon_{\mathfrak{l}}{}^A$ are three-forms.

Regarding the HT symmetry, the infinitesimal transformations are most easily expressed in the following matrix form [67]:

$$
\begin{pmatrix}
\delta_0^{\rm HT} B^\alpha{}_{\mu\nu} \\
\delta_0^{\rm HT} C^a{}_\mu \\
\delta_0^{\rm HT} D^A \\
\delta_0^{\rm HT} \alpha^\alpha{}_\mu \\
\delta_0^{\rm HT} \beta^a{}_{\mu\nu} \\
\delta_0^{\rm HT} \gamma^A{}_{\mu\nu\rho}
\end{pmatrix}
=
\left(
\begin{array}{c|c|c|c|c|c}
\epsilon^{\alpha\beta}{}_{\mu\nu\sigma\lambda} & \epsilon^{\alpha b}{}_{\mu\nu\sigma} & \epsilon^{\alpha B}{}_{\mu\nu} & \epsilon^{\alpha\beta}{}_{\mu\nu\sigma} & \epsilon^{\alpha b}{}_{\mu\nu\sigma\lambda} & \epsilon^{\alpha B}{}_{\mu\nu\sigma\lambda\xi} \\ \hline
\mu^{a\beta}{}_{\mu\sigma\lambda} & \epsilon^{ab}{}_{\mu\sigma} & \epsilon^{aB}{}_{\mu} & \epsilon^{a\beta}{}_{\mu\sigma} & \epsilon^{ab}{}_{\mu\sigma\lambda} & \epsilon^{aB}{}_{\mu\sigma\lambda\xi} \\ \hline
\mu^{A\beta}{}_{\sigma\lambda} & \mu^{Ab}{}_{\sigma} & \epsilon^{AB} & \epsilon^{A\beta}{}_{\sigma} & \epsilon^{Ab}{}_{\sigma\lambda} & \epsilon^{AB}{}_{\sigma\lambda\xi} \\ \hline
\mu^{\alpha\beta}{}_{\mu\sigma\lambda} & \mu^{\alpha b}{}_{\mu\sigma} & \mu^{\alpha B}{}_{\mu} & \epsilon^{\alpha\beta}{}_{\mu\sigma} & \epsilon^{\alpha b}{}_{\mu\sigma\lambda} & \epsilon^{\alpha B}{}_{\mu\sigma\lambda\xi} \\ \hline
\mu^{a\beta}{}_{\mu\nu\sigma\lambda} & \mu^{ab}{}_{\mu\nu\sigma} & \mu^{aB}{}_{\mu\nu} & \mu^{a\beta}{}_{\mu\nu\sigma} & \epsilon^{ab}{}_{\mu\nu\sigma\lambda} & \epsilon^{aB}{}_{\mu\nu\sigma\lambda\xi} \\ \hline
\mu^{A\beta}{}_{\mu\nu\rho\sigma\lambda} & \mu^{Ab}{}_{\mu\nu\rho\sigma} & \mu^{AB}{}_{\mu\nu\rho} & \mu^{A\beta}{}_{\mu\nu\rho\sigma} & \mu^{Ab}{}_{\mu\nu\rho\sigma\lambda} & \epsilon^{AB}{}_{\mu\nu\rho\sigma\lambda\xi}
\end{array}
\right)
$$

$$
\times
\begin{pmatrix}
\frac{1}{2}\frac{\delta S}{\delta B^\beta{}_{\sigma\lambda}} \\
\frac{\delta S}{\delta C^b{}_\sigma} \\
\frac{\delta S}{\delta D^B} \\
\frac{\delta S}{\delta \alpha^\beta{}_\sigma} \\
\frac{1}{2}\frac{\delta S}{\delta \beta^b{}_{\sigma\lambda}} \\
\frac{1}{3!}\frac{\delta S}{\delta \gamma^B{}_{\sigma\lambda\xi}}
\end{pmatrix} .
\tag{44}
$$

Here, in order to ensure the antisymmetry of the parameter matrix, the following identities must hold:

$$
\begin{gathered}
\mu^{b\alpha}{}_{\sigma\mu\nu} = -\epsilon^{\alpha b}{}_{\mu\nu\sigma}\,, \quad \mu^{B\alpha}{}_{\mu\nu} = -\epsilon^{\alpha B}{}_{\mu\nu}\,, \quad \mu^{\beta\alpha}{}_{\sigma\mu\nu} = -\epsilon^{\alpha\beta}{}_{\mu\nu\sigma}\,, \\
\mu^{b\alpha}{}_{\sigma\lambda\mu\nu} = -\epsilon^{\alpha b}{}_{\mu\nu\sigma\lambda}\,, \quad \mu^{B\alpha}{}_{\sigma\lambda\xi\,\mu\nu} = -\epsilon^{\alpha B}{}_{\mu\nu\sigma\lambda\xi}\,. \\
\mu^{Ba}{}_{\mu} = -\epsilon^{aB}{}_{\mu}\,, \quad \mu^{\beta a}{}_{\sigma\mu} = -\epsilon^{a\beta}{}_{\mu\sigma}\,, \\
\mu^{ba}{}_{\sigma\lambda\mu} = -\epsilon^{ab}{}_{\mu\sigma\lambda}\,, \quad \mu^{Ba}{}_{\sigma\lambda\xi\mu} = -\epsilon^{aB}{}_{\mu\sigma\lambda\xi}\,, \\
\mu^{\beta A}{}_{\sigma} = -\epsilon^{A\beta}{}_{\sigma}\,, \quad \mu^{bA}{}_{\sigma\lambda} = -\epsilon^{Ab}{}_{\sigma\lambda}\,, \quad \mu^{BA}{}_{\sigma\lambda\xi} = -\epsilon^{AB}{}_{\sigma\lambda\xi}\,, \\
\mu^{b\alpha}{}_{\sigma\lambda\mu} = -\epsilon^{\alpha b}{}_{\mu\sigma\lambda}\,, \quad \mu^{B\alpha}{}_{\sigma\lambda\xi\mu} = -\epsilon^{\alpha B}{}_{\mu\sigma\lambda\xi}\,, \quad \mu^{Ba}{}_{\sigma\lambda\xi\mu\nu} = -\epsilon^{aB}{}_{\mu\nu\sigma\lambda\xi}\,.
\end{gathered}
\tag{45}
$$

For more information about properties and importance of HT transformations, see [67].

It is straightforward (if algebraically a bit involved) to verify that transformations (43) and (44) keep the topological action (1) invariant. However, this is not the case for the constrained action (19), since each of the constraint terms may explicitly break one or more of these symmetries. In order to determine which symmetries are preserved and which are broken, and by which constraint term, we proceed as follows. For every individual constraint, we introduce the action

$$
S = S_{3BF} + S_{\text{constraint}}\,,
\tag{46}
$$

and we take the variation of this action with respect to (43). The topological part S_{3BF} is known to be already invariant, which means that the invariance of the action S reduces to the requirement

$$
\delta_0^g S_{\text{constraint}} = 0\,.
\tag{47}
$$

This requirement may not be met automatically, but only by fixing the values of certain subset of parameters $\epsilon_{\mathfrak{g}}{}^{\alpha}$, $\epsilon_{\mathfrak{h}}{}^{a}$, $\epsilon_{\mathfrak{l}}{}^{A}$, $\epsilon_{\mathfrak{m}}{}^{\alpha}$, and $\epsilon_{\mathfrak{n}}{}^{a}$. Each parameter that needs to be fixed indicates that the corresponding symmetry subgroup is broken by the constraint. In the following Subsections, we shall investigate each of the constraints (21)–(28), and use (47) to discuss which symmetries are preserved and which are broken.

One should emphasize that the above method based on (47) makes sense only for the ordinary gauge symmetry, whereas the HT symmetry cannot be studied this way. Namely, as was explained in detail in [67], the definition (44) of the HT symmetry explicitly depends on the form of the action. This means that the very process of adding a constraint term to the action changes the HT symmetry group in a nontrivial way, most often by *increasing* its number of generators and parameters, so that the HT group of the constrained theory is *larger* than the HT group of the topological symmetry. This stands in sharp contrast to the ordinary gauge group, which is being broken down to one of its subgroups by the same process. Therefore, the question of explicit symmetry breaking by the introduction of the constraint term makes sense exclusively for the ordinary gauge symmetry, and cannot be even formulated for the HT symmetry.

3.1. Gravitational simplicity constraint

As explained above, the procedure for analyzing the symmetry breaking in the case of the gravitational simplicity constraint boils down to the calculation of the form variation of (21) with respect to (43) and then the discussion of the requirement (47). Specifically, we have:

$$\delta_0^g S_{\text{grav}} = -\int \Bigg(\delta_0^g \lambda_{[ab]} \wedge \left(B^{[ab]} - \frac{1}{16\pi\, l_p^2} \varepsilon^{[ab]cd} e_c \wedge e_d \right) + \lambda_{[ab]} \wedge \left(\delta_0^g B^{[ab]} + \frac{2}{16\pi\, l_p^2} \varepsilon^{[ab]cd} \delta_0^g e_c \wedge e_d \right) \Bigg) . \tag{48}$$

The variation of the Lagrange multiplier $\lambda_{[ab]}$ is not defined initially, so we choose to define it in such way to preserve as many symmetries as we can. Substituting (43) into (48), after some algebra, the variation of the gravitational constraint becomes:

$$\begin{aligned} \delta_0^g S_{\text{grav}} = \int & \left(\delta_0^g \lambda_{[ij]} - \lambda_{[i|h} \epsilon_{\mathfrak{g}\,|j]}{}^h \right) \wedge \left(B^{[ij]} - \frac{1}{16\pi\, l_p^2} \varepsilon^{[ij]nm} e_n \wedge e_m \right) \\ & + \lambda_{[ab]} \wedge e_d \wedge \left(\epsilon_{\mathfrak{h}}{}^{[a|} \eta^{|b]d} - \frac{1}{8\pi\, l_p^2} \varepsilon^{[ab]cd} \eta_{fc} \nabla \epsilon_{\mathfrak{n}}{}^f \right) \\ & + \lambda_{[ab]} \wedge \left(\epsilon_{\mathfrak{n}}{}^{[a|} \beta^{|b]} - \nabla \epsilon_{\mathfrak{m}}{}^{[ab]} - \epsilon_{\mathfrak{l}}{}^A \rhd^{[ab]}{}_A{}^B D_B \right) , \end{aligned} \tag{49}$$

from where we can see that we can choose the variation of the multiplier $\lambda_{[ab]}$ as follows:

$$\delta_0^g \lambda_{[ij]} = -\lambda_{[ab]} f_{[gh][ij]}{}^{[ab]} \epsilon_{\mathfrak{g}}{}^{[gh]} = \lambda_{[i|h} \epsilon_{\mathfrak{g}\,|j]}{}^h . \tag{50}$$

This choice removes the whole first row in (49). However, from the second row one can see that the requirement (47) can only be satisfied if one chooses specific values of $\epsilon_{\mathfrak{h}}{}^a$, $\epsilon_{\mathfrak{n}}{}^a$, $\epsilon_{\mathfrak{m}}{}^\alpha$ and $\epsilon_{\mathfrak{l}}{}^A$. The only parameter that is not fixed is $\epsilon_{\mathfrak{g}}{}^\alpha$. This implies that this constraint breaks all symmetry groups $\tilde{M}$, $\tilde{N}$, $\tilde{L}$, and $\tilde{H}$, except for the group $\tilde{G}$, which remains unbroken.

3.2. Scalar simplicity constraint

Using the above procedure we examine all the remaining constraints. In the case of the constraint for the scalar field, we have:

$$\begin{aligned} \delta_0^g S_{\text{scal}} = \int & \Big[\delta_0^g \tilde{\lambda}_A \wedge \left(\tilde{\gamma}^A - H_{abc}{}^A e^a \wedge e^b \wedge e^c \right) \\ & + \tilde{\lambda}_A \wedge \left(\delta_0^g \tilde{\gamma}^A - \delta_0^g H_{abc}{}^A e^a \wedge e^b \wedge e^c - 3 H_{abc}{}^A \delta_0^g e^a \wedge e^b \wedge e^c \right) \\ & + \delta_0^g \Lambda^{ab}{}_A \wedge \left(H_{abc}{}^A \varepsilon^{cdef} e_d \wedge e_e \wedge e_f - \nabla \phi^A \wedge e_a \wedge e_b \right) \\ & + \Lambda^{ab}{}_A \wedge \Big(\delta_0^g H_{abc}{}^A \varepsilon^{cdef} e_d \wedge e_e \wedge e_f + 3 H_{abc}{}^A \varepsilon^{cdef} \delta_0^g e_d \wedge e_e \wedge e_f \\ & - \nabla \delta_0^g \phi^A \wedge e_a \wedge e_b - \delta_0^g \alpha^{[kl]} \rhd_{[kl]B}{}^A \phi^B \wedge e_a \wedge e_b - 2 \nabla \phi^A \wedge \delta_0^g e_a \wedge e_b \Big) \Big] . \end{aligned} \tag{51}$$

Substituting (43) into (51), we get:

$$\begin{aligned}\delta_0^g S_{\text{scal}} = \int \Big[& \left(\delta_0^g \tilde{\lambda}_A + \tilde{\lambda}_B \epsilon_{\mathfrak{g}}{}^{[ij]} \rhd_{[ij]A}{}^B\right) \wedge \left(\tilde{\gamma}^A - H_{abc}{}^A e^a \wedge e^b \wedge e^c\right) \\ & - \left(\delta_0^g H_{abc}{}^A - \epsilon_{\mathfrak{g}}{}^{[ij]} H_{abc}{}^B \rhd_{[ij]B}{}^A\right) \left(\tilde{\lambda}_A \eta^{ad}\eta^{be}\eta^{cf} - \Lambda^{ab}{}_A \varepsilon^{cdef}\right) \wedge e_d \wedge e_e \wedge e_f \\ & + \left(\delta_0^g \Lambda^{ab}{}_A + \Lambda^{ab}{}_B \epsilon_{\mathfrak{g}}{}^{[cd]} \rhd_{[cd]A}{}^B\right) \wedge \left(H_{abc}{}^A \varepsilon^{cdef} e_d \wedge e_e \wedge e_f - \nabla\phi^A \wedge e_a \wedge e_b\right) \\ & + \left(\tilde{\lambda}_A \eta^{ad}\eta^{be}\eta^{cf} - \Lambda^{ab}{}_A \varepsilon^{cdef}\right) \wedge 3H_{abc}{}^A \left(\nabla \epsilon_{\mathfrak{n}d}\right) \wedge e_e \wedge e_f \\ & + 2\Lambda^{ab}{}_A \wedge \nabla\phi^A \wedge \nabla\epsilon_{\mathfrak{n}a} \wedge e_b + \tilde{\lambda}_A \wedge \nabla \epsilon_{\mathfrak{l}}{}^A \Big] . \end{aligned} \tag{52}$$

It is obvious from fourth and fifth row from above equation that scalar field constraint breaks only $\tilde{N}$ and $\tilde{L}$ symmetries, while $\tilde{H}$ and $\tilde{M}$ symmetries are preserved since their parameters $\epsilon_{\mathfrak{h}}{}^a$ and $\epsilon_{\mathfrak{m}}{}^{\alpha}$ do not even appear in the variation (52). Finally, in order to preserve $\tilde{G}$ symmetry, we choose to define variations of new multipliers as:

$$\delta_0^g \tilde{\lambda}^A = \epsilon_{\mathfrak{g}}{}^{[ij]} \tilde{\lambda}^B \rhd_{[ij]B}{}^A , \qquad \delta_0^g H_{abc}{}^A = H_{abc}{}^B \epsilon_{\mathfrak{g}}{}^{[ij]} \rhd_{[ij]B}{}^A , \qquad \delta_0^g \Lambda^{abA} = \epsilon_{\mathfrak{g}}{}^{[cd]} \Lambda^{abB} \rhd_{[cd]B}{}^A . \tag{53}$$

3.3. *Dirac simplicity constraint*

In the same way, variation of the constraint for the Dirac fields gives us:

$$\begin{aligned}\delta_0^g S_{\text{Dirac}} = \int & \left(\delta_0^g \bar{\lambda}_A\right) \wedge \left(\gamma^A + \frac{\mathrm{i}}{6}\varepsilon_{abcd} e^a \wedge e^b \wedge e^c \left(\gamma^d \psi\right)^A\right) \\ & - \left(\delta_0^g \lambda^A\right) \wedge \left(\bar{\gamma}_A - \frac{\mathrm{i}}{6}\varepsilon_{abcd} e^a \wedge e^b \wedge e^c \left(\bar{\psi}\gamma^d\right)_A\right) \\ & + \bar{\lambda}_A \wedge \left(\left(\delta_0^g \gamma^A\right) + \frac{\mathrm{i}}{6}\varepsilon_{abcd} e^a \wedge e^b \wedge e^c \left(\gamma^d \left(\delta_0^g \psi\right)\right)^A + \frac{\mathrm{i}}{2}\varepsilon_{abcd}\left(\delta_0^g e^a\right) \wedge e^b \wedge e^c \left(\gamma^d \psi\right)^A\right) \\ & - \lambda^A \wedge \left(\left(\delta_0^g \bar{\gamma}_A\right) - \frac{\mathrm{i}}{6}\varepsilon_{abcd} e^a \wedge e^b \wedge e^c \left(\gamma^d \left(\delta_0^g \bar{\psi}\right)\right)_A - \frac{\mathrm{i}}{2}\varepsilon_{abcd}\left(\delta_0^g e^a\right) \wedge e^b \wedge e^c \left(\gamma^d \bar{\psi}\right)_A\right) . \end{aligned} \tag{54}$$

Substituting (43) into (54) gives us:

$$\begin{aligned}\delta_0^g S_{\text{Dirac}} = \int & \left(\delta_0^g \bar{\lambda}_A + \epsilon_{\mathfrak{g}}{}^{\alpha} \rhd_{\alpha A}{}^B \bar{\lambda}_B\right) \wedge \left(\gamma^A + \frac{\mathrm{i}}{6}\varepsilon_{abcd} e^a \wedge e^b \wedge e^c \left(\gamma^d \psi\right)^A\right) \\ & - \left(\delta_0^g \lambda^A + \epsilon_{\mathfrak{g}}{}^{\alpha} \rhd_{\alpha}{}^A{}_B \lambda^B\right) \wedge \left(\bar{\gamma}_A - \frac{\mathrm{i}}{6}\varepsilon_{abcd} e^a \wedge e^b \wedge e^c \left(\bar{\psi}\gamma^d\right)_A\right) \\ & + \bar{\lambda}_A \wedge \left(\nabla \epsilon_{\mathfrak{l}}{}^A + \frac{\mathrm{i}}{2}\varepsilon_{abcd}\left(\nabla \epsilon_{\mathfrak{n}}{}^a\right) \wedge e^b \wedge e^c \left(\gamma^d \psi\right)^A\right) \\ & - \lambda^A \wedge \left(\nabla \bar{\epsilon}_{\mathfrak{l}A} - \frac{\mathrm{i}}{2}\varepsilon_{abcd}\left(\nabla \epsilon_{\mathfrak{n}}{}^a\right) \wedge e^b \wedge e^c \left(\bar{\psi}\gamma^d\right)_A\right) . \end{aligned} \tag{55}$$

This constraint also breaks only $\tilde{N}$ and $\tilde{L}$ symmetries, similar to the scalar constraint. The variation laws for new multipliers are chosen to be:

$$\delta_0^g \bar{\lambda}_A = \epsilon_{\mathfrak{g}}{}^{\alpha} \rhd_{\alpha}{}^{B}{}_{A} \bar{\lambda}_B \,, \qquad \delta_0^g \lambda^A = \epsilon_{\mathfrak{g}}{}^{\alpha} \rhd_{\alpha}{}^{A}{}_{B} \lambda^B \,. \tag{56}$$

3.4. Yang–Mills simplicity constraint

Yang–Mills simplicity constraint is similar to gravitational constraint, but it contains two more Lagrange multipliers. We apply the same variation procedure as above:

$$\begin{aligned}\delta_0^g S_{\text{Yang–Mills}} = \int & \delta_0^g \lambda^{\alpha} \wedge \left(B_{\alpha} - 12 C^{\alpha\beta} M_{\beta ab} e^a \wedge e^b\right) \\ & + \lambda^{\alpha} \wedge \left(\delta_0^g B_{\alpha} - 12 C^{\alpha\beta} \delta_0^g M_{\beta ab} e^a \wedge e^b - 24 C^{\alpha\beta} M_{\beta ab} \delta_0^g e^a \wedge e^b\right) \\ & + \delta_0^g \zeta^{\alpha ab} \left(M_{\alpha ab} \varepsilon_{cdef} e^c \wedge e^d \wedge e^e \wedge e^f - F_{\alpha} \wedge e_a \wedge e_b\right) \\ & + \zeta^{\alpha ab} \left((\delta_0^g M_{\alpha ab}) \, \varepsilon_{cdef} e^c \wedge e^d \wedge e^e \wedge e^f + 4 M_{\alpha ab} \varepsilon_{cdef} (\delta_0^g e^c) \wedge e^d \wedge e^e \wedge e^f \right. \\ & \left. - (\delta_0^g F_{\alpha}) \wedge e_a \wedge e_b - 2 F_{\alpha} \wedge (\delta_0^g e_a) \wedge e_b \right) ,\end{aligned} \tag{57}$$

where

$$\delta_0^g F_{\alpha} = \epsilon_{\mathfrak{g}}{}^{\beta} F^{\gamma} \rhd_{\alpha\beta\gamma} \,. \tag{58}$$

The variation of the field strength (58) is obtained by varying the definition (2) using (43). Combining equations (43), (57), and (58) gives us:

$$\begin{aligned}\delta_0^g S_{\text{Yang–Mills}} = \int & \left(\delta_0^g \lambda^{\alpha} + \lambda^{\gamma} \epsilon_{\mathfrak{g}}{}^{\beta} \rhd_{\gamma\beta}{}^{\alpha}\right) \wedge \left(B_{\alpha} - 12 C^{\alpha\beta} M_{\beta ab} e^a \wedge e^b\right) \\ & + \left(\delta_0^g M_{\alpha ab} - \epsilon_{\mathfrak{g}}{}^{\beta} M^{\gamma}{}_{ab} \rhd_{\alpha\beta\gamma}\right) \left(\zeta^{\alpha ab} \varepsilon_{cdef} e^c \wedge e^d \wedge e^e \wedge e^f - 12 C^{\alpha}{}_{\beta} \lambda^{\beta} \wedge e^a \wedge e^b\right) \\ & + \left(\delta_0^g \zeta^{\alpha ab} + \rhd_{\gamma\beta}{}^{\alpha} \zeta^{\gamma ab} \epsilon_{\mathfrak{g}}{}^{\beta}\right) \left(M_{\alpha ab} \varepsilon_{cdef} e^c \wedge e^d \wedge e^e \wedge e^f - F_{\alpha} \wedge e_a \wedge e_b\right) \\ & + \lambda^{\alpha} \wedge \left(-\nabla \epsilon_{\mathfrak{m}\alpha} - \epsilon_{\mathfrak{l}}{}^{A} \rhd_{\alpha A}{}^{B} D_B + \epsilon_{\mathfrak{n}}{}^{a} \rhd_{\alpha a}{}^{b} \beta_b - \epsilon_{\mathfrak{h}}{}^{a} \wedge e_b \rhd_{\alpha a}{}^{b} + 24 C^{\alpha\beta} M_{\beta ab} (\nabla \epsilon_{\mathfrak{n}}{}^{a}) \wedge e^b\right) \\ & + \zeta^{\alpha ab} \left(-4 M_{\alpha ab} \varepsilon_{cdef} (\nabla \epsilon_{\mathfrak{n}}{}^{c}) \wedge e^d \wedge e^e \wedge e^f + 2 F_{\alpha} \wedge (\nabla \epsilon_{\mathfrak{n} a}) \wedge e_b\right) .\end{aligned} \tag{59}$$

Similar to the case of gravitational constraint, all symmetries, $\tilde{H}$, $\tilde{L}$, $\tilde{N}$ and $\tilde{M}$, except for $\tilde{G}$, are broken. The variations for new multipliers are chosen as:

$$\delta_0^g \lambda^{\alpha} = \epsilon_{\mathfrak{g}}{}^{\beta} \lambda^{\gamma} \rhd_{\beta\gamma}{}^{\alpha} \,, \qquad \delta_0^g M_{\alpha ab} = \epsilon_{\mathfrak{g}}{}^{\beta} M^{\gamma}{}_{ab} \rhd_{\alpha\beta\gamma} \,, \qquad \delta_0^g \zeta^{\alpha ab} = \epsilon_{\mathfrak{g}}{}^{\beta} \zeta^{\gamma ab} \rhd_{\beta\gamma}{}^{\alpha} \,. \tag{60}$$

3.5. Higgs, Yukawa, spin, and CC terms

Variation of Higgs term gives us

$$\delta_0^g S_{\text{Higgs}} = -\frac{1}{3}\chi \int 2\left(\phi_A\phi^A - v^2\right)\phi_A\left(\delta_0^g\phi^A\right)\varepsilon^{abcd}e_a\wedge e_b\wedge e_c\wedge e_d + \left(\phi_A\phi^A - v^2\right)^2\varepsilon^{abcd}\left(\delta_0^g e_a\right)\wedge e_b\wedge e_c\wedge e_d\,. \tag{61}$$

This implies that

$$\delta_0^g S_{\text{Higgs}} = \frac{1}{3}\chi\int\left(\phi_A\phi^A - v^2\right)^2\varepsilon^{abcd}\left(\nabla\epsilon_{\mathfrak{n}a}\right)\wedge e_b\wedge e_c\wedge e_d\,, \tag{62}$$

where we have used the identity $\triangleright_{\alpha AB} = -\triangleright_{\alpha BA}$ from the theorem. This constraint does not break $\tilde{G}$ symmetry since its parameter drops out of (62) despite being present in the form variations for both ϕ^A and e_a, so only $\tilde{N}$ symmetry is broken.

The variation of Yukawa coupling term is:

$$\delta_0^g S_{\text{Yukawa}} = -\frac{2}{4!}\int Y_{ABC}\delta_0^g\left(\bar{\psi}^A\psi^B\right)\phi^C\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\wedge e^d + Y_{ABC}\bar{\psi}^A\psi^B\delta_0^g\phi^C\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\wedge e^d + 4Y_{ABC}\bar{\psi}^A\psi^B\phi^C\varepsilon_{abcd}\delta_0^g e^a\wedge e^b\wedge e^c\wedge e^d\,. \tag{63}$$

Again, substituting the variations of fields (43) into (63) gives:

$$\delta_0^g S_{\text{Yukawa}} = \frac{1}{3}\int Y_{ABC}\bar{\psi}^A\psi^B\phi^C\varepsilon_{abcd}\left(\nabla\epsilon_{\mathfrak{n}}{}^a\right)\wedge e^b\wedge e^c\wedge e^d\,. \tag{64}$$

As above, only $\tilde{N}$ symmetry is broken, because the Y_{ABC} matrix is defined in such a way to preserve $\tilde{G}$ symmetry.

Spin coupling term does not break $\tilde{G}$ symmetry for the same reasons as the Dirac or Higgs terms, but it does break $\tilde{N}$ and $\tilde{H}$ symmetries:

$$\begin{aligned}\delta_0^g S_{\text{spin}} &= 2\pi\,\mathrm{i}l_p^2\varepsilon_{abcd}\int\left(\delta_0^g\left(\bar{\psi}\gamma_5\gamma^a\psi\right)e^b\wedge e^c\wedge\beta^d + \bar{\psi}\gamma_5\gamma^a\psi\delta_0^g\left(e^b\wedge e^c\wedge\beta^d\right)\right),\\ &= -2\pi\,\mathrm{i}l_p^2\varepsilon_{abcd}\int\bar{\psi}\gamma_5\gamma^a\psi\left(2\left(\nabla\epsilon_{\mathfrak{n}}{}^b\right)\wedge e^c\wedge\beta^d + e^b\wedge e^c\wedge\left(\nabla\epsilon_{\mathfrak{h}}{}^d\right)\right).\end{aligned} \tag{65}$$

Finally, the CC term breaks only $\tilde{N}$ symmetry:

$$\delta_0^g S_{\text{CC}} = -\int\frac{1}{24\pi l_p^2}\Lambda\varepsilon_{abcd}\delta_0^g e^a\wedge e^b\wedge e^c\wedge e^d = -\int\frac{1}{24\pi l_p^2}\Lambda\varepsilon_{abcd}\nabla\epsilon_{\mathfrak{n}}{}^a\wedge e^b\wedge e^c\wedge e^d\,. \tag{66}$$

3.6. Overview of symmetry breaking

Summing up all the results from this section, we can make a table of symmetries and constraints. Each field labeled with $\times$ corresponds to the breaking of a given symmetry by a given constraint term:

	S_{grav}	S_{scal}	S_{Dirac}	$S_{\text{Yang–Mills}}$	S_{Higgs}	S_{Yukawa}	S_{spin}	S_{CC}
$\tilde{G}$								
$\tilde{H}$	$\times$			$\times$			$\times$	
$\tilde{L}$	$\times$	$\times$	$\times$	$\times$				
$\tilde{M}$	$\times$			$\times$				
$\tilde{N}$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$

From the above table one can observe several interesting features. First, $\tilde{G}$ symmetry is preserved by all constraints, while $\tilde{N}$ symmetry is broken by all constraints. Second, the gravitational and Yang–Mills constraints break all symmetries except $\tilde{G}$, and these constraints are the only ones to do so. Finally, the Higgs potential constraint, the Yukawa coupling constraint and the cosmological constant constraint break exclusively the $\tilde{N}$ symmetry, while preserving all others.

In addition to the above results, there are three more constraint terms, which do not appear in the action (19) but will appear later on in sections 4 and 5, after we rewrite the action in a form corresponding to the spontaneously broken symmetry. The first of these is the mass term for scalar fields:

$$S_{\text{scalar mass}} = -\frac{m^2}{4!}\varepsilon_{abcd}\int \phi_A\phi^A e^a\wedge e^b\wedge e^c\wedge e^d\,. \tag{67}$$

Variation of this term is

$$\delta_0^g S_{\text{scalar mass}} = -\frac{m^2}{4!}\varepsilon_{abcd}\int \left(2\left(\delta_0^g\phi_A\right)\phi^A e^a\wedge e^b\wedge e^c\wedge e^d + 4\phi_A\phi^A\left(\delta_0^g e^a\right)\wedge e^b\wedge e^c\wedge e^d\right)\,, \tag{68}$$

which reduces to:

$$\delta_0^g S_{\text{scalar mass}} = \frac{m^2}{3!}\varepsilon_{abcd}\int \phi_A\phi^A\left(\nabla\epsilon_{\mathfrak{n}}{}^a\right)\wedge e^b\wedge e^c\wedge e^d\,. \tag{69}$$

We conclude that the term (67) breaks only $\tilde{N}$ symmetry.

The second new constraint term is the Dirac mass term:

$$S_{\text{Dirac mass}} = -\frac{m}{12}\varepsilon_{abcd}\int \bar{\psi}_A\psi^A e^a\wedge e^b\wedge e^c\wedge e^d\,. \tag{70}$$

Its variation is

$$\delta_0^g S_{\text{Dirac mass}} = -\frac{m}{12}\varepsilon_{abcd}\int \Big((\delta_0^g\bar{\psi}_A)\psi^A e^a\wedge e^b\wedge e^c\wedge e^d + \bar{\psi}_A(\delta_0^g\psi^A)e^a\wedge e^b\wedge e^c\wedge e^d + 4\bar{\psi}_A\psi^A(\delta_0^g e^a)\wedge e^b\wedge e^c\wedge e^d\Big)\,, \tag{71}$$

which reduces to:

$$\delta_0^g S_{\text{Dirac mass}} = \frac{m}{3}\varepsilon_{abcd}\int \bar{\psi}_A\psi^A \left(\nabla\epsilon_{\mathfrak{n}}{}^a\right)\wedge e^b\wedge e^c\wedge e^d\,. \tag{72}$$

Similar to the scalar mass term, this constraint also breaks only $\tilde{N}$ symmetry.

The third new constraint term is the Proca constraint. This term will explicitly appear in section 4 in equation (80) below, as part of the discussion of the Proca action. It has the following form (see section 4 for the details of the notation):

$$S_{\text{Proca}} = \int \Theta^{\alpha ab}\wedge\left(\Xi_{\alpha abc}\varepsilon^{cdef}e_d\wedge e_e\wedge e_f + \frac{M}{g}\alpha_\alpha\wedge e_a\wedge e_b\right) + \frac{M}{g}\alpha^\alpha\wedge\Xi_{\alpha abc}e^a\wedge e^b\wedge e^c\,. \tag{73}$$

The variation of this term is:

$$\begin{aligned}\delta_0^g S_{\text{Proca}} = \int \Bigg[&\frac{M}{g}\left(\delta_0^g\alpha^\alpha\wedge\Xi_{\alpha abc}e^a + \alpha^\alpha\wedge\delta_0^g\Xi_{\alpha abc}e^a + 3\alpha^\alpha\wedge\Xi_{\alpha abc}\delta_0^g e^a\right)\wedge e^b\wedge e^c\\
&+\delta_0^g\Theta^{\alpha ab}\wedge\left(\Xi_{\alpha abc}\varepsilon^{cdef}e_d\wedge e_e\wedge e_f + \frac{M}{g}\alpha_\alpha\wedge e_a\wedge e_b\right)\\
&+\Theta^{\alpha ab}\wedge\left(\delta_0^g\Xi_{\alpha abc}\varepsilon^{cdef}e_d\wedge e_d\wedge e_e\wedge e_f + 3\Xi_{\alpha abc}\varepsilon^{cdef}\delta_0^g e_d\wedge e_e\wedge e_f\right)\\
&+\frac{M}{g}\Theta^{\alpha ab}\wedge\left(\delta_0^g\alpha_\alpha\wedge e_a\wedge e_b + 2\alpha_\alpha\wedge\delta_0^g e_a\wedge e_b\right)\Bigg]\,.\end{aligned} \tag{74}$$

Substituting the variations of connection α and tetrad fields, and using the fact that $\triangleright_{\alpha a}{}^b = 0$, we get:

$$\begin{aligned}\delta_0^g S_{\text{Proca}} = \int \Bigg[&\delta_0^g\Theta^{\alpha ab}\wedge\left(\Xi_{\alpha abc}\varepsilon^{cdef}e_d\wedge e_e\wedge e_f + \frac{M}{g}\alpha_\alpha\wedge e_a\wedge e_b\right)\\
&+\delta_0^g\,\Xi_{\alpha abc}\left(\frac{M}{g}\alpha^\alpha\wedge e^a\wedge e^b\wedge e^c + \Theta^{\alpha ab}\wedge\varepsilon^{cdef}e_d\wedge e_e\wedge e_f\right)\\
&+\frac{M}{g}\nabla\epsilon_{\mathfrak{g}}{}^\alpha\wedge\left(\Theta^{\alpha ab}\wedge e_a\wedge e_b - \Xi_{\alpha abc}e^a\wedge e^b\wedge e^c\right)\\
&-3\,\Xi_{\alpha abc}\left(\frac{M}{g}\alpha^\alpha\wedge\nabla\epsilon_{\mathfrak{n}}{}^a\wedge e^b\wedge e^c + \Theta^{\alpha ab}\varepsilon^{cdef}\nabla\epsilon_{\mathfrak{n}d}\wedge e_e\wedge e_f\right)\\
&-2\frac{M}{g}\Theta^{\alpha ab}\wedge\alpha_\alpha\wedge\nabla\epsilon_{\mathfrak{n}a}\wedge e_b\Bigg]\,.\end{aligned} \tag{75}$$

We can eliminate the first two rows by choosing the variations of the new multipliers to be

$$\delta_0^g\Theta^{\alpha ab} = 0\,,\qquad \delta_0^g\,\Xi_{\alpha abc} = 0\,. \tag{76}$$

However, in addition to broken $\tilde{N}$ symmetry, the Proca constraint is the only constraint which breaks $\tilde{G}$ symmetry, since its parameter appears explicitly in the third row.

Finally, let us note that the scalar mass constraint, Dirac mass constraint and the Proca constraint supplement the above table of constraints with three more columns, as follows:

	$S_{\text{scalar mass}}$	$S_{\text{Dirac mass}}$	S_{Proca}
$\tilde{G}$			$\times$
$\tilde{H}$			
$\tilde{L}$			
$\tilde{M}$			
$\tilde{N}$	$\times$	$\times$	$\times$

This concludes the analysis of explicit symmetry breaking of the constrained $3BF$ theory. In what follows, we turn to the detailed analysis of the Proca action, and after that to the spontaneous symmetry breaking, which has completely different nature and properties from the explicit symmetry breaking.

4. Constrained 3*BF* action for the Proca field

In order to study the electroweak theory, spontaneous symmetry breaking and the Higgs mechanism within the framework of higher gauge theory, an important step is to give a review of the Proca action written as a constrained $3BF$ theory, since the Higgs mechanism will naturally generate mass terms for the vector bosons. The $3BF$ formulation of the Proca action will therefore help us recognize these terms when we turn to the details of the Higgs mechanism.

In order to introduce the constrained $3BF$ action for the Proca field, the first step is to specify the choice of a 3-group. The typical choice is the following. The three component Lie groups are given as:

$$G = SO(3,1) \times SU(N)\,, \qquad H = \mathbb{R}^4\,, \qquad L = \{\mathbb{1}_L\}\,. \tag{77}$$

This choice corresponds to the $SU(N)$ Yang–Mills field coupled to Einstein–Cartan gravity, with no scalar or fermion matter (since the group L is trivial). The trivial choice of L implies that the Peiffer lifting and the homomorphism δ are also trivial, as well as the action $\triangleright$ of the group G onto L. What remains to be specified is the homomorphism ∂ and the action $\triangleright$ of the group G onto itself and onto the group H. We choose the homomorphism ∂ to be trivial as well, while the action $\triangleright$ is specified as follows. The action of G onto itself is given via the equations (11) and (12), similar as for the SM, while the action of G onto H is also given via the equations (13).

In order to define the corresponding $3BF$ action, the symmetric nondegenerate invariant bilinear forms $\langle\_\,,\_\rangle_{\mathfrak{g}}$ and $\langle\_\,,\_\rangle_{\mathfrak{h}}$ are specified via the equations (15) and (16), respectively, while $\langle\_\,,\_\rangle_{\mathfrak{l}}$ is trivial. These choices simplify the $3BF$ action into a $2BF$ action, a special case of (20), given as:

$$S_{2BF} = \int B_\alpha \wedge F^\alpha + B^{[ab]} \wedge R_{[ab]} + e_a \wedge \nabla \beta^a\,. \tag{78}$$

Here the first term is the BF term for the $SU(N)$ group corresponding to the Yang–Mills part, while the remaining two terms correspond to the gravitational part.

Once we have specified the topological part of the action, we deform it by adding appropriate constraints. In order to obtain appropriate dynamics for gravity, we have to add the gravitational constraint term (21), while in order to obtain appropriate dynamics for the Yang–Mills field we similarly have to add the Yang–Mills constraint term (24), and in this case one additional constraint, called the Proca constraint term:

$$S = S_{2BF} + S_{\text{grav}} + S_{\text{Yang–Mills}} + S_{\text{Proca}}\,. \tag{79}$$

The new Proca constraint term has the following form

$$S_{\text{Proca}} = \int \Theta^{\alpha ab} \wedge \left(\Xi_{\alpha abc} \varepsilon^{cdef} e_d \wedge e_e \wedge e_f + \frac{M}{g} \alpha_\alpha \wedge e_a \wedge e_b \right) + \frac{M}{g} \alpha^\alpha \wedge \Xi_{\alpha abc} e^a \wedge e^b \wedge e^c \,, \tag{80}$$

where the 1-form $\Theta^{\alpha ab}$ and 0-form $\Xi_{\alpha abc}$ are new Lagrange multipliers, M is the new parameter, while g is the Yang–Mills coupling constant, corresponding to the choice of the coupling constant bilinear form in (24) as:

$$C_{\alpha\beta} = \frac{1}{g^2} g_{\alpha\beta} \,. \tag{81}$$

In order to demonstrate that the action (79) really corresponds to the theory of the Proca field, we compute the corresponding equations of motion. Similarly to the case of the SM, the variations of the action with respect to all fields will give the equations that can be solved for the multipliers,

$$\begin{aligned}
M_{\alpha ab} &= -\frac{1}{48} \varepsilon_{abcd} e^c{}_\mu e^d{}_\nu F_\alpha{}^{\mu\nu} \,, & \lambda_{\alpha\mu\nu} &= -F_{\alpha\mu\nu} \,, & \zeta^{\alpha ab} &= \frac{1}{4g^2} \varepsilon^{abcd} e_{c\mu} e_{d\nu} F^{\alpha\mu\nu} \,, \\
\Theta^{\alpha ab}{}_\mu &= \frac{M}{6g} \varepsilon^{abcd} \alpha^\alpha{}_\nu e_c{}^\nu e_{d\mu} \,, & \lambda_{[ab]\mu\nu} &= R_{[ab]\mu\nu} \,, & \Xi_{\alpha abc} &= \frac{M}{6g} \varepsilon_{abcd} \alpha_{\alpha\mu} e^{d\mu} \,, \\
B_{\alpha\mu\nu} &= -\frac{e}{2g^2} \varepsilon_{\mu\nu\rho\sigma} F_\alpha{}^{\rho\sigma} \,, & \beta^a{}_{\mu\nu} &= 0 \,, & B_{[ab]\mu\nu} &= \frac{1}{8\pi l_p^2} \varepsilon_{abcd} e^c{}_\mu e^d{}_\nu \,,
\end{aligned} \tag{82}$$

then the Einstein field equation (32) for the stress–energy tensor of the form

$$T_{\mu\nu} = \frac{1}{g^2} \left(F^\alpha{}_{\mu\rho} F_{\alpha\nu}{}^\rho - \frac{1}{4} g_{\mu\nu} F^\alpha{}_{\rho\sigma} F_\alpha{}^{\rho\sigma} \right) + \frac{M^2}{g^2} \left(\alpha_{\alpha\mu} \alpha^\alpha{}_\nu - \frac{1}{2} g_{\mu\nu} \alpha_\alpha{}^\rho \alpha^\alpha{}_\rho \right) , \tag{83}$$

as well as the equations for the spin connection (31) and the torsion equation $T^a \equiv \nabla e^a = 0$, and finally the equation of motion for the vector boson field

$$\nabla_\mu F^{\alpha\mu}{}_\nu - M^2 \alpha^\alpha{}_\nu = 0 \,, \tag{84}$$

where $F^\alpha{}_{\mu\nu}$ is the standard Yang–Mills field strength tensor for the $SU(N)$ connection $\alpha^\alpha{}_\mu$. This is precisely the Proca equation for the field with mass M.

In addition to the equations of motion, one can verify that the action (79) corresponds to the Proca theory by eliminating all auxiliary fields. Since auxiliary fields are algebraically determined as functions of the dynamical fields, their equations of motion can be substituted back into the action, leading to the second-order formulation of the theory. In particular, substituting all equations (82) into (79), after a certain amount of straightforward algebra, one obtains precisely the traditional formulation of the action for the Proca field coupled to Einstein–Cartan gravity:

$$S = \int \frac{1}{16\pi l_p^2} \varepsilon^{abcd} R_{ab} \wedge e_c \wedge e_d - \frac{1}{g^2} F_\alpha \wedge \star F^\alpha - \frac{1}{4!} \frac{M^2}{g^2} \alpha^\alpha{}_\mu \alpha_\alpha{}^\mu \varepsilon^{abcd} e_a \wedge e_b \wedge e_c \wedge e_d \,. \tag{85}$$

Here $\alpha_\alpha{}^\mu \equiv \alpha_{\alpha\nu} g^{\mu\nu}$ where $g^{\mu\nu} = \eta^{ab} e_a{}^\mu e_b{}^\nu$. Also, $\star F$ denotes the Hodge dual of the 2-form F:

$$\star F^\alpha = \frac{1}{4} F^\alpha{}_{cd} \varepsilon^{abcd} e_a \wedge e_b \,. \tag{86}$$

When dealing with the electroweak theory and the SM, one encounters multiple Proca fields, with different masses M. In order to account for this, let us generalize the action (79) to the case of multiple Proca fields. This is done by choosing a 3-group with a modified group G of the form:

$$G = SO(3,1) \times \prod_i U(1) \times \prod_j SU(N_j) \,. \tag{87}$$

Compared to (77), one can see that the subgroup $SU(N)$ in G has been substituted with multiple copies of $U(1)$ and $SU(N_j)$, depending on how many Proca fields we wish to have in the theory. The structure of the 3-group remains essentially the same, in the sense that the action $\triangleright$ is extended from the $SU(N)$ case to the more general case in an obvious way, so that equations (11)–(13) remain valid for the general choice (87).

Given this more general choice of the 3-group, the action for the theory formally still has the form (79), but now the terms S_{2BF} and $S_{\text{Yang–Mills}}$ correspond to the new choice of the internal gauge group, and the coupling constant bilinear form $C_{\alpha\beta}$ does not need to have the form (81) anymore, but instead it may depend on multiple coupling constants g_i, one for each term in the products in (87). The only requirements on $C_{\alpha\beta}$ are that it must be symmetric, nondegenerate and G-invariant, since its eigenvalues should be $1/g_i^2$. Finally, the term S_{Proca} becomes more complicated, and has the following form:

$$\begin{aligned} S_{\text{Proca}} = \int \Theta^{\alpha ab} \wedge \left(\Xi_{\alpha abc} \varepsilon^{cdef} e_d \wedge e_e \wedge e_f + N_{\alpha\beta}\, \alpha^\beta \wedge e_a \wedge e_b \right) \\ + \alpha^\alpha \wedge \tilde{N}_\alpha{}^\beta \Xi_{\beta abc} e^a \wedge e^b \wedge e^c \,, \end{aligned} \tag{88}$$

This constraint term features a new bilinear form $N_{\alpha\beta}$ and a new matrix $\tilde{N}_\alpha{}^\beta$, which are constant and arbitrary, representing new free parameters of the action. In order to understand their physical meaning, let us discuss the equations of motion for the action, as follows. First, the equations that can be solved for the multipliers are

$$\begin{aligned} &M_{\alpha ab} = -\frac{1}{48}\varepsilon_{abcd} e^c{}_\mu e^d{}_\nu F_\alpha{}^{\mu\nu} \,, \qquad \lambda_{\alpha\mu\nu} = -F_{\alpha\mu\nu} \,, \qquad \zeta_\alpha{}^{ab} = \frac{1}{4} C_{\alpha\beta} \varepsilon^{abcd} e_{c\mu} e_{d\nu} F^{\beta\mu\nu} \,, \\ &\Theta^{\alpha ab}{}_\mu = \frac{1}{6} \tilde{N}_\beta{}^\alpha \varepsilon^{abcd} \alpha^\beta{}_\nu e_c{}^\nu e_{d\mu} \,, \qquad \lambda_{[ab]\mu\nu} = R_{[ab]\mu\nu} \,, \qquad \Xi_{\alpha abc} = \frac{1}{6} N_{\alpha\beta} \varepsilon_{abcd} \alpha^\beta{}_\mu e^{d\mu} \,, \\ &B_{\alpha\mu\nu} = -\frac{e}{2} C_{\alpha\beta} \varepsilon_{\mu\nu\rho\sigma} F^{\beta\rho\sigma} \,, \qquad \beta^a{}_{\mu\nu} = 0 \,, \qquad B_{[ab]\mu\nu} = \frac{1}{8\pi l_p^2} \varepsilon_{abcd} e^c{}_\mu e^d{}_\nu \,, \end{aligned} \tag{89}$$

where we can see the presence of the parameters $C_{\alpha\beta}, N_{\alpha\beta}$ and $\tilde{N}_\alpha{}^\beta$. Next, the torsion equation $\nabla e^a = 0$ remains unchanged, while the equation of motion and the stress–energy tensor for the vector fields obtain the following form:

$$\nabla_\mu F^{\alpha\mu}{}_\nu - M^\alpha{}_\beta \alpha^\beta{}_\nu = 0 \,, \tag{90}$$

$$T_{\mu\nu} = C_{\alpha\beta} \left(F^\alpha{}_{\mu\rho} F^\beta{}_\nu{}^\rho - \frac{1}{4} g_{\mu\nu} F^\alpha{}_{\rho\sigma} F^{\beta\rho\sigma} \right) + C_{\alpha\beta} M^\beta{}_\gamma \left(\alpha^\alpha{}_\mu \alpha^\gamma{}_\nu - \frac{1}{2} g_{\mu\nu} \alpha^{\alpha\rho} \alpha^\gamma{}_\rho \right) . \tag{91}$$

Here the new matrix $M^\alpha{}_\beta$ is constructed from $C_{\alpha\beta}, N_{\alpha\beta}$ and $\tilde{N}_\alpha{}^\beta$ as follows:

$$M^\alpha{}_\beta = \frac{1}{2} \left(C^{-1} \right)^{\alpha\gamma} \left(\tilde{N}_\gamma{}^\delta N_{\delta\beta} + \tilde{N}_\beta{}^\delta N_{\delta\gamma} \right) . \tag{92}$$

This matrix is interpreted as the squared-mass matrix of the theory. Note that due to the fact that $C_{\alpha\beta}$ is nondegenerate, it is also invertible. In order to interpret $M^{\alpha}{}_{\beta}$ as a matrix whose eigenvalues are squares of masses, the parameters $C_{\alpha\beta}, N_{\alpha\beta}$ and $\tilde{N}_{\alpha}{}^{\beta}$ have to be chosen so that (92) is positive semi-definite. In such a case, choosing a basis in Lie algebra $\mathfrak{g}$ as an eigenbasis of $M^{\alpha}{}_{\beta}$, and denoting the respective eigenvalues as $M^2_{(\alpha)}$, one can rewrite the squared-mass matrix into the form

$$M^{\alpha}{}_{\beta} = M^2_{(\alpha)} \delta^{\alpha}_{\beta} \,, \tag{93}$$

where the parentheses over the index α denote that this index is not summed over. Substituting this into the equation of motion (90) we finally obtain

$$\nabla_{\mu} F^{\alpha\mu}{}_{\nu} - M^2_{(\alpha)} \alpha^{\alpha}{}_{\nu} = 0 \,. \tag{94}$$

This is a set of equations of motion for several Proca fields, with (possibly different) masses $M_{(\alpha)}$, which explains why we can interpret $M^{\alpha}{}_{\beta}$ as the squared-mass matrix. Also, note that the obtained equation of motion (94) and the stress–energy tensor (91) are natural generalizations of their single Proca field counterparts (84) and (83), respectively. Moreover, similarly to the case of a single Proca field, one can substitute the algebraic equations of motion for the auxiliary fields (89) back into the action (79) with (88) to obtain the traditional second-order formulation of the Proca theory coupled to Einstein–Cartan gravity.

In order to be able to successfully compare, term by term, the Proca action with the action that will be obtained in section 5 as a result of the Higgs mechanism, there is one more generalization that we need to do. In particular, we modify the Proca constraint (88) by introducing two additional Lagrange multipliers, a 1-form θ^{α} and a 3-form ρ_{α}, as follows:

$$\begin{aligned} S_{\text{Proca}} = \int \Theta^{\alpha ab} \wedge \left(\Xi_{\alpha abc} \varepsilon^{cdef} e_d \wedge e_e \wedge e_f + N_{\alpha\beta}\, \alpha^{\beta} \wedge e_a \wedge e_b \right) \\ + \alpha^{\alpha} \wedge \tilde{N}_{\alpha}{}^{\beta} \rho_{\beta} + \theta^{\alpha} \wedge \left(\rho_{\alpha} - \Xi_{\alpha abc} e^a \wedge e^b \wedge e^c \right) . \end{aligned} \tag{95}$$

The two additional Lagrange multipliers provide a convenient extension of the configuration space, so that it is compatible with the configuration space that will naturally appear in section 5. Other than that, the multipliers do not modify any other property of the Proca action. In particular, the equations of motion (89) and (94), as well as the stress–energy tensor (91) and the torsion equation remain unchanged. Of course, extending the configuration space also means that we have two additional equations of motion, for the two new multipliers:

$$\theta^{\alpha} = -\tilde{N}_{\beta}{}^{\alpha} \alpha^{\beta}{}_{\mu} \,, \qquad \rho_{\alpha\nu\rho\sigma} = e M_{\alpha\beta}\, \varepsilon_{\mu\nu\rho\sigma} \alpha^{\beta\mu} \,. \tag{96}$$

As before, these two equations can also be readily substituted back into the action in order to obtain the traditional second-order formulation of the Proca action.

This concludes our analysis of the higher gauge theory reformulation of the Proca action. The form of the terms in the Proca constraint (95) are precisely the type of terms one should look for in the SM action after spontaneous symmetry breaking. As we shall see below, these kind of terms will be found precisely for the $W^{\pm}$ and Z^0 bosons in the electroweak theory.

5. Spontaneous symmetry breaking and the Higgs mechanism

The traditional formulation of the action for the SM of elementary particles does not involve the $3BF$ action and simplicity constraint terms, but is rather expressed in the ordinary tensor form of the Lagrangian. One then performs a sequence of steps, comprising the Higgs mechanism, in order to rewrite the Lagrangian in the form where the full gauge symmetry is not manifest. The additional assumption that the vacuum state is not invariant with respect to the full gauge symmetry, but only one of its subgroups, and the corresponding gauge fixing of the Lagrangian, renders the gauge symmetry of the theory spontaneously broken.

In light of the framework of $3BF$ theory with constraints described in section 2, it is natural to ask whether the Higgs mechanism can be applied to the action (19) which represents the SM expressed in this new language. Answering that question is the topic of this section.

5.1. Constrained 3BF action for the electroweak theory

In order to demonstrate the Higgs mechanism in the simplest way possible, let us restrict the action (19) to the electroweak sector, and for the moment ignore the fermion spectrum. In other words, we choose the 3-group in the following way:

$$G = SO(3,1) \times SU(2) \times U(1)\,, \qquad H = \mathbb{R}^4\,, \qquad L = \mathbb{C}^4\,. \tag{97}$$

The group G features the Lorentz subgroup, the weak isospin $SU(2)$ subgroup, and the weak hypercharge $U(1)$ subgroup. The group H remains the same as before, describing translations, while the group L has been reduced to describe only the doublet of complex scalar fields. The homomorphisms δ and ∂ remain trivial, as well as the Peiffer lifting $\{\_\,,\_\}_{\rm pf}$. Finally, the action of the group G is defined as follows. It acts on itself via conjugation, the Lorentz part acts in the standard way onto the group H, and trivially onto the group L, thereby defining that all component fields from L are scalar fields. The weak isospin and hypercharge act trivially on H, while they act in a nontrivial way on L. In order to explicitly state this action, it is useful to introduce the matrix notation for the generators T_A of the group L, in an obvious way, as:

$$T_1 = \begin{pmatrix} 1\\0\\0\\0 \end{pmatrix}, \qquad T_2 = \begin{pmatrix} 0\\1\\0\\0 \end{pmatrix}, \qquad T_3 = \begin{pmatrix} 0\\0\\1\\0 \end{pmatrix}, \qquad T_4 = \begin{pmatrix} 0\\0\\0\\1 \end{pmatrix}. \tag{98}$$

Then, if we denote the generators of weak isospin as τ_i ($i = 1,2,3$), and the generator of hypercharge as τ_0, we have:

$$\tau_\alpha \triangleright T_A = \triangleright_{\alpha A}{}^B T_B\,, \tag{99}$$

where the index α takes values $0,\ldots,3$, and thus conveniently counts all four generators (τ_0,τ_i) of the group $SU(2)\times U(1)$. The coefficients are explicitly given as:

$$\triangleright_{0A}{}^B = \frac{\mathrm{i}}{2}\begin{pmatrix} 1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1 \end{pmatrix}, \qquad \triangleright_{1A}{}^B = \frac{\mathrm{i}}{2}\begin{pmatrix} 0&1&0&0\\1&0&0&0\\0&0&0&-1\\0&0&-1&0 \end{pmatrix},$$
$$\triangleright_{2A}{}^B = \frac{\mathrm{i}}{2}\begin{pmatrix} 0&\mathrm{i}&0&0\\-\mathrm{i}&0&0&0\\0&0&0&\mathrm{i}\\0&0&-\mathrm{i}&0 \end{pmatrix}, \qquad \triangleright_{3A}{}^B = \frac{\mathrm{i}}{2}\begin{pmatrix} 1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&1 \end{pmatrix}. \tag{100}$$

Note also that the generators of $SU(2)\times U(1)$ satisfy the usual commutation relations

$$f_{\alpha\beta\gamma}=\begin{cases}-\varepsilon_{\alpha\beta\gamma}\,, & \text{for } \alpha,\beta,\gamma\neq 0\,,\\ 0 & \text{otherwise.}\end{cases} \tag{101}$$

This fixes the choice of the electroweak 3-group. Next, the bilinear forms are defined in the natural way—for the groups G and H they are defined as in (15) and (16), while for the group L the choice may appear unusual:

$$g_{AB}=\frac{1}{2}\begin{pmatrix}0&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&0\end{pmatrix}. \tag{102}$$

This is only apparent, since we wish to represent an element of the algebra $\mathfrak{l}$ in the form

$$\phi\equiv\phi^A T_A\equiv\phi_+T_1+\phi_0T_2+\phi_+^\dagger T_3+\phi_0^\dagger T_4=\begin{pmatrix}\phi_+\\ \phi_0\\ \phi_+^\dagger\\ \phi_0^\dagger\end{pmatrix}. \tag{103}$$

However, if one switches to a new basis in $\mathfrak{l}$ as

$$\tilde{T}_1=T_1+T_3\,,\qquad \tilde{T}_2=\mathrm{i}T_1-\mathrm{i}T_3\,,\qquad \tilde{T}_3=T_2+T_4\,,\qquad \tilde{T}_4=\mathrm{i}T_2-\mathrm{i}T_4\,, \tag{104}$$

the same algebra element can be rewritten as

$$\phi=\phi_1\tilde{T}_1+\phi_2\tilde{T}_2+\phi_3\tilde{T}_3+\phi_4\tilde{T}_4\,, \tag{105}$$

where $\phi_1,\dots,\phi_4$ are real-valued components, and there is a natural correspondence between the coefficients:

$$\phi_+=\phi_1+\mathrm{i}\phi_2\,,\qquad \phi_0=\phi_3+\mathrm{i}\phi_4\,,\qquad \phi_+^\dagger=\phi_1-\mathrm{i}\phi_2\,,\qquad \phi_0^\dagger=\phi_3-\mathrm{i}\phi_4\,. \tag{106}$$

In the basis $\tilde{T}_A$ the bilinear form (102) becomes the unit diagonal matrix. The basis $\tilde{T}_A$ is convenient because of the diagonal bilinear form and the real-valued components, while the basis T_A is convenient because it is an eigenbasis for the weak isospin and weak hypercharge opeators (and as we shall see, also for the electromagnetic charge operator). We will be frequently switching between these two bases throughout this section.

We should also note that (105) can be understood as an element of the four-dimensional real-valued Lie algebra $L=\mathbb{R}^4$, or equivalently of the two-dimensional complex-valued Lie algebra $L=\mathbb{C}^2$ (which is implicitly being used in most standard textbooks dealing with the Higgs mechanism). On the other hand, (103) is an element of the four-dimensional complex-valued Lie algebra $L=\mathbb{C}^4$, which is a complexification of $\mathbb{R}^4$, and if we wish to be able to seamlessly switch from the basis T_A to $\tilde{T}_A$ and back, it is far more convenient to work with the complexified algebra. Hence the choice $L=\mathbb{C}^4$ in the electroweak 3-group (97).

Once we have specified the choice of the 3-group and the choices for the bilinear forms, the action for the electroweak theory can be written as:

$$S=S_{3BF}+S_{\text{grav}}+S_{\text{scal}}+S_{\text{Yang–Mills}}+S_{\text{Higgs}}+S_{\text{CC}}\,. \tag{107}$$

It is similar in form to (19), where the constraint terms related to fermions have been ommited. The coupling constant bilinear form in $S_{\text{Yang–Mills}}$ is given as

$$C_{\alpha\beta} = \begin{pmatrix} \frac{1}{g_0^2} & 0 & 0 & 0 \\ 0 & \frac{1}{g_1^2} & 0 & 0 \\ 0 & 0 & \frac{1}{g_1^2} & 0 \\ 0 & 0 & 0 & \frac{1}{g_1^2} \end{pmatrix}, \tag{108}$$

reflecting the structure of the $SU(2) \times U(1)$ group.

5.2. Overview of the Higgs mechanism

There are three main steps in the Higgs mechanism:

- discussion of the stable vacuum,
- introduction of a change of variables,
- gauge fixing of the scalar fields.

In order to understand the details of the Higgs mechanism in the framework of the action (107), it is illustrative to repeat these main steps using the new variables and notation.

The analysis of the stable vacuum is essentially identical to the usual case of the Higgs mechanism. The S_{Higgs} constraint introduces the following potential for the scalar field,

$$V(\phi) = 2\chi \left(\phi^A \phi_A - v^2\right)^2 , \tag{109}$$

and one can observe that the stable vacuum is not unique, but is represented by a 3-sphere of points $\phi^A \phi_A = v^2$ in the configuation space. In order to rewrite the action in terms of fields that become equal to zero at some given vacuum point, one is led to introduce a change of variables from $(\phi_1, \phi_2, \phi_3, \phi_4)$ to $(\phi_1, \phi_2, h, \phi_4)$, where $h(x)$ is the new scalar field, obrained by translating ϕ_3 by v:

$$\phi_3(x) = v + h(x) . \tag{110}$$

This corresponds to the point $(0, 0, v, 0)$ on the 3-sphere as our vacuum of choice, by convention. Of course, this convention is completely arbitrary, and nothing in the rest of the analysis depends on this choice. The change of variables is given in terms of the basis $\tilde{T}_A$, while in terms of our original basis T_A we have:

$$\phi^A = \begin{pmatrix} \phi_+ \\ \phi_0 \\ \phi_+^\dagger \\ \phi_0^\dagger \end{pmatrix} = \begin{pmatrix} \phi_1 + \mathrm{i}\phi_2 \\ v + h + \mathrm{i}\phi_4 \\ \phi_1 - \mathrm{i}\phi_2 \\ v + h - \mathrm{i}\phi_4 \end{pmatrix} . \tag{111}$$

Finally, given this relation, one can observe that the components ϕ_1, ϕ_2 and ϕ_4 are in fact equivalent (up to linear order) to three gauge parameters of the $\mathfrak{g}$-gauge transformation

$$\phi \to \phi' = \mathrm{e}^{\xi^\alpha \tau_\alpha} \triangleright \phi . \tag{112}$$

Namely, using the action (99) of the generators of algebra $\mathfrak{g}$ on the generators of algebra $\mathfrak{l}$, one can start from the following state and the choice of the following gauge parameters,

$$\phi^A = \begin{pmatrix} 0 \\ v+h \\ 0 \\ v+h \end{pmatrix}, \qquad \xi^\alpha\left(\phi^A\right) = \frac{1}{v}\begin{pmatrix} \phi_4 \\ 2\phi_2 \\ 2\phi_1 \\ -\phi_4 \end{pmatrix} + \mathcal{O}\left(\phi^2\right), \tag{113}$$

and evaluate gauge transformation (112) on the above state to obtain:

$$\phi'^A = \mathrm{e}^{\xi^\alpha(\phi)\tau_\alpha}\begin{pmatrix} 0 \\ v+h \\ 0 \\ v+h \end{pmatrix} = \begin{pmatrix} \phi_1 + \mathrm{i}\phi_2 \\ v+h+\mathrm{i}\phi_4 \\ \phi_1 - \mathrm{i}\phi_2 \\ v+h-\mathrm{i}\phi_4 \end{pmatrix}. \tag{114}$$

Therefore, we see that the fields ϕ_1, ϕ_2 and ϕ_4 can be understood as gauge degrees of freedom, given by the relation (113). One can conclude that only the field h is physical, since it cannot be removed by a $\mathfrak{g}$-gauge transformation.

Next, one can see that, even after the removal of ϕ_1, ϕ_2 and ϕ_4 by using a gauge transformation, the state ϕ^A in (113) still remains invariant with respect to a $U(1)$ subgroup of G. Denoting the generator of this stabilizer group as Q, one can easily see from (100) that the stabilizer requirement $Q \triangleright \phi = 0$ is satisfied for

$$Q = \tau_0 + \tau_3 . \tag{115}$$

This equation is known by the name Gell–Mann–Nishijima formula (for electroweak interactions). Phenomenologically, Q corresponds to the electromagnetic charge q, specifically q is an eigenvalue of the operator $-\mathrm{i}Q$, and the corresponding $U(1)$ stabilizer group is the gauge group of electrodynamics. From the stabilizer requirement one can observe that the Higgs field $h(x)$ has no electric charge, since it corresponds for the eigenvalue $q = 0$.

Let us note that the above results do not depend in any way on the choice of the vacuum point $(0,0,v,0)$ on the 3-sphere. One could have chosen any other point, in which case the only difference is that the solution of the stabilizer equation $Q \triangleright \phi = 0$ would be slighly more general:

$$Q = \tau_0 + \vec{\alpha}\cdot\vec{\tau}, \qquad \vec{\alpha} \in \mathbb{R}^3, \qquad \|\vec{\alpha}\|^2 = 1 . \tag{116}$$

Here $\vec{\tau}$ is understood as a triple (τ_1, τ_2, τ_3). In particular, the electromagnetic charge of the Higgs field would remain zero even in this case.

5.3. Transformation of the action

Let us now turn to the problem of the transformation of the action with respect to the gauge transformation of the scalar field that removes the components ϕ_1, ϕ_2 and ϕ_4,

$$\phi^A \to \left(\mathrm{e}^{-\xi} \triangleright \phi\right)^A = \begin{pmatrix} 0 \\ v+h \\ 0 \\ v+h \end{pmatrix}, \tag{117}$$

where $\xi \equiv \xi^\alpha \tau_\alpha$, and the parameters ξ^α are given in (113). In order to see what happens to the action (107), let us first note that the remaining variables that enter the action transform as follows. The transformation of the 3-connection variables $(\alpha, \omega, \beta, \tilde{\gamma})$ is given as:

$$\alpha' = \mathrm{e}^{-\xi}(\alpha + \mathrm{d})\mathrm{e}^{\xi}, \qquad \omega' = \omega, \qquad \beta' = \beta, \qquad \tilde{\gamma}' = \mathrm{e}^{-\xi} \triangleright \tilde{\gamma}. \tag{118}$$

The corresponding curvatures transform as:

$$F' = \mathrm{e}^{-\xi} F \mathrm{e}^{\xi}, \qquad R' = R, \qquad \mathcal{G}' = \mathcal{G}, \qquad \mathcal{H}' = \mathrm{e}^{-\xi} \triangleright \mathcal{H}. \tag{119}$$

The transformations of the Lagrange multipliers which appear in the topological sector of the action, namely B_α, $B_{[ab]}$, e_a and ϕ^A, are given as:

$$B'_\alpha = \left(\mathrm{e}^{-\xi} B \mathrm{e}^{\xi}\right)_\alpha, \qquad B'_{[ab]} = B_{[ab]}, \qquad e'_a = e_a, \tag{120}$$

while the transformation of ϕ^A is already spelled out in (117). Next, the $\mathfrak{g}$-valued Lagrange multipliers which appear in the constraint sector of the action, namely λ_α, $\lambda_{[ab]}$, $M_{\alpha ab}$ and $\zeta_{\alpha ab}$, transform as:

$$\lambda'_\alpha = \left(\mathrm{e}^{-\xi}\lambda \mathrm{e}^{\xi}\right)_\alpha, \qquad \lambda'_{[ab]} = \lambda_{[ab]}, \qquad M'_{\alpha ab} = \left(\mathrm{e}^{-\xi} M \mathrm{e}^{\xi}\right)_{\alpha ab}, \qquad \zeta'_{\alpha ab} = \left(\mathrm{e}^{-\xi}\zeta \mathrm{e}^{\xi}\right)_{\alpha ab}. \tag{121}$$

The $\mathfrak{l}$-valued Lagrange multipliers which appear in the constraint sector of the action, namely $\tilde{\lambda}_A$, Λ_{abA} and H_{abcA}, transform as:

$$\tilde{\lambda}'_A = \left(\mathrm{e}^{-\xi} \triangleright \tilde{\lambda}\right)_A, \qquad \Lambda'_{abA} = \left(\mathrm{e}^{-\xi} \triangleright \Lambda\right)_{abA}, \qquad H'_{abcA} = \left(\mathrm{e}^{-\xi} \triangleright H\right)_{abcA}. \tag{122}$$

Finally, the constraint part of the action also features the covariant derivative $\nabla\phi$, which transforms in a covariant way,

$$(\nabla\phi)' = \mathrm{e}^{-\xi} \triangleright (\nabla\phi), \tag{123}$$

as expected for a covariant derivative.

In addition to all of the above fields, the action also features the bilinear form of coupling constants, $C_{\alpha\beta}$, given by (108). One can observe that this bilinear form is in fact term-by-term proportional to the already introduced bilinear form $\langle \_, \_ \rangle_{\mathfrak{g}}$, as follows:

$$C_{\alpha\beta} = \mathcal{C}(\tau_\alpha, \tau_\beta) \equiv \frac{\delta^j_\alpha \delta^k_\beta}{g_1^2} \langle \tau_j, \tau_k \rangle_{\mathfrak{g}} + \frac{\delta^0_\alpha \delta^0_\beta}{g_0^2} \langle \tau_0, \tau_0 \rangle_{\mathfrak{g}}. \tag{124}$$

The two terms in the sum correspond to bilinear forms $\langle \_, \_ \rangle_{\mathfrak{su}(2)}$ and $\langle \_, \_ \rangle_{\mathfrak{u}(1)}$, respectively. Given that the gauge transformation can be represented in the form $\mathrm{e}^{-\xi^i \tau_i} \times \mathrm{e}^{-\xi^0 \tau_0}$, owing to the direct product structure in the group $SU(2) \times U(1)$, each term in the gauge transformation leaves the corresponding bilinear form invariant,

$$\langle \mathrm{e}^{-\xi^i \tau_i} \triangleright \tau_j, \mathrm{e}^{-\xi^i \tau_i} \triangleright \tau_k \rangle_{\mathfrak{su}(2)} = \langle \tau_j, \tau_k \rangle_{\mathfrak{su}(2)}, \qquad \langle \mathrm{e}^{-\xi^0 \tau_0} \triangleright \tau_0, \mathrm{e}^{-\xi^0 \tau_0} \triangleright \tau_0 \rangle_{\mathfrak{u}(1)} = \langle \tau_0, \tau_0 \rangle_{\mathfrak{u}(1)}, \tag{125}$$

as a consequence of the postulated G-invariance property of the bilinear form $\langle \_, \_ \rangle_{\mathfrak{g}}$. This renders the bilinear form of coupling constants gauge invariant:

$$C'_{\alpha\beta} = C_{\alpha\beta} \,. \tag{126}$$

At this point we are ready to discuss the transformation of the action with respect to (117). Namely, the action (107) is a functional of all fields mentioned above,

$$\alpha^{\alpha},\ \omega^{[ab]},\ \beta^{a},\ \tilde{\gamma}^{A},\ B_{\alpha},\ B_{[ab]},\ e_{a},\ \lambda_{\alpha},\ \lambda_{[ab]},\ M_{\alpha ab},\ \zeta_{\alpha ab},\ \tilde{\lambda}_{A},\ \Lambda_{abA},\ H_{abcA},\ \phi^{A}, \tag{127}$$

or in other words, the fields in the above list define a kinematical configuration space of our action. However, not every term in the action is a function of ϕ^A in particular. Therefore, when performing the gauge transformation (117), terms independent of ϕ^A will remain the same, while the terms dependent of ϕ^A will transform in a nontrivial way, reducing the full configuration space to a smaller one, defined by the fields

$$\alpha^{\alpha},\ \omega^{[ab]},\ \beta^{a},\ \tilde{\gamma}^{A},\ B_{\alpha},\ B_{[ab]},\ e_{a},\ \lambda_{\alpha},\ \lambda_{[ab]},\ M_{\alpha ab},\ \zeta_{\alpha ab},\ \tilde{\lambda}_{A},\ \Lambda_{abA},\ H_{abcA},\ h, \tag{128}$$

which differ from the original set in the replacement $(\phi^1,\phi^2,\phi^3,\phi^4) \to (0,0,v+h,0)$. The task is then to determine the form of the action $\tilde{S}$ which is defined on this reduced configuration space, schematically defined by the transformation:

$$S\left[\dots,\phi^A\right] \quad \xrightarrow{\ \mathrm{e}^{-\xi}\ } \quad \tilde{S}[\dots,h] \equiv S\left[\dots,\phi^A\right]\Bigg|_{\substack{\phi^1=\phi^2=\phi^4=0\\ \phi^3=v+h}} . \tag{129}$$

One can immediately observe that S_{grav}, $S_{\text{Yang–Mills}}$ and S_{CC} transform in a trivial way, since they do not depend on ϕ:

$$\begin{aligned} &S_{\text{grav}} \xrightarrow{\mathrm{e}^{-\xi}} \tilde{S}_{\text{grav}} = S_{\text{grav}}\,, \quad S_{\text{Yang–Mills}} \xrightarrow{\mathrm{e}^{-\xi}} \tilde{S}_{\text{Yang–Mills}} = S_{\text{Yang–Mills}}\,,\\ &S_{\text{CC}} \xrightarrow{\mathrm{e}^{-\xi}} \tilde{S}_{\text{CC}} = S_{\text{CC}}\,. \end{aligned} \tag{130}$$

Moreover, the $2BF$ part of S_{3BF} also transforms trivially, for the same reason.

On the other hand, the third term in S_{3BF}, as well as S_{scal} and S_{Higgs} require more attention. Let us discuss first the S_{Higgs} term. Specifically, we have that

$$S_{\text{Higgs}} = -\int \frac{1}{4!} V(\phi)\, \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d \,, \tag{131}$$

where under the transformation (129) the Higgs potential $V(\phi)$ (see (109)) becomes

$$V(\phi) \xrightarrow{\mathrm{e}^{-\xi}} V(h) \equiv 8v^2\chi h^2 + 8v\chi h^3 + 2\chi h^4 \,. \tag{132}$$

Therefore, we see that

$$S_{\text{Higgs}} \xrightarrow{\mathrm{e}^{-\xi}} \tilde{S}_{\text{Higgs}} = -\int \frac{1}{4!} V(h)\, \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d \,. \tag{133}$$

From the form of the quadratic term in the resulting potential (132) and the general form of the mass term for a single real scalar field (67), one can read off the value of the Higgs mass as:

$$m = 2v\sqrt{2\chi}\,. \tag{134}$$

Let us also note here that one could choose a potential which has a form alternative to (109), for example

$$V_{\text{alt}}(\phi) = 2\chi\left(\phi^A\phi_A\right)^2 - 4\chi v^2\phi^A\phi_A\,. \tag{135}$$

This potential differs from (109) by a constant term $2\chi v^4$, which would then combine with S_{CC} to give a different value of the cosmological constant. However, the potential (109) does not suffer from this problem, and in our case the CC term of the action remains the same before and after spontaneous symmetry breaking.

Next let us discuss the S_{3BF} term. Note that only the final term in S_{3BF} depends on ϕ, while the remainder does not and can be denoted as S_{2BF}. Then, using a suitable change of basis $T_A \to \tilde{T}_A$ in the Lie algebra $\mathfrak{l}$ (see (104)), with an additional notation for the indices $A \to (\bar{A}, H)$ where $\bar{A} \in \{1,2,4\}$ and $H \equiv 3$, we have

$$\begin{aligned}
S_{3BF} &= S_{2BF} + \int \phi^A \nabla\tilde{\gamma}_A \xrightarrow{\text{e}^{-\xi}} S_{2BF} + \int (v+h)\left((\nabla\tilde{\gamma})_0 + (\nabla\tilde{\gamma})_{0^\dagger}\right) \\
&= S_{2BF} + \int (v+h)\left(\text{d}\tilde{\gamma}_H + \rhd_\alpha{}^{\bar{A}}{}_H\,\alpha^\alpha \wedge \tilde{\gamma}_{\bar{A}}\right) \\
&= S_{2BF} + \int h\text{d}\tilde{\gamma}_H + v\text{d}\tilde{\gamma}_H + (v+h)\rhd_\alpha{}^{\bar{A}}{}_H\,\alpha^\alpha \wedge \tilde{\gamma}_{\bar{A}} \\
&= \tilde{S}_{3BF} + \int v\text{d}\tilde{\gamma}_H + (v+h)\rhd_\alpha{}^{\bar{A}}{}_H\,\alpha^\alpha \wedge \tilde{\gamma}_{\bar{A}}\,,
\end{aligned} \tag{136}$$

where the new action $\tilde{S}_{3BF}$ is defined as a functional over the reduced configuration space (128) as

$$\tilde{S}_{3BF} = S_{2BF} + \int h\text{d}\tilde{\gamma}_H\,. \tag{137}$$

In section 6 we shall discuss in detail its corresponding 3-group. Therefore, we conclude that

$$S_{3BF} \xrightarrow{\text{e}^{-\xi}} \tilde{S}_{3BF} + \int v\text{d}\tilde{\gamma}_H + (v+h)\rhd_\alpha{}^{\bar{A}}{}_H\,\alpha^\alpha \wedge \tilde{\gamma}_{\bar{A}}\,, \tag{138}$$

where the extra terms will later be grouped together with extra terms from other parts of the action and discussed in detail.

In equations (136)–(138) we have made use of the basis (104) in the Lie algebra $\mathfrak{l}$, so that we can introduce $\tilde{\gamma}_H \equiv \tilde{\gamma}_0 + \tilde{\gamma}_{0^\dagger}$. The action $\rhd$ was represented via the matrices (100) in the original basis T_A, while in this basis it is now broken into the following set of components:

$$\rhd_{\alpha H}{}^{\bar{A}}\,, \qquad \rhd_{\alpha\bar{A}}{}^{H}\,, \qquad \rhd_{\alpha\bar{A}}{}^{\bar{B}}\,, \qquad \rhd_{\alpha H}{}^{H}\,. \tag{139}$$

Since in this basis the bilinear form g_{AB} is diagonal, in fact $g_{AB} = \delta_{AB}$, a consequence of the theorem from section 2 is that all these components have vanishing diagonal elements, in particular $\rhd_{\alpha H}{}^{H} = 0$, which implies that $\nabla\tilde{\gamma}_H \equiv \text{d}\tilde{\gamma}_H$ and justifies the identification (137). Moreover, the components $\rhd_{\alpha\bar{A}}{}^{\bar{B}}$ drop out of equations (136)–(138) and do not appear anywhere. This leaves us with the remaining set of relevant components, which can be represented in matrix form as follows:

$$\rhd_{\alpha H}{}^{\bar{A}} = \frac{1}{2}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad \rhd_{\alpha\bar{A}}{}^{H} = \frac{1}{2}\begin{pmatrix} 0 & 0 & -1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{140}$$

Note that here α is the row index and $\bar{A}$ is the column index, while $H \equiv 3$ is constant.

Finally, let us discuss the S_{scal} term, which originally has the form

$$\begin{aligned} S_{\text{scal}} = \int \tilde{\lambda}^A \wedge \left(\tilde{\gamma}_A - H_{abcA} e^a \wedge e^b \wedge e^c\right) + \Lambda^{abA} \wedge H_{abcA} \varepsilon^{cdef} e_d \wedge e_e \wedge e_f \\ - \Lambda^{abA} \wedge (\nabla\phi)_A \wedge e_a \wedge e_b \,, \end{aligned} \tag{141}$$

and similarly to S_{3BF}, it also depends on ϕ^A only in the final term, while the remainder is independent of ϕ^A. Splitting the index A into $(\bar{A}, H)$, the constraint transforms into:

$$\begin{aligned} S_{\text{scal}} \xrightarrow{\mathrm{e}^{-\xi}} \int \tilde{\lambda}^H \wedge \left(\tilde{\gamma}_H - H_{abcH} e^a \wedge e^b \wedge e^c\right) + \Lambda^{abH} \wedge H_{abcH} \varepsilon^{cdef} e_d \wedge e_e \wedge e_f \\ - \Lambda^{abH} \wedge \mathrm{d}h \wedge e_a \wedge e_b + \tilde{\lambda}^{\bar{A}} \wedge \left(\tilde{\gamma}_{\bar{A}} - H_{abc\bar{A}} e^a \wedge e^b \wedge e^c\right) \\ + \Lambda^{ab\bar{A}} \wedge H_{abc\bar{A}} \varepsilon^{cdef} e_d \wedge e_e \wedge e_f - \Lambda^{ab\bar{A}} \wedge \alpha^\alpha \rhd_\alpha{}^H{}_{\bar{A}} (v+h) \wedge e_a \wedge e_b \,. \end{aligned} \tag{142}$$

Note that the terms in the first row on the right-hand side are precisely the terms that define the scalar constraint for a single real scalar field, as a functional over the reduced configuration space (128). Denoting those terms as $\tilde{S}_{\text{scal}}$, we conclude that

$$\begin{aligned} S_{\text{scal}} \xrightarrow{\mathrm{e}^{-\xi}} \tilde{S}_{\text{scal}} + \int \tilde{\lambda}^{\bar{A}} \wedge \left(\tilde{\gamma}_{\bar{A}} - H_{abc\bar{A}} e^a \wedge e^b \wedge \mathrm{e}^c\right) \\ + \Lambda^{ab\bar{A}} \wedge H_{abc\bar{A}} \varepsilon^{cdef} e_d \wedge e_e \wedge e_f \\ - \Lambda^{ab\bar{A}} \wedge \alpha^\alpha \rhd_\alpha{}^H{}_{\bar{A}} (v+h) \wedge e_a \wedge e_b \,, \end{aligned} \tag{143}$$

where we again have three extra terms which will be grouped together with the remainder of the action.

After we have discussed all parts of the action (107) in the context of the transformation (129), we can put all the pieces together, and compare the full action with the actions for the Proca and massive scalar fields. However, in order to make this comparison more transparent, it is useful to introduce yet some more notation. In particular, let us introduce a bilinear form $\kappa^{\alpha\beta}$ so that it satisfies the following identity:

$$\kappa^{\alpha\beta} \rhd_{\alpha H}{}^{\bar{A}} \rhd_{\beta\bar{B}}{}^{H} = -\frac{1}{4}\delta^{\bar{A}}_{\bar{B}} \,. \tag{144}$$

This bilinear form is not unique. Namely, since the matrices (140) are of rank 3, there exists a projector $P_\alpha{}^\beta$ which satisfies

$$P_\alpha{}^\beta P_\beta{}^\gamma = P_\alpha{}^\gamma \,, \qquad P_\alpha{}^\alpha = 3 \,, \qquad P_{\alpha\beta} = P_{\beta\alpha} \,, \qquad P_\alpha{}^\beta \rhd_{\beta H}{}^{\bar{A}} = \rhd_{\alpha H}{}^{\bar{A}} \,. \tag{145}$$

Note that a projector that satisfies (145) also satisfies the identity $P_\alpha{}^\beta \rhd_{\beta\bar{A}}{}^H = \rhd_{\alpha\bar{A}}{}^H$, since $\rhd_{\alpha\bar{A}}{}^H$ has the same components as $\rhd_{\alpha H}{}^{\bar{A}}$ up to an overall minus sign, see (140). Therefore, the bilinear form $\kappa^{\alpha\beta}$ is defined up to a term of the form

$$\kappa^{\alpha\beta} \to \kappa^{\alpha\beta} + \left[\delta_\gamma^{(\alpha} - P_\gamma{}^{(\alpha}\right] A^{\beta)\gamma}\,, \tag{146}$$

where $A^{\alpha\beta}$ is an arbitrary matrix, while the parentheses on the indices denote symmetrization. One can recognize that the term in the brackets is the orthogonal projector, which maps into the kernel of the matrices (140). This arbitrariness guarantees that the bilinear form $\kappa^{\alpha\beta}$ can be chosen to be invertible. The projector $P_\alpha{}^\beta$ can be explicitly evaluated using the definition (145) and the matrices (140), while one convenient choice of the bilinear form $\kappa^{\alpha\beta}$ can be evaluated from (144), so that they can be written in matrix form as follows:

$$P_\alpha{}^\beta = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}, \qquad \kappa^{\alpha\beta} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{147}$$

In addition to the projector and $\kappa^{\alpha\beta}$, we can now also introduce the following quantities:

$$\begin{aligned} \theta^\alpha &\equiv -2\kappa^{\alpha\beta} \rhd_\beta{}^H{}_{\bar{A}} \tilde{\lambda}^{\bar{A}}\,, \qquad & \Theta^{\alpha ab} &\equiv -2\kappa^{\alpha\beta} \rhd_\beta{}^H{}_{\bar{A}} \Lambda^{ab\bar{A}}\,, \\ \rho_\alpha &\equiv 2 \rhd_\alpha{}^{\bar{A}}{}_H \tilde{\gamma}_{\bar{A}}\,, & \Xi_{\alpha abc} &\equiv 2 \rhd_\alpha{}^{\bar{A}}{}_H H_{abc\bar{A}}\,. \end{aligned} \tag{148}$$

These new quantities satisfy four fundamental identities,

$$\begin{aligned} \theta^\alpha \wedge \rho_\alpha &= \tilde{\lambda}^{\bar{A}} \wedge \tilde{\gamma}_{\bar{A}}\,, \qquad & \theta^\alpha \Xi_{\alpha abc} &= \tilde{\lambda}^{\bar{A}} H_{abc\bar{A}}\,, \\ \Theta^{\alpha ab} \wedge \rho_\alpha &= \Lambda^{ab\bar{A}} \wedge \tilde{\gamma}_{\bar{A}}\,, & \Theta^{\alpha ab}\Xi_{\alpha cde} &= \Lambda^{ab\bar{A}} H_{cde\bar{A}}\,, \end{aligned} \tag{149}$$

which are a straightforward consequence of the identity (144). The purpose of introducing these quantities lies in the fact that they help us eliminate the $\bar{A}$ indices from equations. Note that in both the definitions (148) and the identities (149) the indices $\bar{A}$ are summed over on the right-hand sides, while they do not appear at all on the left-hand sides.

It is important to empasize that the arbitrariness of $\kappa^{\alpha\beta}$ in (146) introduces changes into the action. This is due to the fact that the change of variables (148) introduces additional variables which do not appear in the original action. The requirement that these additional variables are absent, i.e. that the left-hand sides of identities (149) have the same number of components as the corresponding right-hand sides, reduces the arbitrariness (146) of $\kappa^{\alpha\beta}$ to the following more specific form:

$$\kappa^{\alpha\beta} \to \kappa^{\alpha\beta} + \left[\delta_\gamma^\alpha - P_\gamma{}^\alpha\right] A^{\gamma\delta} \left[\delta_\delta^\beta - P_\delta{}^\beta\right]. \tag{150}$$

Note that, although this transformation still allows one to choose $\kappa^{\alpha\beta}$ to be invertible, the action in fact remains invariant with respect to (150), meaning that we can keep working with the same theory. See appendix B for a detailed analysis and proof.

Once all these new quantities and notation have been introduced, we can return to the analysis of the action. Given the transformation (129) of the action (107), one can apply the definitions (148) and the identities (149) to eliminate all indices $\bar{A}, \bar{B}$ and thus rewrite the transformed action $\tilde{S}$ so that it becomes a functional over the reduced configuration space (128). Putting together the results (130), (133), (138) and (143), we obtain the following form of the full action:

$$
\begin{aligned}
\tilde{S} = {} & S_{\text{grav}} + S_{\text{Yang–Mills}} + S_{\text{CC}} + \tilde{S}_{\text{Higgs}} + \tilde{S}_{3BF} + \tilde{S}_{\text{scal}} \\
& + \int \Theta^{\alpha ab} \wedge \left(\Xi_{\alpha abc}\, \varepsilon^{cdef} e_d \wedge e_e \wedge e_f + \frac{v}{2} \kappa^{-1}_{\alpha\beta} P_\gamma{}^\beta \alpha^\gamma \wedge e_a \wedge e_b \right) \\
& + \int \frac{v}{2} \alpha^\alpha P_\alpha{}^\beta \wedge \rho_\beta + \theta^\alpha \wedge \left(\rho_\alpha - \Xi_{\alpha abc} e^a \wedge e^b \wedge e^c \right) \\
& + \frac{1}{2} \int h\, \alpha^\alpha P_\alpha{}^\beta \wedge \left(\rho_\beta - \kappa^{-1}_{\beta\gamma} \Theta^{\gamma ab} \wedge e_a \wedge e_b \right) + v \int \mathrm{d}\tilde{\gamma}_H \, .
\end{aligned} \tag{151}
$$

This form of the action can now be finally compared to the Proca action (79). Specifically, the second and third row of (151) should be compared to the Proca constraint term in the form (95). The term-by-term comparison gives us the identification of the free parameters in the Proca action as follows:

$$
N_{\alpha\beta} = \frac{v}{2} \kappa^{-1}_{\alpha\gamma} P_\beta{}^\gamma \, , \qquad \tilde{N}_\alpha{}^\beta = \frac{v}{2} P_\alpha{}^\beta \, . \tag{152}
$$

Using these, we can construct the squared-mass matrix (92) to obtain

$$
M^\alpha{}_\beta = \frac{v^2}{4} \left(C^{-1} \right)^{\alpha\gamma} P_\gamma{}^\delta \kappa^{-1}_{\delta\epsilon} P_\beta{}^\epsilon \, , \tag{153}
$$

where the coupling constant matrix $C_{\alpha\beta}$ is specified in (108), while the projector and the bilinear form $\kappa^{\alpha\beta}$ are specified in (147). This gives us an explicit form for the squared-mass matrix as:

$$
M^\alpha{}_\beta = \frac{v^2}{4} \begin{pmatrix} g_0^2 & 0 & 0 & -g_0^2 \\ 0 & g_1^2 & 0 & 0 \\ 0 & 0 & g_1^2 & 0 \\ -g_1^2 & 0 & 0 & g_1^2 \end{pmatrix} . \tag{154}
$$

The physically relevant basis in the Lie algebra $\mathfrak{g}$ is the one in which the above squared-mass matrix is diagonal, and the corresponding eigenvalues are interpreted as squares of masses of gauge vector bosons in that basis. Therefore, we wish to explicitly obtain this basis. Given that the first and last column in (154) are proportional, the determinant of the squared-mass matrix is zero, meaning that at least one of its eigenvalues is zero. Moreover, the matrix is already in block-diagonal form, with g_1^2 being two 1-dimensional blocks, from which one can conclude that two eigenvalues are the same and are equal to $v^2 g_1^2/4$. Finally, from the trace of the matrix one can deduce the fourth eigenvalue, so that the whole set is given as:

$$
M_1^2 = 0 \, , \qquad M_2^2 = \frac{v^2}{4} g_1^2 \, , \qquad M_3^2 = \frac{v^2}{4} g_1^2 \, , \qquad M_4^2 = \frac{v^2}{4} \left(g_0^2 + g_1^2 \right) . \tag{155}
$$

The fact that the eigenvalues M_2^2 and M_3^2 are equal implies that the eigenbasis is not uniquely determined, and we need some additional input in order to fix it. A natural choice is the eigenbasis of the stabilizer Q, introduced in (115), since we want to interpret it as the electromagnetic charge, and the value of this charge should be well-defined for each physical state described by our preferred basis. One can reexpress the stabilizer in the matrix form $Q_\alpha{}^\beta$, defined by the action of Q onto the basis vector τ_α:

$$Q \triangleright \tau_\alpha = Q_\alpha{}^\beta \tau_\beta \,. \tag{156}$$

Using (115), one can easily evaluate the components of the matrix $Q_\alpha{}^\beta$ to be

$$Q_\alpha{}^\beta = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} . \tag{157}$$

This matrix has the eigenvalues $(0, \mathrm{i}, -\mathrm{i}, 0)$, so that the electromagnetic charge operator $-\mathrm{i}Q$ has the corresponding eigenvalues:

$$q_1 = 0\,, \qquad q_2 = +1\,, \qquad q_3 = -1\,, \qquad q_4 = 0\,. \tag{158}$$

The stabilizer also has two equal eigenvalues, so that its eigenbasis is not uniquely determined either. Nevertheless, the squared-mass matrix and the stabilizer matrix mutually commute and therefore share a joint eigenbasis, and this joint eigenbasis is uniquely determined. We can express the new basis in terms of the old basis as follows,

$$\tau_A = \tau_0 + \tau_3\,, \qquad \tau_+ = \frac{\tau_1 + \mathrm{i}\tau_2}{\sqrt{2}}\,, \qquad \tau_- = \frac{\tau_1 - \mathrm{i}\tau_2}{\sqrt{2}}\,, \qquad \tau_Z = -\frac{g_0^2}{g_0^2 + g_1^2}\tau_0 + \frac{g_1^2}{g_0^2 + g_1^2}\tau_3\,, \tag{159}$$

and we can express the components of the connection 1-form $\alpha = \alpha^\alpha{}_\mu \mathrm{d}x^\mu \otimes \tau_\alpha$ in the new basis as:

$$\begin{aligned} A_\mu &= \frac{g_1^2}{g_0^2 + g_1^2}\alpha^0{}_\mu + \frac{g_0^2}{g_0^2 + g_1^2}\alpha^3{}_\mu\,, & W_\mu^+ &= \frac{\alpha^1{}_\mu - \mathrm{i}\alpha^2{}_\mu}{\sqrt{2}}\,, \\ W_\mu^- &= \frac{\alpha^1{}_\mu + \mathrm{i}\alpha^2{}_\mu}{\sqrt{2}}\,, & Z_\mu &= -\alpha^0{}_\mu + \alpha^3{}_\mu\,. \end{aligned} \tag{160}$$

Here we have also introduced the traditional notation for the gauge vector bosons. The electromagetic charges of the four bosons are already built into the notation, while their masses can be read from (155):

$$M_A = 0\,, \qquad M_{W^\pm} = \frac{v}{2}g_1\,, \qquad M_Z = \frac{v}{2}\sqrt{g_0^2 + g_1^2}\,. \tag{161}$$

In the new basis, the squared-mass matrix and the stabilizer matrix are diagonal, while the bilinear form $g_{\alpha\beta}$ and the gauge coupling constant bilinear form $C_{\alpha\beta}$ become:

$$
g_{\alpha\beta} = \begin{pmatrix} 2 & 0 & 0 & \frac{g_1^2-g_0^2}{g_1^2+g_0^2} \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{g_1^2-g_0^2}{g_1^2+g_0^2} & 0 & 0 & \frac{g_1^4+g_0^4}{\left(g_1^2+g_0^2\right)^2} \end{pmatrix}, \qquad C_{\alpha\beta} = \begin{pmatrix} \frac{g_0^2+g_1^2}{g_0^2 g_1^2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{g_1^2} & 0 \\ 0 & \frac{1}{g_1^2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{g_0^2+g_1^2} \end{pmatrix}
$$

$$
\equiv \begin{pmatrix} \frac{1}{g_A^2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{g_W^2} & 0 \\ 0 & \frac{1}{g_W^2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{g_Z^2} \end{pmatrix}. \tag{162}
$$

Returning to the action (151), we see that the second and third rows represent the appropriate Proca constraint term, which leaves us with the action in the following final form:

$$
\tilde{S} = S_{\text{grav}} + S_{\text{Yang–Mills}} + S_{\text{CC}} + \tilde{S}_{\text{Higgs}} + \tilde{S}_{3BF} + \tilde{S}_{\text{scal}} + S_{\text{Proca}}
$$

$$
+ \frac{1}{2} \int h\,\alpha^{\alpha} P_{\alpha}{}^{\beta} \wedge \left(\rho_{\beta} - \kappa^{-1}_{\beta\gamma} \Theta^{\gamma ab} \wedge e_a \wedge e_b\right) + v \int \mathrm{d}\tilde{\gamma}_H \,. \tag{163}
$$

The first row in the action contains terms which describe one real scalar field h (the Higgs field) with mass $m = 2v\sqrt{2\chi}$, and four vector bosons with masses specified in (161), coupled to gravity and to each other. The first term in the second row describes the interaction between the Higgs field and the vector bosons, so that all interactions are equvalent to the interactions of the ordinary electroweak theory. The second term in the second row is a boundary term, and as such it does not contribute to the equations of motion of the theory.

This concludes our analysis of the Higgs mechanism in the context of constrained $3BF$ theory. In short, the result is the same as in the textbook approach to the spontaneous symmetry breaking in electroweak theory. Nevertheless, the technical details that enter the analysis are novel and completely different from the textbook approach, since the $3BF$ formulation of the electroweak action is specified over a different configuration space.

6. Conclusions

6.1. Summary of the results

Let us summarize the results of the paper. In section 2, we gave a review of the action representing the SM coupled to Einstein–Cartan gravity, within the framework of higher gauge theory. In particular, the action of the model is written as a constrained $3BF$ action, based on a convenient choice of a 3-group representing the gauge symmetry of the model. Section 2 also features one theorem (proved in appendix A) which is important for the study of spontaneous symmetry breaking within the higher gauge theory framework, and represents a new result. Section 3 was devoted to the study of explicit symmetry breaking of the gauge group of the topological $3BF$ sector, due to the presence of the constraints. Each constraint was studied separately, and we discussed which gauge sector is being broken by which constraint. The results have been summarized in the table. In section 4 we turned our attention to the $3BF$ formulation of the theory for the Proca field coupled to gravity. This was important for the subsequent comparison with the action for the electroweak theory after spontaneous symmetry breaking. Three completely novel and different formulations of the Proca constraint have been discussed, the first for a single Proca field, the second for multiple Proca fields, and the third also for multiple Proca fields in an extended configuration space convenient for comparison with the electroweak model. Finally, in section 5 we took up the main task of studying the spontaneous symmetry breaking and the Higgs mechanism for the 3BF formulation of the electroweak model. While the Higgs mechanism is conceptually the same as in the ordinary textbook presentations of the electroweak theory, the structure and details of the 3BF version of the action are very different from the standard textbook approach, so much that the complete procedure of spontaneous symmetry breaking had to be done anew, with many highly nontrivial details of the calculation. In this sense, the details of the symmetry breaking procedure described in section 5 represent one of the main results of the paper.

6.2. Discussion

Regarding the above results, there are two main comments that need to be addressed. The first comment deals with the question what happens with the structure of the 3-group as a consequence of the spontaneous symmetry breaking. Namely, the initial 2-crossed module corresponding to the electroweak theory was based on the following choice of the groups (see equation (97)):

$$G = SO(3,1) \times SU(2) \times U(1)\,, \qquad H = \mathbb{R}^4\,, \qquad L = \mathbb{C}^4\,. \tag{164}$$

However, after the analysis, the resulting action (163) does not correspond anymore to this 2-crossed module. Instead, it is straightforward to see that the final 2-crossed module is based on the following choice of the groups:

$$G = SO(3,1) \times SU(2) \times U(1)\,, \qquad H = \mathbb{R}^4\,, \qquad L = \mathbb{R}\,. \tag{165}$$

There are two important points to emphasize here. First, the group L has been reduced from the one describing four complex scalar fields to the one describing a single real scalar field. This is a direct consequence of the spontaneous symmetry breaking, in particular of the gauge transformation (117) which was applied to gauge away the fields ϕ_1, ϕ_2 and ϕ_4. The gauge transformation has induced a reduction of the configuration space of the theory, leaving only a single real scalar field h remaining in the action. In turn, the reduction of the configuration space is consistent with the choice (165) of the 2-crossed module, which corresponds precisely to the topological 3BF action (137), obtained by the gauge fixing procedure from the initial 3BF action based on the 2-crossed module (164).

The second point considers the group G. Formally, the group G remains the same in both the initial and the final 2-crossed module. Nevertheless, as we have seen in section 3, the Proca constraint in fact breaks the G symmetry group, and is the only constraint to do so. Therefore, despite the fact that the topological 3BF sector of the initial and final actions shares the same BF term and the same connection 1-form α stemming from the group G, the presence of the Proca constraint in the final action in fact breaks the group G down to its subgroup $SO(3,1) \times U(1)$, whereas the initial action did not feature the Proca constraint and the group G was not broken. The end result is that the final action has a broken G symmetry, despite the fact that it is based on the 2-crossed module (165) featuring the full group G. This happens due to the appearance of the Proca constraint during the spontaneous symmetry breaking of the action.

The second comment that needs to be addressed deals with the question of the spontaneous symmetry breaking of the whole SM action (19). Namely, in section 5 we have studied the details of the spontaneous symmetry breaking and the Higgs mechanism on the special case of the electroweak theory, in order to keep the analysis as simple as possible. Nevertheless, it is straightforward to add the remaining three constraints S_{Dirac}, S_{Yukawa}, and S_{spin}, as well as the corresponding $\langle D \wedge \mathcal{H} \rangle_{\mathfrak{t}}$ term for fermions to the action, and examine the same procedure for the full SM. The resulting action will have the same terms as the action for the electroweak theory, up to terms corresponding to fermions, and up to the overall presence of the color $SU(3)$ gauge symmetry (which remains unbroken and does not play a role in the Higgs mechanism). The $\langle D \wedge \mathcal{H} \rangle_{\mathfrak{t}}$ term for fermions is equal to

$$\langle D_f \wedge \mathcal{H}_f \rangle_{\mathfrak{t}} \equiv \bar{\psi}_A \left(\overrightarrow{\nabla} \gamma \right)^A - \left(\bar{\gamma} \overleftarrow{\nabla} \right)_A \psi^A \,, \tag{166}$$

and it transforms trivially under the transformation (129) as well as the S_{Dirac} and S_{spin} constraints

$$\langle D_f \wedge \mathcal{H}_f \rangle_{\mathfrak{t}} \xrightarrow{\mathrm{e}^{-\xi}} \langle \tilde{D}_f \wedge \tilde{\mathcal{H}}_f \rangle_{\mathfrak{t}} = \langle D_f \wedge \mathcal{H}_f \rangle_{\mathfrak{t}} \,, \qquad S_{\text{Dirac}} \xrightarrow{\mathrm{e}^{-\xi}} \tilde{S}_{\text{Dirac}} = S_{\text{Dirac}} \,,$$
$$S_{\text{spin}} \xrightarrow{\mathrm{e}^{-\xi}} \tilde{S}_{\text{spin}} = S_{\text{spin}} \,. \tag{167}$$

The only constraint which does not transform trivially is S_{Yukawa}, and it splits into two terms:

$$\begin{aligned} S_{\text{Yukawa}} &= -\int \frac{2}{4!} Y_{ABC} \bar{\psi}^A \psi^B \phi^C \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d \\ &\xrightarrow{\mathrm{e}^{-\xi}} -\frac{1}{12} \int v Y_{ABH} \bar{\psi}^A \psi^B \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d \\ &\quad -\frac{1}{12} \int Y_{ABH} \bar{\psi}^A \psi^B h \, \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d \,, \end{aligned} \tag{168}$$

where the first term on the right hand side has the form similar to the Dirac mass term (70), while the second term is the new Yukawa constraint $\tilde{S}_{\text{Yukawa}}$, describing the interaction between fermions and the Higgs field h. Comparing the first term with the Dirac mass term, we conclude that Yukawa couplings Y_{ABH} are proportional to fermion mass matrix

$$M_{AB} = v Y_{ABH} \,, \tag{169}$$

which consists of the actual fermion masses and the corresponding mixing angles. The final form of the transformed Yukawa constraint thus becomes:

$$S_{\text{Yukawa}} \xrightarrow{\mathrm{e}^{-\xi}} \tilde{S}_{\text{Yukawa}} - \frac{1}{12} \int M_{AB} \bar{\psi}^A \psi^B \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d = \tilde{S}_{\text{Yukawa}} + S_{\text{Dirac mass}} \,. \tag{170}$$

Thus, we conclude that the Higgs mechanism described for the electroweak model can be generalized in a straightforward way to the full SM action (19).

Data availability statement

No new data were created or analysed in this study.

Acknowledgments

The authors would like to thank Tijana Radenković and Mihailo Đorđević for useful discussions and suggestions.

This research was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia (MPNTR). MV was additionally supported by the Science Fund of the Republic of Serbia, Grant No. 7745968, project 'Quantum Gravity from Higher Gauge Theory 2021' – QGHG-2021. The contents of this publication are the sole responsibility of the authors and can in no way be taken to reflect the views of the Science Fund of the Republic of Serbia.

Appendix A. Proof of the theorem from the main text

Theorem. *Given a 2-crossed module* $(L \xrightarrow{\delta} H \xrightarrow{\partial} G,\ \triangleright,\ \{\_\,,\_\}_{\mathrm{pf}})$ *and symmetric, nondegenerate bilinear forms* $\langle\_\,,\_\rangle_{\mathfrak{g}}$, $\langle\_\,,\_\rangle_{\mathfrak{h}}$ *and* $\langle\_\,,\_\rangle_{\mathfrak{l}}$, *if the bilinear forms are G-invariant then the components of the action* $\triangleright_{\alpha\beta\gamma}$, $\triangleright_{\alpha ab}$ *and* $\triangleright_{\alpha AB}$ *are antisymmetric with respect to the second and third index. In addition, there exists a choice of basis in Lie algebras* $\mathfrak{g}$, $\mathfrak{h}$ *and* $\mathfrak{l}$ *such that* $\triangleright_{\alpha\beta}{}^{\gamma}$, $\triangleright_{\alpha a}{}^{b}$ *and* $\triangleright_{\alpha A}{}^{B}$ *have vanishing diagonal elements with respect to the second and third index, and in this basis the bilinear form is also diagonal.*

Proof.

Let us first note that the statement of G-invariance of the bilinear form $\langle\_\,,\_\rangle_{\mathfrak{h}}$ is defined as

$$\langle g \triangleright h_1, g \triangleright h_2 \rangle_{\mathfrak{h}} = \langle h_1, h_2 \rangle_{\mathfrak{h}}\,, \tag{A1}$$

for all $g \in G$ and all $h_1, h_2 \in \mathfrak{h}$. Expanding g, h_1 and h_2 in appropriate bases, one can easily see that the left-hand side can be rewritten as:

$$\langle g \triangleright h_1, g \triangleright h_2 \rangle_{\mathfrak{h}} = h_1^a h_2^c \left[g_{ac} + g^{\alpha} \left(\triangleright_{\alpha c}{}^{d} g_{ad} + \triangleright_{\alpha a}{}^{d} g_{dc} \right) \right] + \mathcal{O}\left(g^2\right)\,. \tag{A2}$$

Equating this to the right-hand side, one sees that the right-hand side cancels the first term in the brackets. Then, emloying the symmetry of the bilinear form, the term in the parentheses reduces to:

$$\triangleright_{\alpha ca} + \triangleright_{\alpha ac} = 0\,, \tag{A3}$$

as was stated by the theorem. The antisymmetry of the remaining two actions $\triangleright_{\alpha\beta\gamma}$ and $\triangleright_{\alpha AB}$ is proved analogously.

In addition, the nondegeneracy of the bilinear form g_{ab} implies that there exists its inverse, denoted g^{ab}. Then, contracting (A3) with g^{ac} one immediately obtains a basis-independent statement that the action is traceless,

$$\triangleright_{\alpha a}{}^{a} = 0. \tag{A4}$$

Moreover, one can always choose a basis in a Lie algebra $\mathfrak{h}$ such that the bilinear form g_{ab} and its inverse are diagonal. From the identity

$$\triangleright_{\alpha a}{}^{b} = \triangleright_{\alpha ac} g^{cb} \tag{A5}$$

one can observe that, in this particular basis, $\triangleright_{\alpha a}{}^{b}$ must be proportional to $\triangleright_{\alpha ab}$ since $g^{cb}=0$ for $c \neq b$. Then, since $\triangleright_{\alpha ab}$ is antisymmetric due to (A3), it is equal to zero for $a=b$, which implies that $\triangleright_{\alpha a}{}^{b}$ is also zero in that case. In other words, $\triangleright_{\alpha a}{}^{b}$ has vanishing diagonal elements with respect to second and third index, as was stated by the theorem. The same property for the remaining two actions $\triangleright_{\alpha\beta}{}^{\gamma}$ and $\triangleright_{\alpha A}{}^{B}$ is proved analogously.

Appendix B. Arbitrariness of the κ-matrix

In section 5 we introduced the bilinear form $\kappa^{\alpha\beta}$ via (144) as well as the new variables (148) which satisfy the list of identities (149). In order to maintain one-to-one correspondence between the old and the new variables, in these identities the number of independent variables on the right hand side has to be equal to the number of the independent variables on the left hand side. Moreover, none of these new variables (148) should be explicitly multiplied by zero during the construction of the identities (149). These requirements have nontrivial consequences on arbitrariness of the choice of the bilinear form $\kappa^{\alpha\beta}$. For instance, let us consider the first identity from (149),

$$\theta^{\alpha} \wedge \rho_{\alpha} = \lambda^{\bar{A}} \wedge \gamma_{\bar{A}}. \tag{B1}$$

From the definition of the variable ρ_α in (148) we conclude that the action of the projector (145) does not change ρ_α

$$\rho_{\alpha} = P_{\alpha}{}^{\beta} \rho_{\beta}. \tag{B2}$$

This implies that this projector does not change the left hand side of identity (B1) and it holds that

$$\theta^{\alpha} \wedge \rho_{\alpha} = \theta^{\alpha} \wedge P_{\alpha}{}^{\beta} \rho_{\beta}. \tag{B3}$$

Using the requirement that none of the variables should be multiplied by zero during the construction of the identity (B1), we conclude that the action of projector must also leave the variable θ^α invariant. Using the definition of the θ^α variable we obtain a nontrivial condition on $\kappa^{\alpha\beta}$:

$$\theta^{\alpha} = \theta^{\beta} P_{\beta}{}^{\alpha} \quad \Rightarrow \quad -2\kappa^{\alpha\gamma} \triangleright_{\gamma}{}^{H}{}_{\bar{A}} \lambda^{\bar{A}} \equiv -2\kappa^{\alpha\gamma} P_{\gamma}{}^{\delta} \triangleright_{\delta}{}^{H}{}_{\bar{A}} \lambda^{\bar{A}} = -2 P_{\beta}{}^{\alpha} \kappa^{\beta\gamma} P_{\gamma}{}^{\delta} \triangleright_{\delta}{}^{H}{}_{\bar{A}} \lambda^{\bar{A}}, \tag{B4}$$

which due to arbitrariness of the field $\lambda^{\bar{A}}$ implies that

$$\kappa^{\alpha\gamma} P_{\gamma}{}^{\delta} = P_{\beta}{}^{\alpha} \kappa^{\beta\gamma} P_{\gamma}{}^{\delta}. \tag{B5}$$

Using the fact that $\kappa^{\alpha\beta}$ is symmetric, by transposing (B5) we obtain that the projector and $\kappa^{\alpha\beta}$ commute

$$\kappa^{\alpha\gamma} P_{\gamma}{}^{\delta} = P_{\gamma}{}^{\alpha} \kappa^{\gamma\delta}. \tag{B6}$$

Now, let us consider how this new restriction on $\kappa^{\alpha\beta}$ reduces its arbitrariness. Combining (146) and (B6) we obtain

$$P_{\gamma}{}^{\alpha} A^{\gamma\delta} \left[\delta^{\beta}_{\delta} - P_{\delta}{}^{\beta} \right] = \left[\delta^{\alpha}_{\gamma} - P_{\gamma}{}^{\alpha} \right] A^{\gamma\delta} P_{\delta}{}^{\beta}, \tag{B7}$$

which is satisfied only if

$$A^{\alpha\delta}\left[\delta^{\beta}_{\delta}-P_{\delta}{}^{\beta}\right]=\left[\delta^{\alpha}_{\gamma}-P_{\gamma}{}^{\alpha}\right]A^{\gamma\beta}=\left[\delta^{\alpha}_{\gamma}-P_{\gamma}{}^{\alpha}\right]A^{\gamma\delta}\left[\delta^{\beta}_{\delta}-P_{\delta}{}^{\beta}\right]. \tag{B8}$$

Since the dimension of the subspace of the orthogonal projector is equal to one, the arbitrariness in the choice of the bilinear form $\kappa^{\alpha\beta}$ is reduced to a single free parameter, whose form is given in (150):

$$\kappa^{\alpha\beta}\to\kappa^{\alpha\beta}+\left[\delta^{\alpha}_{\gamma}-P_{\gamma}{}^{\alpha}\right]A^{\gamma\delta}\left[\delta^{\beta}_{\delta}-P_{\delta}{}^{\beta}\right]. \tag{B9}$$

The arbitrariness of this parameter guarantees the existence of the inverse bilinear form $\kappa^{-1}_{\alpha\beta}$, which also commutes with the projector as a consequence of (B6).

Next, we turn to the effect of this arbitrariness onto the action (151). Using the fact that the inverse bilinear form acts on variables (148) in (151), and variables (148) are invariant under projector action, the action (151) of the theory, and squared-mass matrix (153) depend only on the projection of inverse bilinear form $\kappa^{-1}_{\alpha\beta}P_{\gamma}{}^{\beta}$. The arbitrariness of the inverse bilinear form $\kappa^{-1}_{\alpha\beta}$ can be expressed as a power series in terms of the arbitrary bilinear form $A^{\gamma\delta}$, as

$$\kappa^{-1}_{\alpha\beta}\to\kappa^{-1}_{\alpha\gamma}\sum_{n=0}^{\infty}\left[(-1)^{n}\left([\delta-P]A[\delta-P]\kappa^{-1}\right)^{n}\right]^{\gamma}{}_{\beta}, \tag{B10}$$

from where one can obtain that the projection of the inverse bilinear form does not depend on choice of the arbitrary bilinear form $A^{\gamma\delta}$. This in turn implies that the action (151) and the squared-mass matrix (153) are uniquely defined.

ORCID iDs

Pavle Stipsić https://orcid.org/0000-0001-5790-3997
Marko Vojinović https://orcid.org/0000-0001-6977-4870

References

[1] Polchinski J 1998 *String Theory Vol. I: An Introduction to the Bosonic String* (Cambridge University Press)
[2] Polchinski J 1998 *String Theory Vol. II: Superstring Theory and Beyond* (Cambridge University Press)
[3] Rovelli C 2004 *Quantum Gravity* (Cambridge University Press)
[4] Thiemann T 2007 *Modern Canonical Quantum General Relativity* (Cambridge University Press)
[5] Rovelli C and Vidotto F 2014 *Covariant Loop Quantum Gravity* (Cambridge University Press)
[6] Bianchi E, Han M, Magliaro E, Perini C, Rovelli C and Wieland W 2013 Spinfoam fermions *Class. Quantum Grav.* **30** 235023
[7] Vojinović M 2014 Cosine problem in EPRL/FK spin foam model *Gen. Relativ. Gravit.* **46** 1616
[8] Miković A and Vojinović M 2023 *State-sum Models of Piecewise Linear Quantum Gravity* (World Scientific)
[9] Baez J C and Huerta J 2011 An invitation to higher gauge theory *Gen. Relativ. Gravit.* **43** 2335
[10] Crane L and Sheppeard M D 2003 2-Categorical Poincaré representations and state sum applications (arXiv:math/0306440)
[11] Miković A and Vojinović M 2012 Poincaré 2-group and quantum gravity *Class. Quantum Grav.* **29** 165003
[12] Li Z 2019 A global geometric approach to parallel transport of strings in gauge theory (arXiv:1910.14230)

[13] Saemann C and Wolf M 2014 Non-abelian tensor multiplet equations from twistor space *Commun. Math. Phys.* **328** 527
[14] Jurco B, Saemann C and Wolf M 2005 Semistrict higher gauge theory *J. High Energy Phys.* JHE04(2015)087
[15] Hidaka Y, Nitta M and Yokokura R 2020 Higher-form symmetries and 3-group in axion electrodynamics *Phys. Lett. B* **808** 135672
[16] Hidaka Y, Nitta M and Yokokura R 2021 Global 3-group symmetry and't Hooft anomalies in axion electrodynamics *J. High Energy Phys.* JHE01(2021)173
[17] Sämann C and Wolf M 2014 Six-dimensional superconformal field theories from principal 3-bundles over twistor space *Lett. Math. Phys.* **104** 1147
[18] Song D, Wu K and Yang J 2022 3-form Yang-Mills based on 2-crossed modules *J. Geom. Phys.* **178** 104537
[19] Song D, Wu K and Yang J 2022 Higher form Yang-Mills as higher BFYM theories *Eur. Phys. J. C* **82** 1034
[20] Hidaka Y, Nitta M and Yokokura R 2021 Topological axion electrodynamics and 4-group symmetry *Phys. Lett.* B **823** 136762
[21] Hidaka Y, Nitta M and Yokokura R 2022 Global 4-group symmetry and't Hooft anomalies in topological axion electrodynamics *Prog. Theor. Exp. Phys.* **2022** 04A109
[22] Jurco B, Saemann C and Wolf M 2016 Higher groupoid bundles, higher spaces and self-dual tensor field equations *Fortsch. Phys.* **64** 674
[23] Saemann C and Wolf M 2017 Supersymmetric Yang-Mills theory as higher Chern-Simons theory *J. High Energy Phys.* JHE07(2017)111
[24] Jurco B, Macrelli T, Raspollini L, Saemann C and Wolf M 2019 L-infinity algebras, the BV formalism and classical fields *Fortsch. Phys.* **67** 1910025
[25] Radenković T and Vojinović M 2019 Higher gauge theories based on 3-groups *J. High Energy Phys.* JHE10(2019)222
[26] Radenković T and Vojinović M 2022 Topological invariant of 4-manifolds based on a 3-group *J. High Energy Phys.* JHE07(2022)105
[27] Batalin I A and Vilkovisky G A 1981 Gauge algebra and quantization *Phys. Lett. B* **102** 27
[28] Batalin I A and Vilkovisky G A 1983 Feynman rules for reducible gauge theories *Phys. Lett. B* **120** 166
[29] Batalin I A and Vilkovisky G A 1983 Quantization of gauge theories with linearly dependent generators *Phys. Rev.* D **28** 2567
Batalin I A and Vilkovisky G A 1984 Quantization of gauge theories with linearly dependent generators *Phys. Rev.* D **30** 508 (erratum)
[30] Batalin I A and Vilkovisky G A 1984 Closure of the gauge algebra, generalized Lie equations and Feynman rules *Nucl. Phys. B* **234** 106
[31] Batalin I A and Vilkovisky G A 1985 Existence theorem for gauge algebra *J. Math. Phys.* **26** 172
[32] Alexandrov M, Schwarz A, Zaboronsky O and Kontsevich M 1997 The geometry of the master equation and topological quantum field theory *Int. J. Mod. Phys. A* **12** 7
[33] Grigoriev M 2011 Parent formulation at the Lagrangian level *J. High Energy Phys.* JHE07(2011)061
[34] Alkalaev K B and Grigoriev M 2014 Frame-like Lagrangians and presymplectic AKSZ-type sigma models *Int. J. Mod. Phys.* A **29** 18
[35] Dneprov I, Grigoriev M and Gritzaenko V 2024 Presymplectic minimal models of local gauge theories (arXiv:2402.03240)
[36] Cattaneo A S, Menger L and Schiavina M 2024 Gravity with torsion as deformed BF theory *Class. Quantum Grav.* **41** 155001
[37] Borsten L, Kanakaris D and Kim H 2024 Gravity from AKSZ-Manin theories in two, three, and four dimensions (arXiv:2410.10755)
[38] Pulmann J, Ševera P and Valach F 2021 A non-abelian duality for (higher) gauge theories *Adv. Theor. Math. Phys.* **25** 1
[39] Plebanski J F 1977 On the separation of Einsteinian substructures *J. Math. Phys.* **18** 2511
[40] Baez J C 2000 An introduction to spin foam models of quantum gravity and BF theory *Lect. Notes Phys.* **543** 25
[41] Celada M, González D and Montesinos M 2016 BF gravity *Class. Quantum Grav.* **33** 213001
[42] Baez J 1996 4-dimensional BF theory as a topological quantum field theory *Lett. Math. Phys.* **38** 129

[43] Girelli F, Pfeiffer H and Popescu E M 2008 Topological higher gauge theory – from BF to BFCG theory *J. Math. Phys.* **49** 032503
[44] Martins J F and Miković A 2011 Lie crossed modules and gauge-invariant actions for 2-BF theories *Adv. Theor. Math. Phys.* **15** 1059
[45] Miković A and Oliveira M A 2015 Canonical formulation of Poincaré BFCG theory and its quantization *Gen. Relativ. Gravit.* **47** 58
[46] Miković A, Oliveira M A and Vojinović M 2016 Hamiltonian analysis of the *BFCG* theory for the Poincaré 2-group *Class. Quantum Grav.* **33** 065007
[47] Miković A, Oliveira M A and Vojinović M 2022 Hamiltonian analysis of the BFCG theory for a generic Lie 2-group *Adv. Theor. Math. Phys.* **26** 3783
[48] Miković A, Oliveira M A and Vojinović M 2019 Hamiltonian analysis of the BFCG formulation of general relativity *Class. Quantum Grav.* **36** 015005
[49] Asante S K, Dittrich B, Girelli F, Riello A and Tsimiklis P 2020 Quantum geometry from higher gauge theory *Class. Quantum Grav.* **37** 205001
[50] Girelli F, Laudonio M and Tsimiklis P Polyhedron phase space using 2-groups: kappa-Poincare as a Poisson 2-group (arXiv:2105.10616)
[51] Miković A and Vojinović M 2021 Standard model and 4-groups *Europhys. Lett.* **133** 61001
[52] Porter T 1998 Topological quantum field theories from homotopy n-types *J. London Math. Soc.* **58** 723
[53] Radenković T and Vojinović M 2024 The 3BF theory as a TQFT (arXiv:2412.21032)
[54] Blagojević M 2002 *Gravitation and Gauge Symmetries* (Institute of Physics Publishing)
[55] Radenković T and Vojinović M 2020 Hamiltonian analysis for the scalar electrodynamics as 3*BF* theory *Symmetry* **12** 620
[56] Radenković T and Vojinović M 2022 Gauge symmetry of the 3*BF* theory for a generic Lie 3-group *Class. Quantum Grav.* **39** 135009
[57] Martins J F and Picken R 2011 The fundamental Gray 3-groupoid of a smooth manifold and local 3-dimensional holonomy based on a 2-crossed module *Differ. Geom. Appl. J.* **29** 179–206
[58] Wang W 2014 On 3-gauge transformations, 3-curvatures and Gray-categories *J. Math. Phys.* **55** 043506
[59] Baez J and Dolan J 1995 Higher dimensional algebra and topological quantum field theory *J. Math. Phys.* **36** 6073
[60] Conduché D 1984 Modules croisés généralisés de longueur 2 *J. Pure Appl. Algebra* **34** 155
[61] Yetter D N 1993 TQFT's from homotopy 2-types *J. Knot Theory Ramifications* **2** 113
[62] Girelli F and Pfeiffer H 2004 Higher gauge theory—differential versus integral formulation *J. Math. Phys.* **45** 3949
[63] Berends F A, Burgers G J H and van Dam H 1985 On the theoretical problems in constructing interactions involving higher spin massless particles *Nucl. Phys.* **B260** 295
[64] Lada T and Stasheff J 1993 Introduction to SH Lie algebras for physicists *Int. J. Theor. Phys.* **32** 1087
[65] Stasheff J 1998 The (secret?) homological algebra of the Batalin-Vilkovisky approach *Contemp. Math.* **219** 195
[66] Hohm O and Zwiebach B 2017 L-infinity algebras and field theory *Fortsch. Phys.* **65** 1700014
[67] Đorđević M, Radenković T, Stipsić P and Vojinović M 2023 Henneaux-Teitelboim gauge symmetry and its applications to higher gauge theories *Universe* **9** 281

IOP Publishing Classical and Quantum Gravity
Class. Quantum Grav. **42** (2025) 195009 (35pp) https://doi.org/10.1088/1361-6382/ae06e9

Correspondence between 3*BF* and Einstein–Cartan formulations of quantum gravity

Pavle Stipsić and Marko Vojinović*

Institute of Physics, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia

E-mail: vmarko@ipb.ac.rs and pstipsic@ipb.ac.rs

Received 26 June 2025; revised 5 September 2025
Accepted for publication 15 September 2025
Published 25 September 2025

Abstract
We construct a correspondence between the quantized constrained 3*BF* theory and the quantized Einstein–Cartan theory with contact spin–spin interaction, both of which describe the Standard Model coupled to Einstein–Cartan gravitational field. First we introduce the expectation values of observables using the path integral formalism for both theories, and then by integrating out some configuration space variables in the quantum 3*BF* theory we obtain the definition of the corresponding observable in the quantum Einstein–Cartan theory with contact interaction. The correspondence is a rather general result, since it can be established without actually performing the detailed quantization of either theory. Finally, we discuss the differences in the predictions of the two theories on the example of the 4-volume density of spacetime, and on the example of gravitational waves.

Keywords: quantum gravity, higher gauge theory, 3-group, 3*BF* action, spin–spin contact interaction, Einstein–Cartan action, path integral quantization

1. Introduction

Quantization of the gravitational field represents one of the main open problems in modern theoretical physics. Over the years, vast research disciplines aiming to formulate a theory of quantum gravity have been proposed and developed, the most prominent ones being string theory [1, 2], loop quantum gravity (LQG) [3, 4], and others. Within the LQG framework, one of the promising research directions is based on the idea of covariant quantization, i.e. the quantization by providing a rigorous definition of a path integral for the gravitational field.

* Author to whom any correspondence should be addressed.

This is commonly known as the spinfoam quantization programme, and several models of quantized gravitational field have been proposed in the literature [5, 6].

One of the typical drawbacks of the spinfoam quantization programme is the lack of matter fields in the theory, since the quantization method is adapted to work only for the gravitational field itself. Consequently, various strategies have been proposed to circumvent this issue and generalize the spinfoam quantization programme to include both gravity and matter on equal footing. One such promising generalization has recently been developed, and is based on the application of the so-called higher gauge theory [7–9] and topological quantum field theory techniques [10–16]. Higher gauge theory is a mathematical framework that provides one with a way to generalize the notion of a gauge symmetry structure by using the higher category theory analogs of Lie groups, called Lie n-groups [17–29, 30]. In order to successfully implement the spinfoam quantization procedure for a theory that describes both gravity and matter on an equal footing, attention focuses on the notion of a 3-group, and its corresponding topological action called a $3BF$ action [31], which represents a suitable generalization of the well known BF and $2BF$ actions based on an ordinary Lie group and a Lie 2-group [32–45] (see also [46] for the $4BF$ theory and the 4-group approach). A number of recent results [17, 31, 47–53] have successfully implemented several stages of the spinfoam quantization programme for the constrained $3BF$ theory corresponding to the gravitational field coupled to the full Standard Model. Specifically, general relativity (GR) was first rewritten as a constrained $2BF$ model in [17], a first result which emphasized the relevance of nBF models for realistic physics. Next, in [31, 47] the theory was extended to a $3BF$ model in order to include matter fields and couple them to gravity, in particular all fields present in the Standard Model (gauge bosons, fermions and scalar fields). After that, the properties of the resulting classical theory have been studied in detail—the phase space, Hamiltonian analysis, and gauge symmetries of the theory have been discussed in [48, 49], the additional trivial gauge symmetries of the theory were discussed in [50], while symmetry breaking and the Higgs mechanism were studied in [51]. Finally, the quantization of the topological $3BF$ theory, and the construction of the topological invariant and its corresponding TQFT, have been done in [52, 53]. All this research has demonstrated that the approach to quantum gravity based on the 3-group and the $3BF$ action is technically tangible—the goal of providing a rigorous formulation of a theory of quantum gravity with matter seems to be a viable and achievable prospect, using the $3BF$ action and a 3-group as the starting point.

In this work, we will focus on one interesting property of the $3BF$ formulation of a theory of quantum gravity. Specifically, we will establish a correspondence between the quantization of the suitable constrained $3BF$ action, and the quantization of the standard Einstein–Cartan (EC) formulation of GR coupled to the Standard Model. While we will not study the actual details of the quantization of either the $3BF$ theory nor the EC theory, we will nevertheless be able to introduce a very precise relationship between the two quantum theories. Namely, given any quantum observable that is defined within the context of one specific version of EC theory (called the EC contact theory), we will introduce a corresponding observable defined within the context of the quantized $3BF$ theory, such that the expectation values of the two observables match exactly (and vice versa). This correspondence is established at a full nonperturbative level, and it is a surprising feature of the $3BF$ theory that one can in fact formulate such a correspondence using only some general assumptions, i.e. without introducing all details of the actual quantization of either theory.

The obtained correspondence has two important consequences. First, it enables one to in fact *define* the quantization of the EC contact theory coupled to the Standard Model (originally a very hard problem to solve) by passing to the $3BF$ version of the theory, and performing the quantization of the $3BF$ theory instead (a slightly easier problem to solve). In this way, one can circumvent a number of problems that render the quantization of EC theory non-feasible, and establish it instead by exploiting the obtained correspondence to the quantum $3BF$ theory. The second important consequence of the correspondence is that there exists a regime where the two quantum theories could be experimentally distinguished from each other, at least in principle. To that end, after establishing the correspondence relations as the main result of this work, we will apply those relations to study a few interesting example observables, and discuss in what sense and under which conditions the two quantum theories could be experimentally distinguishable. For example, one can apply the correspondence to compare the magnitude of the strain generated by gravitational waves. In principle, given a source of gravitational waves that is both strong in magnitude and has large quantum uncertainty, one can evaluate the differences in the quantum corrections for the strain amplitude in $3BF$ theory and EC contact theory, and test them against the experimental data. In this sense, the correspondence predicts observable signatures that distinguish the two theories. Of course, we do not have actual access to a gravitational wave source with the required properties, so any such experimental proposal is still far away from the practical capabilities of current technology, but as a matter of principle, this question can be studied at least theoretically, and it does illustrate the phenomenological significance of the obtained correspondence.

The layout of the paper is as follows. In section 2, we present a short review of the classical $3BF$ and EC theories, and demonstrate that they give rise to equivalent sets of classical equations of motions (EoMs). In section 3, we turn to the main analysis of the expectation values of an arbitrary quantum observable defined in the two quantum theories. After some mathematical preliminaries, we establish a correspondence between the expectation value of the observable in one theory, and the expectation value of a similar observable in the other theory, where 'similar' means that the observable is weighted by some power of the absolute value of the determinant of the tetrad, $|e|^{\pm M}$. This correspondence is established in a fully nonperturbative way, and represents the main result of the paper. Section 4 deals with some illustrative example observables that one can study in order to compare the two quantum theories. First, we discuss the spacetime 4-volume density operator as a simple example, and also the classical limit of the two theories. Then, we discuss the case of the gravitational waves, and give an estimate of how large their quantum uncertainties must be in order to be able to experimentally distinguish between the two quantum theories. In section 5 we give our concluding remarks and some topics for future research. The appendix A contains proofs of some technical results used in the main text, while the appendix B contains some additional mathematical and notational details.

Our notation and conventions are as follows. Spacetime indices, denoted by the mid-alphabet Greek letters $\mu,\nu,\dots$, are raised and lowered by the spacetime metric $g_{\mu\nu}$, once it is defined. The Lorentz metric is denoted as $\eta_{ab} = \mathrm{diag}(-1,+1,+1,+1)$. The indices that are counting the generators of Lie groups G, H, and L are denoted with initial Greek letters $\alpha,\beta,\dots$, lowercase initial Latin letters $a,b,c,\dots$, and uppercase Latin indices $A,B,C,\dots$, respectively. The generators themselves are typically denoted as τ_α, t_a and T_A, respectively. We work in the natural system of units, defined by $c = \hbar = 1$ and $G = l_p^2$, where l_p is the Planck length.

The indices which correspond to the Lorentz group are pairs of indices ab and the quantities that depend on them are antisymmetric with respect to their interchange. This means that all independent components of these quantities, according to Einstein summation convention, are summed over twice. Because of this, the result of the summation should be divided by two. Alternatively, in order to avoid this problem, one can introduce the notation $[ab]$ which represents the pair of indices as a single index for which we always assume that $a > b$. Summation over such indices takes into account every independent component precisely once, so it is not necessary to divide the total by two. For example, given some quantity K^{ab}, one has

$$K^{[ab]}\sigma_{[ab]} = \frac{1}{2}K^{ab}\sigma_{ab}\,. \tag{1}$$

In this work, the square brackets will exclusively denote the pairs of Lorentz indices, rather than the usual antisymmetrization over those indices.

All additional notation and conventions used throughout the paper are explicitly defined in the text where they first appear. See also appendix B.

2. Review of the classical 3*BF* and EC actions

In this section, we will provide a short review of four classical theories that will be relevant for subsequent analysis. We will begin by introducing the topological 3*BF* action, based on the notion of a 3-group. This will then be employed to introduce the so-called constrained 3*BF* action, which gives rise to physically relevant EoM and is one of the main theories that we will subsequently study in sections 3 and 4. Then, we will introduce the standard EC action, coupled to the Standard Model in the usual way. Finally, we will introduce its corresponding second-order theory, called the EC contact action. The latter will be the second main theory that we will subsequently study in sections 3 and 4.

2.1. Topological 3BF action

In order to introduce the topological 3*BF* action, one first needs to introduce the notion of a strict Lie 3-group, a generalization of the notion of a Lie group stemming from higher category theory, which is equivalent to the algebraic structure called a Lie 2-crossed module. A Lie 2-crossed module is a triple of Lie groups, G, H and L, together with two homomorphisms between them,

$$\partial : H \to G\,, \qquad \delta : L \to H\,, \tag{2}$$

the actions of the group G on all three groups,

$$\triangleright : G \times X \to X\,, \qquad X = G, H, L\,, \tag{3}$$

as well as the Peiffer lifting map,

$$\{\_\,,\_\}_{\mathrm{pf}} : H \times H \to L\,. \tag{4}$$

All these maps are subject to a certain set of axioms, and together they make up a Lie 2-crossed module, denoted as

$$\left(L \xrightarrow{\delta} H \xrightarrow{\partial} G\,,\, \triangleright\,,\, \{\_\,,\_\}_{\mathrm{pf}} \right)\,. \tag{5}$$

This structure represents the notion of a 3-group in the most convenient way for our purposes. An interested reader can find further mathematical details for example in [7, 8, 31, 49, 52–55].

Given the mathematical structure of a 3-group, it gives rise to a natural choice of an action, called a 3*BF* action (see appendix B for a more detailed explanation of the notation used in this section and throughout the text). The 3*BF* action is purely topological, and defined as:

$$S^{\mathrm{top}}_{3BF} = \int_{\mathcal{M}_4} \langle B \wedge \mathcal{F} \rangle_{\mathfrak{g}} + \langle C \wedge \mathcal{G} \rangle_{\mathfrak{h}} + \langle D \wedge \mathcal{H} \rangle_{\mathfrak{l}}. \tag{6}$$

The Lagrange multipliers B, C and D are two-, one- and zero-forms, and simultaneously they are elements of Lie algebras $\mathfrak{g}$, $\mathfrak{h}$ and $\mathfrak{l}$, corresponding to the Lie groups G, H and L, respectively. The field strengths $\mathcal{F}$, $\mathcal{G}$ and $\mathcal{H}$ are defined as

$$\mathcal{F} = \mathrm{d}\alpha + \alpha \wedge \alpha - \partial\beta\,, \quad \mathcal{G} = \mathrm{d}\beta + \alpha \wedge^{\triangleright} \beta - \delta\gamma\,, \quad \mathcal{H} = \mathrm{d}\gamma + \alpha \wedge^{\triangleright} \gamma + \{\beta \wedge \beta\}_{\mathrm{pf}}\,, \tag{7}$$

and they are called fake curvatures for the connection one-form α, two-form β and three-form γ, which are also valued in algebras $\mathfrak{g}$, $\mathfrak{h}$ and $\mathfrak{l}$, respectively. Bilinear forms $\langle\_\,,\_\rangle_{\mathfrak{g}}$, $\langle\_\,,\_\rangle_{\mathfrak{h}}$ and $\langle\_\,,\_\rangle_{\mathfrak{l}}$ are assumed to be symmetric, nondegenerate and G-invariant, and they map a pair of algebra elements into a real number. Let us also note that, given the structure of the 3-group, one can introduce the notion of a covariant derivative as

$$\nabla = \mathrm{d} + \alpha \wedge^{\triangleright} \tag{8}$$

in the sense that, when ∇ acts for example on the components ϕ^A of the object $\phi \in \mathfrak{l}$, the action $\triangleright$ is being applied as the action from the Lie algebra $\mathfrak{g}$ to Lie algebra $\mathfrak{l}$, giving:

$$\nabla\phi^A = \mathrm{d}\phi^A + \triangleright_{\alpha B}{}^{A}\,\alpha^{\alpha} \wedge \phi^B\,, \tag{9}$$

and similarly for objects that are elements of algebras $\mathfrak{g}$ and $\mathfrak{h}$. Given this notation, one can rewrite the fake curvatures (7) in terms of ordinary curvatures as:

$$\mathcal{F} = \nabla^2 - \partial\beta\,, \qquad \mathcal{G} = \nabla\beta - \delta\gamma\,, \qquad \mathcal{H} = \nabla\gamma + \{\beta \wedge \beta\}_{\mathrm{pf}}\,. \tag{10}$$

We point the reader to the appendix B, which contains more detailed explanation of the above notation, including some examples.

In order to discuss the field content that corresponds to the Standard Model and EC gravity, one makes the following choice of the 3-group, called the Standard Model 3-group (see [31, 47, 51] for further details). The three Lie groups G, H and L are chosen as:

$$G = SO(3,1) \times SU(3) \times SU(2) \times U(1)\,, \qquad H = \mathbb{R}^4\,, \qquad L = \mathbb{C}^4 \times \mathbb{G}^{64} \times \mathbb{G}^{64} \times \mathbb{G}^{64}\,. \tag{11}$$

The physical interpretation of this choice is as follows. The group G represents the usual Standard Model gauge group, together with the local Lorentz group. The group H represents the spacetime translations, while the group L corresponds to the matter fields. Specifically, $\mathbb{C}^4$ corresponds to the Higgs sector, while the three Grassmann algebras $\mathbb{G}^{64}$ correspond to the three families of fermions.

In order to fully specify the Standard Model 3-group, one also needs to define all relevant maps. The homomorphisms ∂ and δ are chosen to to be trivial, as well as the Peiffer lifting $\{\_\,,\_\}_{\mathrm{pf}}$. Regarding the action $\triangleright$, it is defined as follows. The group G can be naturally split into the Lorentz part $SO(3,1)$ (generators counted using the indices $[ab]$) and the internal gauge part $SU(3) \times SU(2) \times U(1)$ (generators counted collectively using indices $\alpha, \beta, \ldots$). The action of G on itself is then given by specifying the action of the Lorentz part on itself and on the internal gauge part, as

$$\triangleright_{[ab][cd]}{}^{[ef]} \equiv f_{[ab][cd]}{}^{[ef]} = \frac{1}{2}\left(\eta_{[a|c}\delta^{[f}_{|b]}\delta^{|e]}_{d} - \eta_{[a|d}\delta^{[f}_{|b]}\delta^{|e]}_{c}\right)\,, \qquad \triangleright_{[ab]\beta}{}^{\gamma} = 0\,, \tag{12}$$

while the action of the internal gauge part on itself and on the Lorentz part is given as

$$\triangleright_{\alpha\beta}{}^{\gamma} = f_{\alpha\beta}{}^{\gamma}\,, \qquad \triangleright_{\alpha[ab]}{}^{[cd]} = 0\,. \tag{13}$$

Further, the action of G on H is specified naturally, assuming that the group H is interpreted as the group of 4-dimensional translations. Then the Lorentz part of G acts in the standard way on translations, while the internal part of G acts trivially:

$$\rhd_{[cd]a}{}^{b} = \frac{1}{2}\eta_{a[d|}\delta^{b}_{|c]}\,, \qquad \rhd_{\alpha a}{}^{b} = 0\,. \tag{14}$$

Finally, the action of the Lorentz and internal subgroups of G on L is also given in a natural way, in accordance with the transformation properties of various fermions and the Higgs scalar. For example, the action of G on left-isospin fermions is given as:

$$\rhd_{[cd]A}{}^{B} = \left(\sigma_{cd}\right)_{A}{}^{B}\,, \qquad \rhd_{\alpha A}{}^{B} = \frac{1}{2}\left(\sigma_{\alpha}\right)_{A}{}^{B}\,. \tag{15}$$

Here the matrices $(\sigma_\alpha)_A{}^B$ are Pauli matrices, and $(\sigma_{ab})_A{}^B = \frac{1}{4}[\gamma_a,\gamma_b]_A{}^B$, where γ_a are the standard Dirac matrices satisfying the anticommutation rule $\gamma_a\gamma_b + \gamma_b\gamma_a = -2\eta_{ab}$. Here we also introduce $\gamma_5 \equiv -\gamma_0\gamma_1\gamma_2\gamma_3$. In a similar way, one defines the action of group G for all other fermions and scalars in the group L, depending on their precise transformation properties (see [31] for details).

Given the Standard Model 3-group, one can rewrite the corresponding topological $3BF$ action (6) in the following form:

$$S_{3BF}^{\text{top}} = \int B_\alpha \wedge F^\alpha + B^{[ab]} \wedge R_{[ab]} + e_a \wedge \nabla\beta^a + \phi^A \left(\nabla\tilde{\gamma}\right)_A + \bar{\psi}_A \left(\overrightarrow{\nabla}\gamma\right)^A - \left(\bar{\gamma}\,\overleftarrow{\nabla}\right)_A \psi^A\,, \tag{16}$$

where we have introduced the following new notation. First, $\mathcal{F}$ is split into the internal symmetry field strength F^α (which is a function of the internal symmetry connection α^α) and the Riemann curvature two-form $R_{[ab]}$ (which is a function of the spin connection $\omega^{[ab]}$). The Lagrange multiplier C is rewritten as the tetrad field one-form e_a, and the Lagrange multiplier D is rewritten as a tuple of scalar and fermion fields $(\phi^A, \psi^A, \bar{\psi}_A)$. This change of notation also suggests the physical interpretation of the fields in (6).

2.2. *Constrained 3BF action*

While the action (16) does correspond to the Standard Model 3-group and features all relevant gravitational, gauge and matter fields, it does not provide the correct classical dynamics for those fields. Namely, this action is an example of a *topological* $3BF$ action, and as such it has trivial EoM, with no propagating degrees of freedom in the theory. In order to remedy this, one introduces additional terms to the action, called *simplicity constraint* terms. By conveniently choosing and adding simplicity constraints, we can introduce the full classical action corresponding to the Standard Model coupled to EC gravity, with the correct classical dynamics. Such an action is then commonly called the *constrained* $3BF$ action, and has the following form:

$$S_{3BF} = S_{3BF}^{\text{top}} + S_{\text{grav}} + S_{\text{scal}} + S_{\text{Dirac}} + S_{\text{Yang–Mills}} + S_{\text{Higgs}} + S_{\text{Yukawa}} + S_{\text{spin}} + S_{\text{CC}}\,. \tag{17}$$

Here we have:

$$S_{3BF}^{\text{top}} = \int B_\alpha \wedge F^\alpha + B^{[ab]} \wedge R_{[ab]} + e_a \wedge \nabla\beta^a + \phi^A \left(\nabla\tilde{\gamma}\right)_A + \bar{\psi}_A \left(\overrightarrow{\nabla}\gamma\right)^A - \left(\bar{\gamma}\,\overleftarrow{\nabla}\right)_A \psi^A\,, \tag{18}$$

$$S_{\rm grav} = -\int \lambda_{[ab]} \wedge \left(B^{[ab]} - \frac{1}{8\pi l_p^2} \varepsilon^{[ab]cd} e_c \wedge e_d \right), \tag{19}$$

$$S_{\rm scal} = \int \tilde{\lambda}^A \wedge \left(\tilde{\gamma}_A - H_{abcA} e^a \wedge e^b \wedge e^c \right) + \Lambda^{abA} \wedge \left(H_{abcA} \varepsilon^{cdef} e_d \wedge e_e \wedge e_f - (\nabla\phi)_A \wedge e_a \wedge e_b \right), \tag{20}$$

$$S_{\rm Dirac} = \int \bar{\lambda}_A \wedge \left(\gamma^A + \frac{i}{6} \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \left(\gamma^d \psi \right)^A \right) - \lambda^A \wedge \left(\bar{\gamma}_A - \frac{i}{6} \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \left(\bar{\psi} \gamma^d \right)_A \right), \tag{21}$$

$$S_{\rm Yang\text{–}Mills} = \int \lambda^\alpha \wedge \left(B_\alpha - 12 C_{\alpha\beta} M^\beta{}_{ab} e^a \wedge e^b \right) + \zeta_\alpha{}^{ab} \left(M^\alpha{}_{ab} \varepsilon_{cdef} e^c \wedge e^d \wedge e^e \wedge e^f - F^\alpha \wedge e_a \wedge e_b \right), \tag{22}$$

$$S_{\rm Higgs} = -\int \frac{2}{4!} \chi \left(\phi^A \phi_A - v^2 \right)^2 \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d, \tag{23}$$

$$S_{\rm Yukawa} = -\int \frac{2}{4!} Y_{ABC} \bar{\psi}^A \psi^B \phi^C \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d, \tag{24}$$

$$S_{\rm spin} = \int 2\pi i l_p^2 \bar{\psi}_A \gamma_5 \gamma^a \psi^A \varepsilon_{abcd} e^b \wedge e^c \wedge \beta^d, \tag{25}$$

$$S_{\rm CC} = -\int \frac{1}{96\pi l_p^2} \Lambda \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d. \tag{26}$$

While the form of the full action may appear quite complicated, one can recognize the meaning and purpose of each part of the action, as follows:

- the topological $3BF$ term (18) is identical to (16), tabulating all fields present in the theory (as dictated by the structure of the Standard Model 3-group),
- the gravitational constraint term (19) gives rise to the dynamics of the gravitational degrees of freedom,
- the scalar constraint (20) gives rise to the dynamics of massless scalar degrees of freedom,
- the Dirac constraint (21) gives rise to the dynamics of massless fermions,
- the Yang–Mills constraint (22) gives rise to the dynamics of massless gauge bosons,
- the Higgs potential constraint (23) contains the self-interactions and the mass of the Higgs field,
- the Yukawa constraint (24) contains the interactions between the Higgs field and fermions, as well as fermion mixing matrices,
- the spin constraint (25) is necessary for the appropriate coupling between fermion spins and torsion, and
- the CC constraint (26) introduces the cosmological constant.

The following free parameters are present in the action:

- l_p is the Planck length, featuring in $S_{\rm grav}$, $S_{\rm spin}$ and $S_{\rm CC}$,
- $C_{\alpha\beta}$ represents the gauge coupling constant bilinear form, featuring in $S_{\rm Yang\text{–}Mills}$,
- χ is the coupling constant for the quartic self-interaction of the Higgs field, featuring in $S_{\rm Higgs}$,
- v is the vacuum expectation value of the Higgs field, also featuring in $S_{\rm Higgs}$,
- Y_{ABC} represent the Yukawa couplings and fermion mixing matrices, featuring in $S_{\rm Yukawa}$, and
- Λ is the cosmological constant, featuring in $S_{\rm CC}$.

The topological part S_{3BF}^{top} and the constraints S_{scal} and S_{Dirac} do not contain any free parameters.

Let us discuss the equations of motion (EoMs) for this action. One can obtain the EoMs solved for all Lagrange multiplier fields, in terms of the dynamical fields and their derivatives (see for example [31, 51] for details):

$$
\begin{aligned}
M_{\alpha ab} &= -\frac{1}{48}\varepsilon_{abcd}F_\alpha{}^{\mu\nu}e^c{}_\mu e^d{}_\nu\,, & \zeta^{\alpha ab} &= \frac{1}{4}C_\beta{}^\alpha\varepsilon^{abcd}F^\beta{}_{\mu\nu}e_c{}^\mu e_d{}^\nu\,,\\
\lambda_{\alpha\mu\nu} &= -F_{\alpha\mu\nu}\,, & B_{\alpha\mu\nu} &= -\frac{e}{2}C_\alpha{}^\beta\varepsilon_{\mu\nu\rho\sigma}F_\beta{}^{\rho\sigma}\,,\\
\lambda_{[ab]\mu\nu} &= R_{[ab]\mu\nu}\,, & B_{[ab]\mu\nu} &= \frac{1}{8\pi\, l_p^2}\varepsilon_{[ab]cd}e^c{}_\mu e^d{}_\nu\,,\\
\tilde{\lambda}^A{}_\mu &= (\nabla_\mu\phi)^A\,, & \tilde{\gamma}^A{}_{\mu\nu\rho} &= -e\varepsilon_{\mu\nu\rho\sigma}\left(\nabla^\sigma\phi\right)^A\,,\\
H^{abcA} &= \frac{1}{6e}\varepsilon^{\mu\nu\rho\sigma}\left(\nabla_\mu\phi\right)^A e^a{}_\nu e^b{}_\rho e^c{}_\sigma\,, & \Lambda^{abA}{}_\mu &= \frac{1}{6e}g_{\mu\lambda}\varepsilon^{\lambda\nu\rho\sigma}\left(\nabla_\nu\phi\right)^A e^a{}_\rho e^b{}_\sigma\,,\\
\gamma^A{}_{\mu\nu\rho} &= -i\varepsilon_{abcd}e^a{}_\mu e^b{}_\nu e^c{}_\rho\left(\gamma^d\psi\right)^A\,, & \bar{\gamma}_{A\mu\nu\rho} &= i\varepsilon_{abcd}e^a{}_\mu e^b{}_\nu e^c{}_\rho\left(\bar{\psi}\gamma^d\right)_A\,,\\
\lambda^A{}_\mu &= \left(\overrightarrow{\nabla}_\mu\psi\right)^A\,, & \bar{\lambda}_{A\mu} &= \left(\bar{\psi}\overleftarrow{\nabla}_\mu\right)_A\,,\\
\beta^a{}_{\mu\nu} &= 0\,.
\end{aligned}
\tag{27}
$$

Next we look at the EoMs for the dynamical fields. The spin connection $\omega^{[ab]}{}_\mu$ is not equivalent to the Levi–Civita connection, since fermionic fields give rise to nonzero torsion. We therefore first split the spin connection into a sum of Ricci rotation coefficients $\Delta^{[ab]}{}_\mu$ and contorsion tensor $K^{[ab]}{}_\mu$:

$$\omega^{[ab]}{}_\mu = \Delta^{[ab]}{}_\mu + K^{[ab]}{}_\mu\,. \tag{28}$$

Here the Ricci rotation coefficients are given as

$$\Delta^{ab}{}_\mu = \frac{1}{2}\left(c^{abc} - c^{bac} - c^{cab}\right)e_{c\mu}\,, \tag{29}$$

where the commutation coefficients are defined as

$$c^{abc} = e^{b\mu}e^{c\nu}\left(\partial_\mu e^a{}_\nu - \partial_\nu e^a{}_\mu\right)\,. \tag{30}$$

The contorsion tensor is given as:

$$K^{ab}{}_\mu = \frac{1}{2}\left(T^{cab} + T^{bac} - T^{abc}\right)e_{c\mu}\,. \tag{31}$$

Here $T^{abc} \equiv T^a{}_{\mu\nu}e^{b\mu}e^{c\nu}$, where $T^a{}_{\mu\nu}$ are the components of the torsion 2-form, defined as:

$$T^a \equiv \nabla e^a = \frac{1}{2}T^a{}_{\mu\nu}\,\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\,, \qquad T^a{}_{\mu\nu} \equiv \nabla_\mu e^a{}_\nu - \nabla_\nu e^a{}_\mu\,. \tag{32}$$

Given all of the above quantities, one can write the EoM for torsion as:

$$T^a = 2\pi\, i\, l_p^2 s^a\,, \qquad s^a \equiv \varepsilon^{abcd}\,e_b\wedge e_c\,\bar{\psi}_A\gamma_5\gamma_d\psi^A\,. \tag{33}$$

We can see that the torsion 2-form is proportional to the spin 2-form s^a. As we shall see below, this is the same as in EC gravity. Also, using (31), the components of the contorsion 1-form are given as

$$K^{ab}{}_{\mu} = -2\pi i l_p^2 \bar{\psi}_A \gamma_5 \gamma_d \psi^A \varepsilon^{abcd} e_{c\mu} \,, \tag{34}$$

so the relationship between contorsion and torsion simplifies and we obtain:

$$T^a = K^{ab} \wedge e_b \,. \tag{35}$$

Next, the Einstein field equation has the usual form:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \Lambda g_{\mu\nu} = 8\pi l_p^2 T_{\mu\nu} \,, \tag{36}$$

where the stress-energy tensor is given as:

$$\begin{aligned} T_{\mu\nu} = {} & F^{\alpha}{}_{\mu\rho} C_{\alpha}{}^{\beta} F_{\beta\nu}{}^{\rho} - \frac{1}{4} g_{\mu\nu} F^{\alpha}{}_{\rho\sigma} C_{\alpha}{}^{\beta} F_{\beta}{}^{\rho\sigma} \\ & + \nabla_{\mu}\phi^A \nabla_{\nu}\phi_A - \frac{1}{2} g_{\mu\nu} \left(\nabla_{\rho}\phi^A \nabla^{\rho}\phi_A + 2\chi \left(\phi^A\phi_A - v^2 \right)^2 \right) \\ & + \frac{i}{2} \left(\bar{\psi}_A \overset{\leftrightarrow}{\nabla}_{\mu} \gamma_d \psi^A \right) e^d_{\nu} - \frac{1}{2} g_{\mu\nu} \left(i \left(\bar{\psi}_A \overset{\leftrightarrow}{\nabla}_{\rho} \gamma^d \psi^A \right) e_d{}^{\rho} - 2 Y_{ABC} \bar{\psi}^A \psi^B \phi^C \right) . \end{aligned} \tag{37}$$

It features three parts, describing the Yang–Mills, scalar and fermion stress-energy, respectively.

EoMs for fermion and scalar fields are

$$\left(i\gamma^{\mu} \overset{\rightarrow}{\nabla}_{\mu} \delta^A_B - Y^A{}_{BC} \phi^C \right) \psi^B = 0 \,, \tag{38}$$

$$\bar{\psi}_B \left(\delta^B_A i \overset{\leftarrow}{\nabla}_{\mu} \gamma^{\mu} + Y_{BAC} \phi^C \right) = 0 \,, \tag{39}$$

$$\nabla_{\mu}\nabla^{\mu}\phi^A - 4\chi \left(\phi^B\phi_B - v^2 \right) \phi^A = 0 \,, \tag{40}$$

while the EoM for Yang–Mills fields is:

$$\nabla_{\mu} F_{\alpha}{}^{\mu\nu} + \frac{1}{2} C^{-1}{}_{\alpha}{}^{\beta} \left(\rhd_{\beta AB} \left(\phi^A \nabla^{\nu} \phi^B - \phi^B \nabla^{\nu} \phi^A \right) + i \bar{\psi}_A \psi_B \left(\rhd_{\beta C}{}^{A} \gamma^{\nu CB} - \gamma^{\nu AC} \rhd_{\beta C}{}^{B} \right) \right) = 0 \,. \tag{41}$$

One can observe that all these EoMs correspond precisely to the Standard Model coupled to EC gravity, along with the cosmological constant term.

From the definition of the action (17) one can see that the full configuration space of the theory is defined over the non-dynamical Lagrange multiplier fields

$$\begin{aligned} & M_{\alpha ab} \,, \zeta^{\alpha ab} \,, \lambda_{\alpha\mu\nu} \,, B_{\alpha\mu\nu} \,, \lambda_{[ab]\mu\nu} \,, B_{[ab]\mu\nu} \,, \tilde{\lambda}^A{}_{\mu} \,, \tilde{\gamma}^A{}_{\mu\nu\rho} \,, H^{abcA} \,, \\ & \Lambda^{abA}{}_{\mu} \,, \gamma^A{}_{\mu\nu\rho} \,, \bar{\gamma}_{A\mu\nu\rho} \,, \lambda^A{}_{\mu} \,, \bar{\lambda}_{A\mu} \,, \beta^a{}_{\mu\nu} \,, \end{aligned} \tag{42}$$

as well as the dynamical fields

$$\omega^{[ab]}{}_{\mu} \,, e^a{}_{\mu} \,, \phi^A \,, \psi^A \,, \bar{\psi}_A \,, \alpha^{\alpha}{}_{\mu} \,. \tag{43}$$

The distinction between dynamical and non-dynamical fields is a consequence of the EoM, since the EoMs for the Lagrange multiplier fields are algebraic equations, while the EoMs for the dynamical fields are partial differential equations. One exception from this convention is the torsion equation (33), which can be explicitly solved for the spin connection as a function of the tetrads and fermion fields, rendering the spin connection as a non-dynamical field as well. This is a well known property in EC theory, but it is customary to regard the spin connection as a dynamical field nonetheless, because in more general theories in Riemann–Cartan spacetimes it often actually becomes a proper dynamical field [56, 57].

2.3. EC actions

The standard EC action is typically introduced in the literature (see for example [56]) as the sum of the actions for the Standard Model minimally coupled to gravity, and the Einstein–Hilbert action for the gravitational field (expressed using the tetrad formalism). In our notation, it reads:

$$\begin{aligned} S_{\mathrm{EC}}\left[e,\omega,\phi,\psi,\bar{\psi},\alpha\right] = \int & \frac{1}{16\pi l_p^2}\varepsilon^{abcd}R_{ab}\wedge e_c\wedge e_d - F^{\alpha}\wedge C_{\alpha\beta}\star F^{\beta} - (\nabla\phi)^A\wedge(\star\nabla\phi)_A \\ & -\frac{i}{6}\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\wedge\left[\left(\bar{\psi}\overleftarrow{\nabla}\right)_A\gamma^d\psi^A - \bar{\psi}_A\gamma^d\left(\overrightarrow{\nabla}\psi\right)^A\right] \\ & -\frac{1}{12}\left[\chi\left(\phi^A\phi_A - v^2\right)^2 + Y_{ABC}\bar{\psi}^A\psi^B\phi^C + \frac{\Lambda}{8\pi l_p^2}\right]\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\wedge e^d. \end{aligned} \tag{44}$$

Here $\star F$ denotes the Hodge dual of the 2-form F, and similarly for the 1-form $\nabla\phi$:

$$\star F^{\alpha} = \frac{1}{4}F^{\alpha}{}_{cd}\varepsilon^{abcd}e_a\wedge e_b, \qquad (\star\nabla\phi)_A = \frac{1}{3!}\left(\nabla_d\phi\right)_A\varepsilon^{dabc}e_a\wedge e_b\wedge e_c. \tag{45}$$

The configuration space of this theory is equivalent to the configuration space of the dynamical fields of the 3BF theory (43), where the spin connection again satisfies the algebraic equation

$$\omega_{ab\mu} = \Delta_{ab\mu} - 2\pi i l_p^2\bar{\psi}_A\gamma_5\gamma^d\psi^A\varepsilon_{abcd}e^c{}_{\mu}, \tag{46}$$

obtained from the equation of motion for torsion analogous to equation (33), so it does not represent a dynamical field. Because of this, by substituting this algebraic relation back into the action, it is possible to construct an equivalent classical theory with the reduced configuration space and equivalent EoM. Such equivalent theory is usually called the *second order theory* in the literature [56]. This substitution is performed by explicitly partitioning the spin connection into the contorsion and Ricci coefficients in the action, and by separating the contributions of the individual terms. This operation is equivalent to the following substitution:

$$\nabla_{(\omega)} = \nabla_{(\Delta)} + \frac{1}{2}K^{ab}\sigma_{ab}. \tag{47}$$

In the action of the EC theory there are only two terms which depend on the spin connection, namely the term with the Riemann tensor and the Dirac Lagrangian term. After the substitution, these terms become:

$$\begin{aligned} R_{ab} &= \mathrm{d}\omega_{ab} + (\omega\wedge\omega)_{ab} = \mathrm{d}\Delta_{ab} + (\Delta\wedge\Delta)_{ab} + \mathrm{d}K_{ab} + (\Delta\wedge K)_{ab} + (K\wedge\Delta)_{ab} + K_a{}^c\wedge K_{cb} \\ &= R_{ab}(\Delta) + \left(\nabla_{(\Delta)}K\right)_{ab} + K_a{}^c\wedge K_{cb}, \end{aligned} \tag{48}$$

and

$$\left(\bar{\psi}\overleftarrow{\nabla}\right)_A \gamma^d \psi^A - \bar{\psi}_A \gamma^d (\nabla\psi)^A = \left(\bar{\psi}\overleftarrow{\nabla}_{(\Delta)}\right)_A \gamma^d \psi^A - \bar{\psi}_A \gamma^d \left(\nabla_{(\Delta)}\psi\right)^A - \frac{1}{2}\bar{\psi}_A \left\{\sigma^{ab}, \gamma^d\right\} \psi^A K_{ab}$$
$$= \left(\bar{\psi}\overleftarrow{\nabla}_{(\Delta)}\right)_A \gamma^d \psi^A - \bar{\psi}_A \gamma^d \left(\nabla_{(\Delta)}\psi\right)^A + \frac{1}{2}\varepsilon^{abcd}\bar{\psi}_A \gamma_5 \gamma_c \psi^A K_{ab} \quad (49)$$

When these terms are substituted back into the standard EC action, after some algebra one obtains the action of the same form as the initial one, and with an extra term representing the contact spin–spin interaction and the fixed spin connection Δ_{ab}. This extra term, expressed using contorsion, is $\frac{1}{8\pi l_p^2} K^{ab} \wedge \star K_{ab}$, so using the relation between the spin tensor and contorsion from the standard EC theory one can eliminate the contorsion from the action:

$$S_{\text{ECC}}\left[e, \phi, \psi, \bar{\psi}, \alpha\right] = \int \frac{1}{16\pi l_p^2}\varepsilon^{abcd} R_{ab} \wedge e_c \wedge e_d - F^\alpha \wedge C_{\alpha\beta} \star F^\beta - (\nabla\phi)^A \wedge (\star\nabla\phi)_A$$
$$- \frac{i}{6}\varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge \left[\left(\bar{\psi}\overleftarrow{\nabla}\right)_A \gamma^d \psi^A - \bar{\psi}_A \gamma^d \left(\overrightarrow{\nabla}\psi\right)^A\right] + \frac{3}{\left(4\pi l_p^2\right)^3} s^a \wedge \star s_a$$
$$- \frac{1}{12}\left[\chi\left(\phi^A \phi_A - v^2\right)^2 + Y_{ABC}\bar{\psi}^A \psi^B \phi^C + \frac{\Lambda}{8\pi l_p^2}\right]\varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d. \quad (50)$$

In this fashion, one obtains the EC contact action (ECC), with a fourth degree contact interaction terms between fermions, featuring within the term $s^a \wedge \star s_a$. The EoM of this theory are equivalent to the EoM for the standard EC theory and the constrained $3BF$ theory, but the configuration space is further reduced and equal to the dynamical configuration space:

$$e^a{}_\mu, \phi^A, \psi^A, \bar{\psi}_A, \alpha^\alpha{}_\mu. \quad (51)$$

This concludes the review of the topological and constrained $3BF$ actions, as well as the EC and ECC actions, regarded as classical theories. In what follows, we will focus our attention to the constrained $3BF$ action (from now on just called $3BF$ action for simplicity), and the ECC action. Specifically, in the next section we will explore the relationship between their corresponding quantum theories, namely the quantized $3BF$ theory and the quantized ECC theory. This will be done within in a framework which consistently defines the quantization of both theories in the same way, so that they can be compared.

3. Quantum observables

Although the $3BF$ and ECC theories are classically equivalent, their quantization may give rise to potentially different quantum theories. In this section we will establish a correspondence between the expectation values of observables in the two quantum theories, demonstrating that although the theories are not fully equivalent at the quantum level, there exists a well-defined correspondence between them.

The process of construction of quantum theories requires the generalization of mathematical results related to multiple integrals over real and Grassmann numbers to the corresponding functional integrals over bosonic and fermionic fields. Therefore, in the first Subsection we will provide a review of the integrals that will be generalized to functional level. In the second Subsection we will apply those integrals to the definition of the expectation value of an arbitrary observable in the quantum $3BF$ theory, step by step, so that the obtained result can be interpreted as the expectation value of a related observable in the quantum ECC theory. In this way, we will construct a correspondence between the two quantum theories, at the full nonperturbative level.

3.1. Mathematical preliminaries

Let us begin from the definitions of certain mathematical identities. We will discuss a total of four generalizations of the properties of the Dirac delta function in several special cases, as well as one identity related to the Stokes theorem. The identities can be classified into two groups, bosonic and fermionic. The two bosonic identities can be obtained by generalizing the following properties of the Dirac delta function to the bosonic fields. The first identity is

$$\int_{\mathbb{R}} \mathrm{d}y \mathrm{e}^{\mathrm{i}yF} = 2\pi\,\delta\left(F\right), \qquad F \in \mathbb{R}, \tag{52}$$

based on which one can obtain the following multiple integral:

$$\int_{\mathbb{R}} \mathrm{d}y \int_{\mathbb{R}^n} \mathrm{d}x_k H\left(x_k\right) \mathrm{e}^{\mathrm{i}\left(yF\left(x_k\right)+G\left(x_k\right)\right)} = 2\pi \int_{\mathbb{R}^n} \mathrm{d}x_k H\left(x_k\right)\delta\left(F\left(x_k\right)\right) \mathrm{e}^{\mathrm{i}G\left(x_k\right)}. \tag{53}$$

Then, generalizing this identity to the functional level, one obtains:

$$\int D\varphi D\phi_k H\left(\phi_k\right) \mathrm{e}^{\mathrm{i}\int\left(\varphi\wedge F\left(\phi_k\right)+G\left(\phi_k\right)\right)} = \mathcal{N}\int D\phi_k H\left(\phi_k\right)\delta\left(F\left(\phi_k\right)\right)\mathrm{e}^{\mathrm{i}\int G\left(\phi_k\right)}. \tag{54}$$

The second identity that we need is:

$$\int_{\mathbb{R}} \mathrm{d}y\delta\left(yF-G\right)H\left(y\right) = \frac{H\left(G/F\right)}{|F|}, \qquad F, G \in \mathbb{R}, \tag{55}$$

which can be easily proved using a simple change of variables. The corresponding multiple integral is

$$\int_{\mathbb{R}} \mathrm{d}y \int_{\mathbb{R}^n} \mathrm{d}x_k \delta\left(y^{aB}F_B{}^A\left(x_k\right) - G^{aA}\left(x_k\right)\right) H\left(y, x_k\right) = \int_{\mathbb{R}^n} \mathrm{d}x_k \frac{H\left(G^{aB}\left(x_k\right)F^{-1}{}_B{}^A\left(x_k\right), x_k\right)}{\left|F\left(x_k\right)\right|^{|a|}}, \tag{56}$$

while the generalization to the functional level is given as:

$$\int D\varphi D\phi_k \delta\left(\varphi^{aB}F_B{}^A\left(\phi_k\right) - G^{aA}\left(\phi_k\right)\right) H\left(\varphi, \phi_k\right) = \int D\phi_k \frac{1}{\left|F\left(\phi_k\right)\right|^{|a|}} H\left(G^{aB}\left(\phi_k\right)F^{-1}{}_B{}^A\left(\phi_k\right), \phi_k\right), \tag{57}$$

where $F_B{}^A(\phi_k)$ is an arbitrary invertible matrix function, $G^{aA}(\phi_k)$ and $H(\varphi, \phi_k)$ are arbitrary functions, and $|a|$ denotes the number of possible values of the index a.

Next, we need an identity related to the Stokes theorem, namely

$$\int D\varphi D\phi_k H(\varphi, \phi_k) \mathrm{e}^{\mathrm{i}\int(\nabla\varphi)\wedge E(\phi_k)+\varphi\wedge F(\phi_k)+G(\phi_k)} = \int D\varphi_\partial D\phi_{k\partial} \mathrm{e}^{\mathrm{i}\oint \varphi_\partial\wedge E(\phi_{k\partial})} \times \int D\varphi D\phi_k H(\varphi, \phi_k) \mathrm{e}^{\mathrm{i}\int(-1)^{p-1}\varphi\wedge\nabla E(\phi_k)+\varphi\wedge F(\phi_k)+G(\phi_k)}, \tag{58}$$

which holds equally for both bosonic and fermionic fields. Here it is assumed that φ is a p-form, while φ_∂ and $\phi_{k\partial}$ are values of the fields on the integration boundary. Note that the sole purpose of this identity is to move the covariant derivative ∇ from acting on φ to acting on $E(\phi_k)$.

Finally, the two fermionic identities which can be equivalently generalized from Grassmann numbers to Grassmann fields are given as

$$\int_{\mathbb{G}^n} \mathrm{d}\theta_1 \mathrm{d}\theta_2 \dots \mathrm{d}\theta_n \, \mathrm{e}^{\mathrm{i}\theta_1(\theta_2-\theta_3-\dots-\theta_k)} \, F(\theta_2,\dots,\theta_n)$$
$$= (-1)^{n-1} \, i \int_{\mathbb{G}^{n-1}} \mathrm{d}\theta_2 \dots \mathrm{d}\theta_n \, \delta(\theta_2-\theta_3-\dots-\theta_k) \, F(\theta_2,\dots,\theta_n), \tag{59}$$

and

$$\int_{\mathbb{R}^m} \mathrm{d}^m y \int_{\mathbb{R}^m} \mathrm{d}^m x \int_{\mathbb{G}^k} \mathrm{d}\theta_1 \mathrm{d}\theta_2 \dots \mathrm{d}\theta_k \int_{\mathbb{G}^{n-k}} \mathrm{d}\theta_{k+1} \dots \mathrm{d}\theta_n \, \mathrm{e}^{\mathrm{i}y_a\left(x^a - M^{aij}\theta_i\theta_j\right)} \, F(x,\theta_1,\dots,\theta_n)$$
$$= (2\pi)^m \int_{\mathbb{R}^m} \mathrm{d}^m x \int_{\mathbb{G}^k} \mathrm{d}\theta_1 \mathrm{d}\theta_2 \dots \mathrm{d}\theta_k \int_{\mathbb{G}^{n-k}} \mathrm{d}\theta_{k+1} \dots \mathrm{d}\theta_n \prod_{a=1}^{m} \delta\left(x^a - M^{aij}\theta_i\theta_j\right) \, F(x,\theta_1,\dots,\theta_n). \tag{60}$$

For proof see appendix A. The corresponding functional identities for the Grassmann fields look the same as the above two identities, up to the replacement of the integration measure $\mathrm{d}\theta \to D\theta$, the overall coefficient $(2\pi)^m \to \mathcal{N}$, and the notation $\prod \delta(x^a) \to \delta(\phi)$. One should note that the last identity contains the Dirac delta function with a combination of real and Grassmann numbers where the order of integration is important, namely, one must first integrate over the real numbers x^a, and only afterwards over the Grassmann numbers. Also, the Dirac delta function for the Grassmann numbers is odd (i.e. skew-Hermitian) because of the anticommutativity of the Grassmann numbers, so it always appears paired with an imaginary unit $i \equiv \sqrt{-1}$ in equation (59).

3.2. Expectation values of the observables

The correspondence between two quantum theories can be obtained by comparing the expectation values of the observables between the two theories. To that end, one defines the expectation values as

$$\langle F \rangle_{3BF} = \frac{1}{Z_{3BF}} \int D\phi_i \; F(\phi_k) \; \mathrm{e}^{\mathrm{i}S_{3BF}[\phi_i]}, \quad \langle F \rangle_{\mathrm{ECC}} = \frac{1}{Z_{\mathrm{ECC}}} \int D\phi_i \; F(\phi_k) \; \mathrm{e}^{\mathrm{i}S_{\mathrm{ECC}}[\phi_i]}, \tag{61}$$

where the state sums are given as

$$Z_{3BF} = \int D\phi_i \; \mathrm{e}^{\mathrm{i}S_{3BF}[\phi_i]}, \qquad Z_{\mathrm{ECC}} = \int D\phi_i \; \mathrm{e}^{\mathrm{i}S_{\mathrm{ECC}}[\phi_i]}. \tag{62}$$

Here at the beginning it is important to make two comments. First, for the purpose of subsequent analysis one does not need to discuss the precise definition of the path integral itself, which means that we do not need to specify the quantum $3BF$ and ECC theories explicitly. The only requirement that we assume is that the measures in the path integrals are defined in the same way in both theories, and that they are defined such that the functional identities (54), (57), (58), (59) and (60) hold. In this way, we can discuss the properties and compare the expectation values of the observables in two quantum theories in the full nonperturbative regime, despite the fact that we do not specify the details of the quantizations of the two theories.

Second, the fields ϕ_i for $3BF$ theory and ECC theory belong to their corresponding configurations spaces (42)–(43) and (51), respectively. It should be clear that the observables $F(\phi_k)$ in (61) can be compared only if they both live in the common configuration subspace of both theories, or in other words, if the observable F depends only on the fields ϕ_k belonging to the reduced configuration subspace (51) which is defined for the ECC theory.

In what follows, we will apply the functional identities (54), (57), (58), (59) and (60), step by step, in order to reduce the definition of the expectation value of the observable F in quantum $3BF$ theory, to the definition of the expectation value of the corresponding observable in quantum EC theory, and then in the quantum ECC theory.

We begin by explicitly writing the product of the state sum and the expectation value of the observable F in quantum $3BF$ theory. According to the definition (61), it is equal to:

$$
\begin{aligned}
Z_{3BF}\langle F\rangle_{3BF} = \mathcal{N}\int & D\alpha D\omega D\beta De D\tilde{\gamma} D\gamma D\bar{\gamma} D\phi D\psi D\bar{\psi} DBD\tilde{\lambda}D\lambda D\bar{\lambda}D\Lambda D\zeta DHDM \\
& \times \exp\left(i\int B_\alpha\wedge F^\alpha + B_{[ab]}\wedge R^{[ab]} + e_a\wedge(\nabla\beta)^a + \phi^A\left(\nabla\tilde{\gamma}\right)_A + \bar{\psi}_A\left(\nabla\gamma\right)^A - \left(\bar{\gamma}\overleftarrow{\nabla}\right)_A\psi^A \right. \\
& + \lambda^\alpha\wedge\left(B_\alpha - 12C_{\alpha\beta}M^\beta{}_{ab}e^a\wedge e^b\right) - \lambda_{[ab]}\wedge\left(B^{[ab]} - \frac{1}{8\pi l_p^2}\varepsilon^{[ab]cd}e_c\wedge e_d\right) \\
& + \tilde{\lambda}^A\wedge\left(\tilde{\gamma}_A - H_{abcA}e^a\wedge e^b\wedge e^c\right) + \bar{\lambda}_A\wedge\left(\gamma^A + \frac{i}{6}\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\left(\gamma^d\psi\right)^A\right) \\
& - \lambda^A\wedge\left(\bar{\gamma}_A - \frac{i}{6}\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\left(\bar{\psi}\gamma^d\right)_A\right) + 2\pi i\, l_p^2\bar{\psi}_A\gamma_5\gamma^a\psi^A\varepsilon_{abcd}e^b\wedge e^c\wedge\beta^d \\
& + \zeta_\alpha{}^{ab}\left(M^\alpha{}_{ab}\varepsilon_{cdef}e^c\wedge e^d\wedge e^e\wedge e^f - F^\alpha\wedge e_a\wedge e_b\right) \\
& + \Lambda^{abA}\wedge\left(H_{abcA}\varepsilon^{cdef}e_d\wedge e_e\wedge e_f - (\nabla\phi)_A\wedge e_a\wedge e_b\right) \\
& \left. - \frac{1}{12}\left(\chi\left(\phi^A\phi_A - v^2\right)^2 + Y_{ABC}\bar{\psi}^A\psi^B\phi^C + \frac{\Lambda}{8\pi l_p^2}\right)\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\wedge e^d\right) F(\phi_k)\,. \qquad (63)
\end{aligned}
$$

Applying the Stokes theorem functional identity (58) over the terms in the second row, we remove the covariant derivative from the variables β_a, $\bar{\gamma}_A$, γ^A and $\tilde{\gamma}^A$, in order to be able to perform the integration over these variables in the next step. We obtain:

$$
\begin{aligned}
Z_{3BF}\langle F\rangle_{3BF} = \mathcal{N}\int & D\beta_\partial De_\partial D\tilde{\gamma}_\partial D\gamma_\partial D\bar{\gamma}_\partial D\phi_\partial D\psi_\partial D\bar{\psi}_\partial\, e^{i\oint \phi^A_\partial\tilde{\gamma}_{A\partial} + \bar{\psi}_{A\partial}\gamma^A_\partial + \bar{\gamma}_{A\partial}\psi^A_\partial - e_{a\partial}\wedge\beta^a_\partial} \\
& \times\int D\alpha D\omega D\beta De D\tilde{\gamma} D\gamma D\bar{\gamma} D\phi D\psi D\bar{\psi} DBD\tilde{\lambda}D\lambda D\bar{\lambda}D\Lambda D\zeta DHDM \\
& \times \exp\left(i\int B_\alpha\wedge F^\alpha + B_{[ab]}\wedge R^{[ab]} + (\nabla e)_a\wedge\beta^a - (\nabla\phi)^A\wedge\tilde{\gamma}_A - \left(\bar{\psi}\overleftarrow{\nabla}\right)_A\wedge\gamma^A + \bar{\gamma}_A\wedge(\nabla\psi)^A \right. \\
& + \lambda^\alpha\wedge\left(B_\alpha - 12C_{\alpha\beta}M^\beta{}_{ab}e^a\wedge e^b\right) - \lambda_{[ab]}\wedge\left(B^{[ab]} - \frac{1}{8\pi l_p^2}\varepsilon^{[ab]cd}e_c\wedge e_d\right) \\
& + \tilde{\lambda}^A\wedge\left(\tilde{\gamma}_A - H_{abcA}e^a\wedge e^b\wedge e^c\right) + \bar{\lambda}_A\wedge\left(\gamma^A + \frac{i}{6}\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\left(\gamma^d\psi\right)^A\right) \\
& - \lambda^A\wedge\left(\bar{\gamma}_A - \frac{i}{6}\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\left(\bar{\psi}\gamma^d\right)_A\right) + 2\pi i\, l_p^2\bar{\psi}_A\gamma_5\gamma^a\psi^A\varepsilon_{abcd}e^b\wedge e^c\wedge\beta^d \\
& + \zeta_\alpha{}^{ab}\left(M^\alpha{}_{ab}\varepsilon_{cdef}e^c\wedge e^d\wedge e^e\wedge e^f - F^\alpha\wedge e_a\wedge e_b\right) \\
& + \Lambda^{abA}\wedge\left(H_{abcA}\varepsilon^{cdef}e_d\wedge e_e\wedge e_f - (\nabla\phi)_A\wedge e_a\wedge e_b\right) \\
& \left. - \frac{1}{12}\left(\chi\left(\phi^A\phi_A - v^2\right)^2 + Y_{ABC}\bar{\psi}^A\psi^B\phi^C + \frac{\Lambda}{8\pi l_p^2}\right)\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\wedge e^d\right) F(\phi_k)\,. \qquad (64)
\end{aligned}
$$

By careful inspection of the above equation one can note that, in the case when the observable $F(\phi_k)$ does not depend on the fields at the boundary, the integration over the boundary gives no contribution at all. Also, formulated in this way, the $3BF$ theory enforces the restrictions that the matter fields and tetrad fields must be zero on the manifold boundary. These restrictions can be removed by adding appropriate boundary terms to the classical $3BF$ action. The detailed discussion of the boundary conditions is given in section 3.3 below.

Next, we perform the integration over the fields β_a, $B_{[ab]}$, B_α, $\tilde{\gamma}^A$, $\bar{\gamma}_A$, γ_A, ${\zeta_\alpha}^{ab}$ and Λ^{abA} by applying the functional identities (54) and (59), which gives:

$$
\begin{aligned}
Z_{3BF}\langle F\rangle_{3BF} = \mathcal{N}\int & De_\partial D\phi_\partial D\psi_\partial D\bar{\psi}_\partial\, \delta(\phi_\partial)\,\delta(\psi_\partial)\,\delta(\bar{\psi}_\partial)\,\delta(e_\partial) \\
&\times \int D\alpha D\omega De D\phi D\psi D\bar{\psi} D\tilde{\lambda} D\lambda D\bar{\lambda} DHDM \\
&\times \delta(F^\alpha + \lambda^\alpha)\,\delta\left(R_{[ab]} - \lambda_{[ab]}\right)\delta\left(\tilde{\lambda}_A - (\nabla\phi)_A\right)\delta\left(\bar{\lambda}_A - \left(\bar{\psi}\overleftarrow{\nabla}\right)_A\right)\delta\left(\lambda^A - (\nabla\psi)^A\right) \\
&\times \exp\left(i\int -12\lambda^\alpha \wedge C_{\alpha\beta}{M^\beta}_{ab} e^a \wedge e^b + \lambda_{[ab]} \wedge \frac{1}{8\pi l_p^2}\varepsilon^{[ab]cd} e_c \wedge e_d - \tilde{\lambda}^A \wedge H_{abcA} e^a \wedge e^b \wedge e^c\right. \\
&+ \bar{\lambda}_A \wedge \frac{i}{6}\varepsilon_{abcd} e^a \wedge e^b \wedge e^c \left(\gamma^d\psi\right)^A + \lambda^A \wedge \frac{i}{6}\varepsilon_{abcd} e^a \wedge e^b \wedge e^c \left(\bar{\psi}\gamma^d\right)_A \\
&\left. - \frac{1}{12}\left(\chi\left(\phi^A\phi_A - v^2\right)^2 + Y_{ABC}\bar{\psi}^A\psi^B\phi^C + \frac{\Lambda}{8\pi l_p^2}\right)\varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d\right) \\
&\times \delta\left({M^\alpha}_{ab}\varepsilon_{cdef} e^c \wedge e^d \wedge e^e \wedge e^f - F^\alpha \wedge e_a \wedge e_b\right) \\
&\times \delta\left(H_{abcA}\varepsilon^{cdef} e_d \wedge e_e \wedge e_f - (\nabla\phi)_A \wedge e_a \wedge e_b\right) \\
&\times \delta\left((\nabla e)_a - 2\pi i l_p^2 \bar{\psi}_A \gamma_5 \gamma^d \psi^A \varepsilon_{abcd} e^b \wedge e^c\right) F(\phi_k)\,.
\end{aligned} \tag{65}
$$

Then, further integration over λ^α, $\lambda_{[ab]}$, $\tilde{\lambda}_A$, $\bar{\lambda}_A$ and λ^A removes the Dirac delta terms from the third row, thus giving the following:

$$
\begin{aligned}
Z_{3BF}\langle F\rangle_{3BF} = \mathcal{N}\int & De_\partial D\phi_\partial D\psi_\partial D\bar{\psi}_\partial\, \delta(\phi_\partial)\,\delta(\psi_\partial)\,\delta(\bar{\psi}_\partial)\,\delta(e_\partial) \\
&\times \int D\alpha D\omega De D\phi D\psi D\bar{\psi} DHDM \\
&\times \exp\left(i\int 12F^\alpha \wedge C_{\alpha\beta}{M^\beta}_{ab} e^a \wedge e^b + R_{ab} \wedge \frac{1}{16\pi l_p^2}\varepsilon^{abcd} e_c \wedge e_d - (\nabla\phi)^A \wedge H_{abcA} e^a \wedge e^b \wedge e^c\right. \\
&\left. - \frac{1}{12}\left(\chi\left(\phi^A\phi_A - v^2\right)^2 + Y_{ABC}\bar{\psi}^A\psi^B\phi^C + \frac{\Lambda}{8\pi l_p^2}\right)\varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d\right) \\
&\times \delta\left({M^\alpha}_{ab}\varepsilon_{cdef} e^c \wedge e^d \wedge e^e \wedge e^f - F^\alpha \wedge e_a \wedge e_b\right) \\
&\times \delta\left(H_{abcA}\varepsilon^{cdef} e_d \wedge e_e \wedge e_f - (\nabla\phi)_A \wedge e_a \wedge e_b\right) \\
&\times \delta\left(2\pi i l_p^2 \bar{\psi}_A \gamma_5 \gamma^d \psi^A \varepsilon_{abcd} e^b \wedge e^c - (\nabla e)_a\right) F(\phi_k)
\end{aligned} \tag{66}
$$

The multipliers ${M^\alpha}_{ab}$ and H_{abcA}, which have so far not been integrated over, are related to the Hodge duals of the field strengths F^α and $(\nabla\phi)_A$, respectively. Applying the functional identity (57) one can integrate out also these remaining multipliers, in favor of the Hodge duals (45), giving:

$$
\begin{aligned}
Z_{3BF}\langle F\rangle_{3BF} = \mathcal{N} \int & De_\partial D\phi_\partial D\psi_\partial D\bar{\psi}_\partial \, \delta\left(\phi_\partial\right) \delta\left(\psi_\partial\right) \delta\left(\bar{\psi}_\partial\right) \delta\left(e_\partial\right) \\
& \times \int D\alpha D\omega De D\phi D\psi D\bar{\psi} \frac{1}{|e|^{(|\alpha|+|A|(D-1))|[ab]|}} \\
& \times \exp\left(i \int -F^\alpha \wedge C_{\alpha\beta} \star F^\beta + R_{ab} \wedge \frac{1}{16\pi l_p^2} \varepsilon^{abcd} e_c \wedge e_d - (\nabla\phi)^A \wedge (\star\nabla\phi)_A \right. \\
& - \frac{i}{6} \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge \left(\left(\bar{\psi} \overleftarrow{\nabla} \right)_A \gamma^d \psi^A - \bar{\psi}_A \gamma^d (\nabla\psi)^A \right) \\
& \left. - \frac{1}{12} \left(\chi \left(\phi^A \phi_A - v^2 \right)^2 + Y_{ABC} \bar{\psi}^A \psi^B \phi^C + \frac{\Lambda}{8\pi l_p^2} \right) \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d \right) \\
& \times \delta\left(2\pi i l_p^2 \bar{\psi}_A \gamma_5 \gamma^d \psi^A \varepsilon_{abcd} e^b \wedge e^c - (\nabla e)_a \right) F(\phi_k) \, . \qquad (67)
\end{aligned}
$$

The quantity D (which appears in the exponent of the determinant of the tetrad) is the dimension of the spacetime manifold, $D = 4$, so the value of the exponent of the determinant of the tetrad is $N = (|\alpha| + |A|(D-1))|[ab]| = 144$, taking into account that $|\alpha| = 12$, $|A| = 4$ and $|[ab]| = 6$. One can recognize that the obtained argument of the exponent is now the action (44) of the EC theory, so we can write:

$$
\begin{aligned}
Z_{3BF}\langle F\rangle_{3BF} = \mathcal{N} \int & De_\partial D\phi_\partial D\psi_\partial D\bar{\psi}_\partial \, \delta\left(\phi_\partial\right) \delta\left(\psi_\partial\right) \delta\left(\bar{\psi}_\partial\right) \delta\left(e_\partial\right) \\
& \times \int D\alpha D\omega De D\phi D\psi D\bar{\psi} \, \frac{1}{|e|^N} \, \delta\left(2\pi i l_p^2 \bar{\psi}_A \gamma_5 \gamma^d \psi^A \varepsilon_{abcd} e^b \wedge e^c - (\nabla e)_a \right) \\
& \times F(\phi_k) \, \mathrm{e}^{\mathrm{i} S_{\mathrm{EC}}[\phi_k]} \, . \qquad (68)
\end{aligned}
$$

At this point, one could be tempted to try to establish a correspondence between the quantum $3BF$ theory and the quantum EC theory. Unfortunately, this is not a viable option since the Dirac delta term under the integral enforces an additional strong constraint between torsion and the spin tensor, and must be integrated out before one can attempt to establish the correspondence. This can be done by integrating out the spin connection 1-form $\omega^{[ab]}$, which is present inside the Dirac delta term as part of the covariant derivative ∇ acting on the tetrad 1-form. In order to perform this integration, the expression inside the Dirac delta term must be transformed, since in its original form it depends on the antisymmetric part of the spin connection over the second index and the spacetime index. This dependence can be removed by passing to the locally inertial coordinate system, where this antisymmetric piece can be evaluated, by introducing a change of variables $\omega_{abc} = \omega_{ab\mu} e_c{}^\mu$. This change of variables induces the following change of the path integral measure:

$$
D\omega_{ab\mu} = D\omega_{abc} \left| \frac{\delta\left(\omega_{abc} e^c{}_\mu\right)}{\delta\omega_{efg}} \right| = D\omega_{abc} \left| \delta^{[ef]}_{[ab]} e^g{}_\mu \right| = D\omega_{abc} |e|^{|[ab]|} \, . \qquad (69)
$$

Now one can introduce the quantity A_{abc} which is antisymmetric with respect to the second and third indices of the spin connection:

$$
A_{abc} = \frac{1}{2} \left(\omega_{abc} - \omega_{acb} \right) \, . \qquad (70)
$$

One can easily demonstrate that these fields contain all components of the spin connection. To see this, it is sufficient to apply the antisymmetry of the spin connection with respect to the first two indices onto the following linear combination:

$$A_{abc} - A_{bac} - A_{cab} = \frac{1}{2}\left(\omega_{abc} - \omega_{acb} - \omega_{bac} + \omega_{bca} - \omega_{cab} + \omega_{cba}\right) = \omega_{abc}\,. \quad (71)$$

The Jacobian $\mathcal{J}$ of this change of variables is a constant, and it can be absorbed in the normalization factor $\mathcal{N}$, so the path integral measure remains essentially the same:

$$\mathcal{N}\int D\omega_{abc} = \mathcal{N}\int |\mathcal{J}| DA_{abc} = \mathcal{N}'\int DA_{abc}\,. \quad (72)$$

Substituting this back into (68) we obtain the expression which depends on the fields A_{abc}:

$$Z_{3BF}\langle F\rangle_{3BF} = \mathcal{N}\int De_{\partial}D\phi_{\partial}D\psi_{\partial}D\bar{\psi}_{\partial}\,\delta(\phi_{\partial})\,\delta(\psi_{\partial})\,\delta(\bar{\psi}_{\partial})\,\delta(e_{\partial})\int D\alpha DADeD\phi D\psi D\bar{\psi}\,\frac{|e|^{|[ab]|}}{|e|^{N}}$$
$$\times\,\delta\left(2\pi i l_p^2\bar{\psi}_A\gamma_5\gamma^d\psi^A\varepsilon_{abcd}e^b{}_\mu e^c{}_\nu\varepsilon^{\mu\nu\rho\sigma} - (\partial_\mu e_{a\nu})\varepsilon^{\mu\nu\rho\sigma} + A_{abc}|e|e_d{}^\rho e_e{}^\sigma\varepsilon^{bcde}\right)$$
$$\times\,F(\phi_k)\,\mathrm{e}^{\mathrm{i}S_{\mathrm{EC}}[\phi_k]}\,. \quad (73)$$

Applying the functional identities (53) and (56) we then obtain the expression which can be integrated over the fields A_{abc} in a straightforward manner:

$$Z_{3BF}\langle F\rangle_{3BF} = \mathcal{N}\int De_{\partial}D\phi_{\partial}D\psi_{\partial}D\bar{\psi}_{\partial}\,\delta(\phi_{\partial})\,\delta(\psi_{\partial})\,\delta(\bar{\psi}_{\partial})\,\delta(e_{\partial})\int D\alpha DADeD\phi D\psi D\bar{\psi}\,\frac{|e|^{|[ab]|}}{|e|^{N}}$$
$$\times\,\delta\left(A_{abc} - \left(\frac{1}{2}c_{abc} - 2\pi i l_p^2\bar{\psi}_A\gamma_5\gamma^d\psi^A\varepsilon_{abcd}\right)\right)\frac{1}{\left|2|e|\varepsilon^{[bc][de]}e_{[d}{}^{[\rho}e_{e]}{}^{\sigma]}\right|^{|a|}}$$
$$\times\,F(\phi_k)\,\mathrm{e}^{\mathrm{i}S_{\mathrm{EC}}[\phi_k]}\,. \quad (74)$$

The next step is the evaluation of the determinant of the product of two tetrad fields and the Levi–Civita tensor. One first evaluates the determinants of the Levi–Civita tensor as $\left|\varepsilon^{[ab][cd]}\right| = \left|\varepsilon^{[\mu\nu][\rho\sigma]}\right| = 1$, and then one constructs the identity:

$$\frac{1}{|e|^{|[\mu\nu]|}} = \left|\frac{1}{|e|}\varepsilon^{[\mu\nu][\rho\sigma]}\right| = \left|e_a{}^{[\mu}e_b{}^{\nu]}e_c{}^{[\rho}e_d{}^{\sigma]}\varepsilon^{abcd}\right|$$
$$= \left|2e_{[a}{}^{[\mu}e_{b]}{}^{\nu]}\right|\left|2e_{[c}{}^{[\rho}e_{d]}{}^{\sigma]}\right|\left|\varepsilon^{[ab][cd]}\right| = \left|2e_{[a}{}^{[\mu}e_{b]}{}^{\nu]}\right|^2, \quad (75)$$

from which it follows that

$$\left|2e_{[a}{}^{[\mu}e_{b]}{}^{\nu]}\right| = \frac{1}{|e|^{\frac{1}{2}|[\mu\nu]|}} = \frac{1}{|e|^{\frac{1}{2}|[ab]|}}\,. \quad (76)$$

Using the obtained relation (76) the path integral becomes

$$Z_{3BF}\langle F\rangle_{3BF} = \mathcal{N}\int De_{\partial}D\phi_{\partial}D\psi_{\partial}D\bar{\psi}_{\partial}\,\delta(\phi_{\partial})\,\delta(\psi_{\partial})\,\delta(\bar{\psi}_{\partial})\,\delta(e_{\partial})\int D\alpha DADeD\phi D\psi D\bar{\psi}$$
$$\times\,\delta\left(A_{abc} - \left(\frac{1}{2}c_{abc} - 2\pi i l_p^2\bar{\psi}_A\gamma_5\gamma^d\psi^A\varepsilon_{abcd}\right)\right)\frac{1}{|e|^{N+|[ab]|\left(\frac{|a|}{2}-1\right)}}F(\phi_k)\,\mathrm{e}^{\mathrm{i}S_{\mathrm{EC}}[\phi_k]}\,. \quad (77)$$

In the case of the Standard Model, the exponent of the determinant of the tetrad is given as $M = N + |[ab]| \left(\frac{|a|}{2} - 1\right) = N + 6 = 150$, taking into account that $|a| = 4$. The integration over the field A_{abc} substitutes the connection A_{abc} in the action with:

$$A_{abc} = \frac{1}{2} c_{abc} - 2\pi i l_p^2 \bar{\psi}_A \gamma_5 \gamma^d \psi^A \varepsilon_{abcd} \,, \tag{78}$$

which corresponds to the substitution

$$\omega_{ab\mu} = \Delta_{ab\mu} - 2\pi i l_p^2 \bar{\psi}_A \gamma_5 \gamma^d \psi^A \varepsilon_{abcd} e^c{}_\mu \,, \tag{79}$$

which is in turn the previously described procedure of obtaining the action for ECC theory from the action of the EC theory (see equation (46)). We thus end up with the expression featuring the action of the ECC theory:

$$\begin{aligned} Z_{3BF} \langle F \rangle_{3BF} &= \mathcal{N} \int De_\partial D\phi_\partial D\psi_\partial D\bar{\psi}_\partial \, \delta(\phi_\partial) \, \delta(\psi_\partial) \, \delta(\bar{\psi}_\partial) \, \delta(e_\partial) \\ &\quad \times \int D\alpha De D\phi D\psi D\bar{\psi} \, \frac{1}{|e|^M} \, F(\phi_k) \, \mathrm{e}^{\mathrm{i} S_{\mathrm{ECC}}[\phi_k]} \\ &= \mathcal{N}' Z_{\mathrm{ECC}} \left\langle \frac{1}{|e|^M} F \right\rangle_{\mathrm{ECC}} . \end{aligned} \tag{80}$$

As a special case, substituting the unit observable, $F(\phi_k) = 1$, one can evaluate the state sum as

$$\begin{aligned} Z_{3BF} &= \mathcal{N} \int De_\partial D\phi_\partial D\psi_\partial D\bar{\psi}_\partial \, \delta(\phi_\partial) \, \delta(\psi_\partial) \, \delta(\bar{\psi}_\partial) \, \delta(e_\partial) \int D\alpha De D\phi D\psi D\bar{\psi} \, \frac{1}{|e|^M} \, \mathrm{e}^{\mathrm{i} S_{\mathrm{ECC}}[\phi_k]} \\ &= \mathcal{N}' Z_{\mathrm{ECC}} \left\langle \frac{1}{|e|^M} \right\rangle_{\mathrm{ECC}} . \end{aligned} \tag{81}$$

As another special case, substituting the observable $F(\phi_k) = |e|^M$, one can evaluate the state sum in a different way, as

$$\begin{aligned} Z_{3BF} \left\langle |e|^M \right\rangle_{3BF} &= \mathcal{N} \int De_\partial D\phi_\partial D\psi_\partial D\bar{\psi}_\partial \, \delta(\phi_\partial) \, \delta(\psi_\partial) \, \delta(\bar{\psi}_\partial) \, \delta(e_\partial) \int D\alpha De D\phi D\psi D\bar{\psi} \, \mathrm{e}^{\mathrm{i} S_{\mathrm{ECC}}[\phi_k]} \\ &= \mathcal{N}' Z_{\mathrm{ECC}} . \end{aligned} \tag{82}$$

Combining (81) and (82), we obtain the normalization relation

$$\langle |e|^M \rangle_{3BF} \left\langle \frac{1}{|e|^M} \right\rangle_{\mathrm{ECC}} = 1 \,. \tag{83}$$

As the last step, substituting (81) and (82) back into (80), and remembering the definitions (61), we finally obtain the correspondence between the quantum $3BF$ theory and the quantum ECC theory, in the following form:

$$\langle F \rangle_{3BF} = \frac{\left\langle \frac{1}{|e|^M} F \right\rangle_{\mathrm{ECC}}}{\left\langle \frac{1}{|e|^M} \right\rangle_{\mathrm{ECC}}} \,, \qquad \langle F \rangle_{\mathrm{ECC}} = \frac{\langle |e|^M F \rangle_{3BF}}{\langle |e|^M \rangle_{3BF}} \,. \tag{84}$$

The equations (84) are the nonperturbative correspondence between the expectation values of observables in the quantum $3BF$ and quantum ECC theories that we were looking for, and represent the main result of the paper. The existence of this correspondence emphasizes the importance of the 3*BF* theory, since its quantization automatically gives rise to a quantization of the ECC theory, which is a physically relevant model of quantum gravity with matter of the Standard Model. Namely, it can be argued that the explicit construction of path integral for the quantum 3*BF* theory is, all else being equal, easier to perform than the direct construction of the path integral for the quantum ECC theory itself.

3.3. Boundary conditions

One technical detail that should be discussed is the presence of the boundary terms in the state sum (80) and eventual dependence of the observable F on values of fields at the boundary of a spacetime manifold (if it features a boundary). As they stand, the boundary terms in (80) feature Dirac delta functions for the boundary values of certain fields, essentially claiming that those fields should vanish at the boundary. While this is not a big issue for fermion fields (since we typically assume them to be localized to some finite region in the manifold bulk anyway), the situation is rather different for the scalar field, and most importantly, the tetrad field. Namely, the scalar field describes the Higgs sector, which is known to have a nonzero vacuum expectation value, which should be constant throughout spacetime. This is in obvious conflict with the statement that it should be zero at the spacetime boundary. Similarly, the tetrad field is typically assumed to be nondegenerate (the tetrad determinant e is assumed to be nonzero everywhere), since otherwise there would be singularities in the metric structure of spacetime. This is again in obvious conflict with the statement that tetrad fields should become zero at the spacetime boundary.

There are two ways one can think of this issue. One way would be to postulate that spacetime does not (or should not) have a boundary to begin with. Then one could simply drop the boundary terms from (80), and the issue of the values of Higgs and tetrad fields at the boundary would become immaterial. Another way would be to modify the 3*BF* action by adding suitable boundary counterterms. These extra terms would not influence the dynamics of the theory in the bulk, but would influence the boundary values of fields. For example, this is a property of the well known Gibbons–Hawking–York (GHY) boundary term [58, 59]. It would thus be interesting to introduce such modifications to the 3*BF* action, and study their influence on the expectation values of observables in (80), with a possibility of fixing the issues related to the Higgs and tetrad fields.

In order to discuss GHY term, we first need some elementary formalism to describe the boundary of the spacetime manifold. Looking at some coordinate patch x^μ of $\mathcal{M}$ which intersects the boundary $\partial\mathcal{M}$, one can introduce a new set of coordinates ξ^i ($i = 1,2,3$) on the intersection patch in $\partial\mathcal{M}$. Given this, a point with coordinates ξ^i on $\partial\mathcal{M}$ can be assigned coordinates x^μ on $\mathcal{M}$ using the parametric equations

$$x^\mu = z^\mu\left(\xi^i\right), \tag{85}$$

where $z^\mu(\xi)$ are some functions encoding the 'position' of $\partial\mathcal{M}$ in $\mathcal{M}$. In the tangent space of point ξ^i one can introduce the natural coordinate basis $u_i \equiv \partial_i$. These vectors also live in the tangent space of the same point in $\mathcal{M}$, so one can expand them in the coordinate basis ∂_μ as $u_i = u_i^\mu \partial_\mu$, where the components u_i^μ can be evaluated using (85) as

$$u_i^\mu = \frac{\partial z^\mu(\xi)}{\partial \xi^i}. \tag{86}$$

Next, given a metric $g_{\mu\nu}(x)$ on $\mathcal{M}$, one introduces the induced metric $\gamma_{ij}(\xi)$ on $\partial\mathcal{M}$ as a pullback:

$$\gamma_{ij}(\xi) = g_{\mu\nu}(z(\xi))\, u_i^\mu u_j^\nu \,. \tag{87}$$

This metric (in older literature also called the *first fundamental form* on $\partial\mathcal{M}$) is assumed to be nondegenerate, with its inverse denoted as γ^{ij}, and can be used to raise and lower the boundary indices $i, j, \ldots$. One can additionally introduce the induced triads, connection, covariant derivative, curvature, torsion, and various other induced quantities on the boundary, but this is not necessary for the purpose of this work. In addition to all these standard geometric notions associated to $\partial\mathcal{M}$ as a manifold, the boundary has some additional properties that stem from its embedding into $\mathcal{M}$. Namely, given that $\mathcal{M}$ is 4-dimensional and $\partial\mathcal{M}$ is 3-dimensional, there will be one additional vector in the tangent space of $\mathcal{M}$ that is linearly independent of the three tangent vectors u_i. Calling it the *normal vector*, and denoting its components as n^μ, one can choose it to be orthogonal to all three tangent vectors, $n_\mu u_i^\mu = 0$, so that the following resolution of the identity holds:

$$\delta^\mu_\nu = \epsilon n^\mu n_\nu + u_i^\mu u^i_\nu \,, \qquad (\epsilon = \pm 1)\,. \tag{88}$$

The normal vector is normalized as $n^\mu n_\mu = \epsilon$, and the boundary $\partial\mathcal{M}$ is called spacelike if its normal vector is timelike ($\epsilon = -1$, recall that we work with the $(-,+,+,+)$ signature convention throughout the paper), while it is called timelike if its normal vector is spacelike ($\epsilon = +1$). We will not introduce lightlike boundary since it is not necessary for our purposes.

The normal vector allows us to introduce one more notion specific to the boundary, called *extrinsic curvature* (in older literature also called the *second fundamental form*), as a projection of the covariant derivative of the normal vector onto the tangent space of $\partial\mathcal{M}$,

$$K_{ij} = u_i^\mu u_j^\nu \nabla_\mu n_\nu \,, \tag{89}$$

where ∇_μ denotes the standard covariant derivative on $\mathcal{M}$, compatible with the metric $g_{\mu\nu}$. The extrinsic curvature scalar is defined as $K = \gamma^{ij} K_{ij}$.

Now we are ready to introduce the GHY boundary term. In the context of the traditional Einstein–Hilbert formulation of the action for GR, one can write:

$$S_{\text{GR}} = -\frac{1}{16\pi G}\int_{\mathcal{M}} d^4x\, \sqrt{-g}R - \frac{1}{8\pi G}\oint_{\partial\mathcal{M}} d^3\xi\, \epsilon\sqrt{|\gamma|}K. \tag{90}$$

The first term is the standard Einstein–Hilbert action for GR, while the second term is the GHY boundary term. Its purpose is to make sure that the action S_{GR} has well-defined functional derivatives with respect to the metric $g_{\mu\nu}$, since the curvature scalar R contains its second derivatives. Namely, in the variation of the action δS_{GR} the second derivatives of the metric in R give rise to the variation of the first derivatives of the metric on the boundary, which are then canceled by the GHY term, rendering δS_{GR} ultimately dependent solely on the variation $\delta g_{\mu\nu}$ of the metric itself, rather than its derivatives [58, 59].

The form of the GHY term in (90) has been designed precisely to correspond to the Einstein–Hilbert formulation of GR, and does not *a priori* fulfill its purpose when the gravitational interaction is formulated in terms of the EC action, or the 3*BF* action, or otherwise. Nevertheless, in [60] the GHY boundary term has been reformulated to match a completely arbitrary theory that may contain curvature, torsion, and even nonmetricity. In our work, nonmetricity is absent, but curvature and torsion are present, so the results of [60] lend themselves to be applied in a straightforward manner to the case of the $3BF$ action (17). Writing (17) in the form

$$S_{3BF} = \int_{\mathcal{M}} \mathcal{L}_{3BF} \,, \tag{91}$$

where $\mathcal{L}_{3BF}$ is the Lagrangian 4-form of the $3BF$ action, the GHY term is given as:

$$S^{3BF}_{\text{GHY}} = 2 \oint_{\partial\mathcal{M}} \epsilon K^i \wedge n^\mu u_i^\nu \star\varphi_{\mu\nu} \Big|_{\partial\mathcal{M}} . \tag{92}$$

Here K^i is the extrinsic curvature boundary 1-form

$$K^i \equiv K^i{}_j \mathrm{d}\xi^j = u^{i\mu} u_j^\nu \nabla_\mu n_\nu \mathrm{d}\xi^j , \tag{93}$$

while $\star\varphi_{\mu\nu}$ are the 2-forms obtained from the specific details of the Lagrangian 4-form $\mathcal{L}_{3BF}$, see [60]. For the purpose of our work, we are not interested in the precise form of $\star\varphi_{\mu\nu}$. We merely need to know that it depends on all fields present in the action (17), i.e. $\star\varphi_{\mu\nu}$ is a function over the whole kinematical configuration space (42) and (43).

At this point we can study the influence of the GHY boundary term on the derivation of our main result (84). We begin by modifying the original $3BF$ action by adding to it the GHY boundary term,

$$S_{3BF} \to S_{3BF} + S^{3BF}_{\text{GHY}} . \tag{94}$$

We then proceed through the calculation described in the previous subsection. Starting from (63) with the added GHY boundary term, we proceed to (64), where the boundary contribution to the path integral (the first row of (64)) now generalizes to:

$$\int D\beta_\partial De_\partial D\tilde{\gamma}_\partial D\gamma_\partial D\bar{\gamma}_\partial D\phi_\partial D\psi_\partial D\bar{\psi}_\partial DB_\partial D\zeta_\partial D\Lambda_\partial D\alpha_\partial D\omega_\partial D\tilde{\lambda}_\partial D\lambda_\partial D\bar{\lambda}_\partial DH_\partial DM_\partial \\ \times \mathrm{e}^{\mathrm{i}S^{3BF}_{\text{GHY}} + i\oint \phi^A_\partial \tilde{\gamma}_{A\partial} + \bar{\psi}_{A\partial}\gamma^A_\partial + \bar{\gamma}_{A\partial}\psi^A_\partial - e_{a\partial}\wedge\beta^a_\partial} . \tag{95}$$

The difference from the original boundary contribution in (64) consists of the fact that the additional S^{3BF}_{GHY} term is present, and since it depends on the full configuration space, we explicitly denote the path integrals over all boundary fields.

The subsequent steps, given by (65)–(68) all the way to (80), involve integrating out a range of variables living in the bulk, without any contributions to the boundary. This means that the boundary term (95) appears in (80) instead of the old one. Note that (80) features also the ECC action (50), which should arguably also be corrected by its own version of the GHY boundary term,

$$S_{\text{ECC}} \to S_{\text{ECC}} + S^{\text{ECC}}_{\text{GHY}} , \tag{96}$$

where

$$S^{\text{ECC}}_{\text{GHY}} = 2 \oint_{\partial\mathcal{M}} \epsilon K^i \wedge n^\mu u_i^\nu \star\tilde{\varphi}_{\mu\nu} \Big|_{\partial\mathcal{M}} . \tag{97}$$

Here $\star\tilde{\varphi}_{\mu\nu}$ are the 2-forms obtained from the specific details of the ECC Lagrangian 4-form $\mathcal{L}_{\text{ECC}}$. It is a function over the kinematical configuration space (51) of ECC theory. Subtracting the term $S^{\text{ECC}}_{\text{GHY}}$ from the boundary term (95), and reabsorbing it into a redefinition (96) of S_{ECC}, the final generalized form of equation (80) reads:

$$
\begin{aligned}
Z_{3BF}\langle F\rangle_{3BF} = \mathcal{N}\int & D\beta_\partial De_\partial D\tilde{\gamma}_\partial D\gamma_\partial D\bar{\gamma}_\partial D\phi_\partial D\psi_\partial D\bar{\psi}_\partial DB_\partial D\zeta_\partial D\Lambda_\partial D\alpha_\partial D\omega_\partial D\tilde{\lambda}_\partial D\lambda_\partial D\bar{\lambda}_\partial DH_\partial DM_\partial \\
& \times \mathrm{e}^{\mathrm{i}S^{3BF}_{\mathrm{GHY}} - iS^{\mathrm{ECC}}_{\mathrm{GHY}} + i\oint \phi^A_\partial \tilde{\gamma}_{A\partial} + \bar{\psi}_{A\partial}\gamma^A_\partial + \bar{\gamma}_{A\partial}\psi^A_\partial - e_{a\partial}\wedge\beta^a_\partial} \\
& \times \int D\alpha De D\phi D\psi D\bar{\psi}\, \frac{1}{|e|^M}\, F(\phi_k)\, \mathrm{e}^{\mathrm{i}S_{\mathrm{ECC}}[\phi_k]} \\
= \mathcal{N}' Z_{\mathrm{ECC}} & \left\langle \frac{1}{|e|^M} F \right\rangle_{\mathrm{ECC}} .
\end{aligned} \tag{98}
$$

The subsequent analysis leading to (84) remains unchanged—one evaluates (98) for observables $F(\phi_k) = 1$ and $F(\phi_k) = |e^M|$ and combines them to obtain both (83) and (84), in an unchanged form, completing the main result.

Looking at the boundary term in (98), we can observe that it is substantially different from the boundary term in (80). Namely, both GHY terms in the exponent depend on all variables in the configuration space, which means that the four remaining terms cannot be integrated in a straightforward manner, and one does not obtain the problematic Dirac delta functions $\delta(\phi_\partial)\delta(\psi_\partial)\delta(\bar{\psi}_\partial)\delta(e_\partial)$. The evaluation of the explicit form of S^{3BF}_{GHY} and $S^{\mathrm{ECC}}_{\mathrm{GHY}}$, and the detailed study of its contribution to the boundary integral, are out of the scope of the current work, and we postpone them for future research. Nevertheless, even without explicit calculation, it is reasonably obvious that the GHY terms will give nontrivial contribution and will either completely remove the Dirac delta functions, or substantially modify their arguments, allowing the tetrad, scalar and fermion fields to have potentially nonzero values on the boundary.

4. Examples

In this section we will compare the predictions of the quantum $3BF$ and ECC theories on the example of the spacetime volume density and discuss the classical limit. Then we will study the example of the gravitational waves.

4.1. Spacetime 4-volume density

As a simplest example, let us define the 4-volume density observable defined as

$$
F(\phi_k) = \rho \equiv |e|, \tag{99}
$$

and the value of the 4-volume density as the expectation value of this operator, in a given quantum theory of gravity. The name is motivated by the fact that the 4-volume of some 4-dimensional region $\mathcal{R}$ of spacetime is given as

$$
V(\mathcal{R}) = \int_{\mathcal{R}} \mathrm{d}^4x\sqrt{-g} = \int_{\mathcal{R}} \mathrm{d}^4x |e|, \tag{100}
$$

hence one can loosely call $|e|$ as the 'density' of spacetime 4-volume of the region $\mathcal{R}$.

The correspondence relations (84) can be applied to obtain the ratio of the expectation values of 4-volume density in the two quantum theories:

$$
\frac{\rho_{3BF}}{\rho_{\mathrm{ECC}}} \equiv \frac{\langle\rho\rangle_{3BF}}{\langle\rho\rangle_{\mathrm{ECC}}} = \frac{\left\langle \frac{1}{|e|^{M-1}} \right\rangle_{\mathrm{ECC}}}{\langle|e|\rangle_{\mathrm{ECC}} \left\langle \frac{1}{|e|^M} \right\rangle_{\mathrm{ECC}}} = \frac{\langle|e|\rangle_{3BF}\left\langle|e|^M\right\rangle_{3BF}}{\left\langle|e|^{M+1}\right\rangle_{3BF}} . \tag{101}
$$

At this point it is useful to remember the definitions of the statistical quantities of covariance and variance, which are useful to separate the quantum corrections from the classical quantities:

$$\text{Cov}\,(X,Y) = \langle XY\rangle - \langle X\rangle\langle Y\rangle\,, \qquad \text{Var}\,(X) = \text{Cov}\,(X,X) = (\Delta X)^2\,. \tag{102}$$

Here ΔX represents the standard deviation, i.e. the uncertainty of the observable X. Covariance and variance satisfy the Cauchy–Schwarz inequality

$$|\text{Cov}\,(X,Y)| \leqslant \Delta X \Delta Y, \tag{103}$$

which can be used to estimate the covariance. Based on the definitions (102), it is clear that the ratio of the 4-volume densities in $3BF$ and ECC theories is given as:

$$\frac{\rho_{3BF}}{\rho_{\text{ECC}}} = 1 + \frac{\text{Cov}\left(|e|, \frac{1}{|e|^M}\right)_{\text{ECC}}}{\langle|e|\rangle_{\text{ECC}}\left\langle\frac{1}{|e|^M}\right\rangle_{\text{ECC}}}, \qquad \frac{\rho_{\text{ECC}}}{\rho_{3BF}} = 1 + \frac{\text{Cov}\left(|e|, |e|^M\right)_{3BF}}{\langle|e|\rangle_{3BF}\langle|e|^M\rangle_{3BF}}. \tag{104}$$

Then, using (103), we can write

$$\frac{\rho_{3BF}}{\rho_{\text{ECC}}} \leqslant 1 + \left(\frac{\Delta|e|}{\langle|e|\rangle}\,\frac{\Delta\frac{1}{|e|^M}}{\left\langle\frac{1}{|e|^M}\right\rangle}\right)_{\text{ECC}}, \qquad \frac{\rho_{\text{ECC}}}{\rho_{3BF}} \leqslant 1 + \left(\frac{\Delta|e|}{\langle|e|\rangle}\,\frac{\Delta|e|^M}{\langle|e|^M\rangle}\right)_{3BF}. \tag{105}$$

In the classical limit one can assume that the uncertainties tend to zero, and we see that the 4-volume densities have approximately the same value in both theories.

In fact, the above example indicates that the classical limits of the two theories are the same. In order to demonstrate this in full generality, one can repeat the above analysis for the case of an arbitrary observable $F(\phi_k)$. Starting from (84), we have

$$\frac{\langle F\rangle_{3BF}}{\langle F\rangle_{\text{ECC}}} = \frac{\left\langle\frac{1}{|e|^M}F\right\rangle_{\text{ECC}}}{\left\langle\frac{1}{|e|^M}\right\rangle_{\text{ECC}}\langle F\rangle_{\text{ECC}}}, \qquad \frac{\langle F\rangle_{\text{ECC}}}{\langle F\rangle_{3BF}} = \frac{\langle|e|^M F\rangle_{3BF}}{\langle|e|^M\rangle_{3BF}\langle F\rangle_{3BF}}. \tag{106}$$

The expectation value of the product in the numerator can be removed in favor of covariance using (102), and the covariance can be estimated using the Cauchy–Schwarz inequality (103), leading us to:

$$\frac{\langle F\rangle_{3BF}}{\langle F\rangle_{\text{ECC}}} \leqslant 1 + \left(\frac{\Delta F}{\langle F\rangle}\,\frac{\Delta\frac{1}{|e|^M}}{\left\langle\frac{1}{|e|^M}\right\rangle}\right)_{\text{ECC}}, \qquad \frac{\langle F\rangle_{\text{ECC}}}{\langle F\rangle_{3BF}} \leqslant 1 + \left(\frac{\Delta F}{\langle F\rangle}\,\frac{\Delta|e|^M}{\langle|e|^M\rangle}\right)_{3BF}. \tag{107}$$

In the classical limit we expect the uncertainties of the observables to become negligible, $\Delta F \to 0$, giving us

$$\frac{\langle F\rangle_{3BF}}{\langle F\rangle_{\text{ECC}}} \leqslant 1, \qquad \frac{\langle F\rangle_{\text{ECC}}}{\langle F\rangle_{3BF}} \leqslant 1. \tag{108}$$

Taken together, these two inequalities enforce the result

$$\langle F\rangle_{3BF} = \langle F\rangle_{\text{ECC}}\,, \tag{109}$$

claiming that in the classical limit all observables have the same value in both theories, as expected. The two theories differ only at the level of quantum correction terms.

4.2. Gravitational waves

In the example of gravitational waves, we study the two (mutually related) fields, namely the tetrad field and the metric field, as well as their excitations over the flat spacetime configuration:

$$e^a{}_\mu = \delta^a_\mu + \varepsilon^a{}_\mu, \qquad g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}. \tag{110}$$

We expand the determinant to the second order (which is incidentally the full expansion to all orders):

$$|e| = 1 + \varepsilon^a{}_a + \frac{1}{2}\left(\varepsilon^a{}_a\varepsilon^b{}_b - \varepsilon^a{}_b\varepsilon^b{}_a\right) + \frac{1}{6}\left(\varepsilon^a{}_a\varepsilon^b{}_b\varepsilon^c{}_c - 3\varepsilon^a{}_a\varepsilon^b{}_c\varepsilon^c{}_b + 2\varepsilon^a{}_b\varepsilon^b{}_c\varepsilon^c{}_a\right) + |\varepsilon|. \tag{111}$$

Let us note here that in principle one can discuss gravitational wave perturbations over some more general curved background spacetime, rather than flat spacetime. In other words, given some background tetrad $\hat{e}^a{}_\mu$ and its corresponding background metric $\hat{g}_{\mu\nu}$, instead of (110) we could write

$$e^a{}_\mu = \hat{e}^a{}_\mu + \varepsilon^a{}_\mu, \qquad g_{\mu\nu} = \hat{g}_{\mu\nu} + h_{\mu\nu}. \tag{112}$$

While the subsequent analysis is in principle similar, it is technically more complicated since the evaluation of the determinant (111) would not have 1 as the leading term, but rather $\det \hat{e}^a{}_\mu$. Because of this, we opt not to work with a generic curved background, but rather with the special case of the flat background. Qualitatively speaking, all results and conclusions of the analysis remain the same.

One can introduce the quantity E that collects all corrections of the tetrad determinant with respect to unity. This is useful for the study of the convergence properties of the power series of $\pm M$th degree of the tetrad determinant over such a parameter, so the corresponding exponents of the tetrad determinant in second order are given as:

$$\frac{1}{|e|^M} = \frac{1}{(1+E)^M} = \sum_{n=0}^{+\infty}\binom{M+n}{n}(-E)^n = 1 - M\varepsilon^a{}_a + \frac{M}{2}\left(M\varepsilon^a{}_a\varepsilon^b{}_b + \varepsilon^a{}_b\varepsilon^b{}_a\right) + o\left(\varepsilon^2\right). \tag{113}$$

The requirement that this series converges is that $|E| < 1$, while the terms in the series begin to decrease when the following condition is satisfied:

$$|E| < \frac{n+1}{M+n+1}. \tag{114}$$

From here it follows that the contribution of the second order will be greater that the contribution of the third order in case when $|E| < 0.0196$ (assuming that $n = 2$, $M = 150$), or in other words, if one requires that the second order contributes k times more than the third order, then $|E| < 0.0196/k$. Also, in the case of the tetrad determinant with a positive exponent, we similarly have:

$$|e|^M = (1+E)^M = \sum_{n=0}^{+\infty}\binom{M}{n}E^n = 1 + M\varepsilon^a{}_a + \frac{M}{2}\left(M\varepsilon^a{}_a\varepsilon^b{}_b - \varepsilon^a{}_b\varepsilon^b{}_a\right) + o\left(\varepsilon^2\right). \tag{115}$$

This series is finite, since the binomial coefficient is equal to zero when $n > M$, so there are no convergence issues. In addition, the terms in the series decrease when:

$$|E| < \frac{n+1}{M-n}, \tag{116}$$

which is a weaker requirement from the previous inequality (114) and therefore always satisfied.

By applying the correspondence equations (84) to the metric perturbation observable, $F(\phi_k) = h_{\mu\nu}$, we obtain the following:

$$\langle h_{\mu\nu}\rangle_{3BF} = \frac{\langle h_{\mu\nu}\rangle_{\rm ECC} - M\langle \varepsilon^a{}_a h_{\mu\nu}\rangle_{\rm ECC}}{1 - M\langle \varepsilon^a{}_a\rangle_{\rm ECC} + \frac{M}{2}\langle M\varepsilon^a{}_a\varepsilon^b{}_b + \varepsilon^a{}_b\varepsilon^b{}_a\rangle_{\rm ECC}}. \tag{117}$$

After the expansion of the denominator into power series one obtains:

$$\begin{aligned}\langle h_{\mu\nu}\rangle_{3BF} &= \left(\langle h_{\mu\nu}\rangle_{\rm ECC} - M\langle \varepsilon^a{}_a h_{\mu\nu}\rangle_{\rm ECC}\right)\\ &\quad\times\left(1 + M\langle \varepsilon^a{}_a\rangle_{\rm ECC} - \frac{M}{2}\langle M\varepsilon^a{}_a\varepsilon^b{}_b + \varepsilon^a{}_b\varepsilon^b{}_a\rangle_{\rm ECC} + \frac{M^2}{2}\langle \varepsilon^a{}_a\rangle^2_{\rm ECC}\right)\\ &= \langle h_{\mu\nu}\rangle_{\rm ECC}\left(1 + M\langle \varepsilon^a{}_a\rangle_{\rm ECC}\right) - M\langle \varepsilon^a{}_a h_{\mu\nu}\rangle_{\rm ECC}\end{aligned} \tag{118}$$

Also, in the opposite case we have

$$\begin{aligned}\langle h_{\mu\nu}\rangle_{\rm ECC} &= \left(\langle h_{\mu\nu}\rangle_{3BF} + M\langle \varepsilon^a{}_a h_{\mu\nu}\rangle_{3BF}\right)\\ &\quad\times\left(1 - M\langle \varepsilon^a{}_a\rangle_{3BF} - \frac{M}{2}\langle M\varepsilon^a{}_a\varepsilon^b{}_b - \varepsilon^a{}_b\varepsilon^b{}_a\rangle_{3BF} - \frac{M^2}{2}\langle \varepsilon^a{}_a\rangle^2_{3BF}\right)\\ &= \langle h_{\mu\nu}\rangle_{3BF}\left(1 - M\langle \varepsilon^a{}_a\rangle_{3BF}\right) + M\langle \varepsilon^a{}_a h_{\mu\nu}\rangle_{3BF}.\end{aligned} \tag{119}$$

These expressions simplify to the following final form:

$$\langle h_{\mu\nu}\rangle_{3BF} = \langle h_{\mu\nu}\rangle_{\rm ECC} - M\mathrm{Cov}\left(\varepsilon^a{}_a, h_{\mu\nu}\right)_{\rm ECC}, \tag{120}$$

$$\langle h_{\mu\nu}\rangle_{\rm ECC} = \langle h_{\mu\nu}\rangle_{3BF} + M\mathrm{Cov}\left(\varepsilon^a{}_a, h_{\mu\nu}\right)_{3BF}. \tag{121}$$

Using the Cauchy–Schwarz inequality (103), we can estimate the order of magnitude of the perturbation necessary for the experimental comparison between quantum $3BF$ and quantum ECC theories. Namely, up to second order, the deviation of the predictions between two theories is given as

$$\mathrm{Cov}\left(\varepsilon^a{}_a, h_{\mu\nu}\right)_{\rm ECC} = \mathrm{Cov}\left(\varepsilon^a{}_a, h_{\mu\nu}\right)_{3BF} = \mathrm{Cov}\left(\varepsilon^a{}_a, h_{\mu\nu}\right), \tag{122}$$

i.e. the correction term is the same in both theories up to second order, which can be seen from (120) and (121). Besides, based on the relations $h_{\mu\nu} = \eta_{\mu a}\varepsilon^a{}_\nu + \eta_{a\nu}\varepsilon^a{}_\mu$, $\varepsilon_{\mu\nu} = \eta_{\mu a}\varepsilon^a{}_\nu$ and the Cauchy–Schwarz inequality, assuming also that the uncertainty of each component of the tetrad is approximately the same, one obtains that:

$$\begin{aligned}\langle h_{\mu\nu}\rangle_{\rm ECC} - \langle h_{\mu\nu}\rangle_{3BF} &= M\mathrm{Cov}\left(\varepsilon^a{}_a, h_{\mu\nu}\right) \leqslant 2M\Delta\varepsilon^a{}_a\Delta\varepsilon_{\mu\nu}\\ &\approx 8M\left(\Delta\varepsilon_{\mu\nu}\right)^2 \approx 2M\left(\Delta h_{\mu\nu}\right)^2.\end{aligned} \tag{123}$$

The inequality (123) can in principle be used to experimentally distinguish between the quantum $3BF$ theory and quantum ECC theory, by measuring the gravitational waves and comparing the outcome to theoretical predictions. In order to obtain some intuition of the orders of magnitude involved, let us start from some ballpark orders of magnitude for current technological state of the art measurements of gravitational waves, taking for example LIGO/Virgo detectors as reference. According to [61], a typical precision of the strain measurement can be estimated to be 10^{-21}, which means that the right-hand side of (123) should be

$$2M(\Delta h_{\mu\nu})^2 \geqslant 10^{-21}. \tag{124}$$

Remembering that for the Standard Model we have $M = 150$ (see discussion below equation (77)), this gives us an estimate for the minimal quantum correction that can be detectable:

$$\Delta h_{\mu\nu} \geqslant \sqrt{\frac{10^{-21}}{2\cdot 150}} \approx 10^{-12}. \tag{125}$$

This is a huge value, as can be seen from the fact that the strain amplitude of the black hole merger signal in [61] is of order 10^{-18}. While the distance of GW150914 source was estimated to $r_{\text{GW}} \approx 410\text{Mpc}$, which is far outside of our galaxy, one can infer the strain amplitude of a hypothetical similar event happening within the Milky Way galaxy, i.e. at a distance of $r_{\text{MW}} \approx 34\text{Kpc}$. Since the amplitude of the strain of a spherical wave falls off as $1/r$ from the source, one could simply estimate that a similar black hole merger within our galaxy would give rise to the signal with strain amplitude of the order

$$h_{\text{MW}} \approx h_{\text{GW}}\frac{r_{\text{GW}}}{r_{\text{MW}}} = 10^{-18} \times \frac{4.1\cdot 10^5\text{Kpc}}{3.4\cdot 10^1\text{Kpc}} \approx 10^{-14}. \tag{126}$$

This is still two orders of magnitude smaller than the needed scale of the quantum correction $\Delta h_{\mu\nu}$. Moreover, there is nothing in the theory to suggest why a system of two merging black holes would even have a quantum uncertainty that big, to begin with.

In other words, using current technology, one would need a gravitational wave source that

(a) generates strain $\langle h_{\mu\nu}\rangle$ of the order of at least 10^{-11}, and
(b) gives rise to quantum uncertainty of the strain, $\Delta h_{\mu\nu}$, of the order of at least 10^{-12}.

Obviously, there are no known candidates for such a source of gravitational waves in nature. Nevertheless, at least in principle, if one were to have such a source, it would be possible to apply (123) to experimentally distinguish between the quantum $3BF$ theory and the quantum ECC theory.

5. Conclusions

5.1. Summary of the results

Let us summarize the results of the paper. After the Introduction, in section 2 we gave a short review of four classical actions—the topological $3BF$ action, the physically relevant constrained $3BF$ action, the EC action featuring the Standard Model matter sector, and the EC contact action, featuring the four-fermion contact interaction. We have demonstrated that the constrained $3BF$ theory and the ECC theory give rise to equivalent sets of classical EoM. In section 3, we have turned to the main analysis of the expectation values of an arbitrary quantum observable $F(\phi_k)$ that can be defined in both theories. After introducing some mathematical identities needed for the analysis, we have established a correspondence between the expectation value of the observable F in one theory, and the expectation value of a corresponding observable $|e|^{\pm M}F$ in the other theory, where e is the determinant of the tetrad fields, while the coefficient M has been determined to be $M = 150$. This correspondence has been established in a fully nonperturbative way, and represents the main result of the paper. Section 3 closes with an analysis of the boundary terms present in the theory. In section 4 we discussed some illustrative example observables, in order to compare the two quantum theories. First, we have discussed the spacetime 4-volume density as a simple example, and also the classical limit of the two theories. Then, we turned our attention to the example of gravitational waves, and we gave an estimate of how large their quantum uncertainties must be in order to be able to experimentally distinguish between the two quantum theories.

5.2. Discussion

The main relevance of our results is reflected in the following. On one hand, the classical ECC theory is arguably the phenomenologically very relevant model for the construction of a realistic full theory of quantum gravity with matter. Needless to say, the quantization of this action is rather hard, and so far remains an open problem in modern theoretical physics. On the other hand, the classical constrained $3BF$ theory represents a model that is slightly more tangible for efficient and rigorous quantization, at least within the path integral formalism. This is partly because it is based on a somewhat novel algebraic structure, a 3-group, which represents the generalization of the notion of a Lie group within the framework of higher gauge theory. Initial steps towards the path integral quantization of the $3BF$ theory for the case of the Standard Model 3-group have already been taken [31, 47–53], with promising results. Given this context, establishing a fully nonperturbative correspondence between the quantum $3BF$ and quantum ECC theory represents a quite useful result, since it allows us to sidestep the hard question of quantization of ECC theory itself, and instead define it in terms of the quantization of the $3BF$ theory.

The second important consequence of the correspondence is that there exists a regime where the $3BF$ and ECC theories could be experimentally distinguished from each other, at least in principle. As we have seen in section 4.2, given a source of gravitational waves that is both strong in magnitude and has large quantum uncertainty, the difference in the quantum corrections for the strain amplitude can in principle be large enough to be detected using current technology. This would allow us to compare $3BF$ and ECC theories against experimental data. In this sense, the correspondence predicts observable signatures that distinguish the two theories. Obviously, we do not have actual access to a gravitational wave source with the required properties, so any such experimental proposal is outside of the realm of practical feasibility. But as a matter of principle, the fact that such scenarios can be studied at least theoretically illustrates the phenomenological significance of the obtained correspondence.

Going beyond the $3BF$ and ECC theories, one can also ask is it possible to establish a similar correspondence in the context of other approaches to quantum gravity, such as the canonical LQG, or causal set theory (CST)? This is an interesting question, with no obvious answer. Namely, the correspondence between $3BF$ and ECC theories has been established using the language of path integrals. In the canonical LQG approach one starts essentially from a variant of EC theory, but the path integral language is not used. Instead, the quantization is being performed by foliating the spacetime into space and time, and by imposing canonical commutation relations on appropriate variables on each spatial hypersurface [3, 4].

This canonical quantization programme has not been developed for the 3*BF* theory so far. Nevertheless, assuming it could be developed, one could in principle study the same correspondence within the canonical language—given an observable in the canonical quantization of 3*BF* theory, can one find a corresponding observable in canonical LQG, such that their expectation values are equal? This may be an interesting topic for future research.

Regarding the possible correspondence between 3*BF* theory and CST, the situation is more complicated. While CST can be formulated using the (discrete) path integral language, there are two main issues that arise when trying to establish the correspondence between observables in 3*BF* and CST. First, the CST formalism has so far not been developed enough to describe non-scalar matter fields [62]. This is problematic, since the 3*BF* theory features gravity coupled to the full Standard Model, including fermions and gauge bosons. Therefore, a correspondence between most of the observables in 3*BF* and CST theories is impossible to establish. Second, in the classical limit, the 3*BF* theory gives rise to standard Einstein field equations of GR. On the other hand, the classical limit of CST cannot reproduce full Einstein field equations, since the fundamental CST assumption of a causal order relation excludes some of their solutions. For example, closed timelike curves are well known solutions of Einstein equations, but cannot be present within the CST formalism because they describe spacetime geometry with causal structure containing cycles, which is incompatible with a poset-type order relation required by CST. In this sense, 3*BF* and CST theories have different classical limits (more precisely, their sets of classical solutions are inequivalent), suggesting that the general correspondence between observables of the two theories is unlikely to exist.

When looking at the procedure of deriving the correspondence relations (84) in section 3, there are some technical details that merit further attention. Namely, the relation (68) comes tantalizingly close to establishing a correspondence between the 3*BF* theory and the standard EC theory, as opposed to the ECC theory. As was discussed below (68), such a prospect is thwarted due to the presence of the Dirac delta function encoding the algebraic relationship between the spin connection and fermion spin current, which is a consequence of the typical coupling between fermions and torsion in EC theory. The constrained 3*BF* action (17) has been constructed to encode precisely that same coupling, so the presence of the corresponding Dirac delta function should come as no surprise. Nevertheless, it would be an interesting avenue of research to extend both the 3*BF* theory and the EC theory to include some different type of coupling between fermions and torsion, which could potentially lead to a proper dynamical equation of motion for the spin connection. This is a valid possibility, since phenomenologically nobody in fact knows precisely how fermions couple to torsion—this interaction (if it exists to begin with) is too small and has so far eluded any experimental detection. In this context, one could formulate some alternative 3*BF* and EC models, which feature a different type of fermion-torsion coupling. Then one could attempt to repeat the analysis presented in this paper for those models, possibly establishing the correspondence between the quantum versions of the modified 3*BF* theory and the modified EC theory, rather than the ECC theory. This seems to be an interesting topic for future investigation.

Data availability statement

No new data were created or analysed in this study.

Acknowledgments

The authors would like to thank Tijana Radenković and Mihailo Đorđević for useful discussions and suggestions. This research was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia (MPNTR). MV was additionally supported by the Science Fund of the Republic of Serbia, Grant Number 7745968, project 'Quantum Gravity from Higher Gauge Theory 2021' – QGHG-2021. The contents of this publication are the sole responsibility of the authors and can in no way be taken to reflect the views of the Science Fund of the Republic of Serbia.

Appendix A. Proofs of the identities featuring Dirac delta functions

The proofs of the identities featuring Dirac delta functions for real and Grassmann fields can be established by the following straightforward computations:

- the identity (57):

$$
\begin{aligned}
&\int D\varphi D\phi_k \delta\left(\varphi^{aB} F_B{}^A(\phi_k) - G^{aA}(\phi_k)\right) H(\varphi,\phi_k) \\
&= \int D\left(\varphi^{aB} F_B{}^A(\phi_k)\right) D\phi_k \begin{vmatrix} \frac{\delta\left(\varphi F(\phi_k) F^{-1}(\phi_k)\right)^{aA}}{\delta\phi_n} & \frac{\delta\phi_m}{\delta\phi_n} \\ \frac{\delta\left(\varphi F(\phi_k) F^{-1}(\phi_k)\right)^{aA}}{\delta(\varphi F(\phi_k))^{bB}} & 0 \end{vmatrix} \delta\left(\varphi^{aB} F_B{}^A(\phi_k) - G^{aA}(\phi_k)\right) H(\varphi,\phi_k) \\
&= \int D\hat{\varphi} D\phi_k \left|\delta^a_b F^{-1}{}_B{}^A(\phi_k)\right| \delta\left(\hat{\varphi}^{aA} - G^{aA}(\phi_k)\right) H\left(\hat{\varphi}^{aB} F^{-1}{}_B{}^A(\phi_k), \phi_k\right) \\
&= \int D\phi_k \frac{1}{|F(\phi_k)|^{|a|}} H\left(G^{aB}(\phi_k) F^{-1}{}_B{}^A(\phi_k), \phi_k\right), \qquad \text{(A1)}
\end{aligned}
$$

- the identity (59):

$$
\begin{aligned}
&\int_{\mathbb{G}^n} \mathrm{d}\theta_1 \mathrm{d}\theta_2 \mathrm{d}\theta_3 \ldots \mathrm{d}\theta_n \mathrm{e}^{\mathrm{i}\theta_1(\theta_2-\theta_3-\ldots-\theta_k)} F(\theta_2,\theta_3,\ldots,\theta_n) \\
&= \int_{\mathbb{G}^n} \mathrm{d}\theta_1 \mathrm{d}\theta_2 \mathrm{d}\theta_3 \ldots \mathrm{d}\theta_n \left(1 + i\theta_1(\theta_2 - \theta_3 - \ldots - \theta_k)\right) \\
&\quad \times \left(\theta_3 \ldots \theta_k f_{01\ldots1}(\theta_{k+1},\ldots,\theta_n) + \ldots + \theta_2 \ldots \theta_{k-1} f_{1\ldots10}(\theta_{k+1},\ldots,\theta_n)\right) \\
&= i \int_{\mathbb{G}^n} \mathrm{d}\theta_1 \mathrm{d}\theta_2 \mathrm{d}\theta_3 \ldots \mathrm{d}\theta_n \theta_1 \theta_2 \theta_3 \ldots \theta_k \left(f_{01\ldots1}(\theta_{k+1},\ldots,\theta_n) + \ldots + (-1)^l f_{11\ldots10_l1\ldots1}(\theta_{k+1},\ldots,\theta_n)\right) \\
&= -i \int_{\mathbb{G}^n} \mathrm{d}\theta_3 \ldots \mathrm{d}\theta_n \mathrm{d}\theta_2 d\theta_1 \theta_1 \theta_2 \theta_3 \ldots \theta_k \left(f_{01\ldots1}(\theta_{k+1},\ldots,\theta_n) + \ldots + (-1)^l f_{11\ldots10_l1\ldots1}(\theta_{k+1},\ldots,\theta_n)\right) \\
&= -i \int_{\mathbb{G}^{n-2}} \mathrm{d}\theta_3 \ldots \mathrm{d}\theta_n \theta_3 \ldots \theta_k \left(f_{01\ldots1}(\theta_{k+1},\ldots,\theta_n) + \ldots + (-1)^l f_{11\ldots10_l1\ldots1}(\theta_{k+1},\ldots,\theta_n)\right) \\
&= -i \int_{\mathbb{G}^{n-2}} \mathrm{d}\theta_3 \ldots \mathrm{d}\theta_n F(\theta_3 + \ldots + \theta_k, \theta_3, \ldots, \theta_n) \\
&= -i \int_{\mathbb{G}^{n-1}} [d]\theta_3 \ldots \mathrm{d}\theta_n \mathrm{d}\theta_2 \,\delta(\theta_2 - \theta_3 - \ldots - \theta_k) F(\theta_2,\theta_3,\ldots,\theta_n) \\
&= (-1)^{n-1} i \int_{\mathbb{G}^{n-1}} \mathrm{d}\theta_2 \ldots \mathrm{d}\theta_n \,\delta(\theta_2 - \theta_3 - \ldots - \theta_k) F(\theta_2,\theta_3,\ldots,\theta_n), \qquad \text{(A2)}
\end{aligned}
$$

- the identity (60):

$$
\begin{aligned}
&\int_{\mathbb{R}^m} \mathrm{d}^m y \int_{\mathbb{R}^m} \mathrm{d}^m x \int_{\mathbb{G}^k} \mathrm{d}\theta_1 \mathrm{d}\theta_2 \dots \mathrm{d}\theta_k \int_{\mathbb{G}^{n-k}} \mathrm{d}\theta_{k+1} \dots \mathrm{d}\theta_n \, \mathrm{e}^{\mathrm{i}y_a \left(x^a - M^{aij}\theta_i\theta_j\right)} F(x, \theta_1, \dots, \theta_n) \\
&= \int_{\mathbb{R}^m} \mathrm{d}^m y \int_{\mathbb{R}^m} \mathrm{d}^m x \int_{\mathbb{G}^k} \mathrm{d}\theta_1 \mathrm{d}\theta_2 \dots \mathrm{d}\theta_k \int_{\mathbb{G}^{n-k}} \mathrm{d}\theta_{k+1} \dots \mathrm{d}\theta_n \, \mathrm{e}^{\mathrm{i}y_a x^a} \sum_{b=0}^{+\infty} \frac{1}{b!} \left(-iy_a M^{aij}\theta_i\theta_j\right)^b F(x, \theta_1, \dots, \theta_n) \\
&= \int_{\mathbb{R}^m} \mathrm{d}^m y \int_{\mathbb{R}^m} \mathrm{d}^m x \int_{\mathbb{G}^k} \mathrm{d}\theta_1 \mathrm{d}\theta_2 \dots \mathrm{d}\theta_k \int_{\mathbb{G}^{n-k}} \mathrm{d}\theta_{k+1} \dots \mathrm{d}\theta_n \, \mathrm{e}^{\mathrm{i}y_a x^a} \sum_{b=0}^{+\infty} \frac{1}{b!} \left(M^{aij}\theta_i\theta_j \frac{\partial}{\partial x^a}\right)^b F(x, \theta_1, \dots, \theta_n) \\
&= (2\pi)^m \int_{\mathbb{R}^m} \mathrm{d}^m x \int_{\mathbb{G}^k} \mathrm{d}\theta_1 \mathrm{d}\theta_2 \dots \mathrm{d}\theta_k \int_{\mathbb{G}^{n-k}} \mathrm{d}\theta_{k+1} \dots \mathrm{d}\theta_n \prod_{a=1}^{m} \delta(x^a) \sum_{b=0}^{+\infty} \frac{1}{b!} \left(M^{aij}\theta_i\theta_j \frac{\partial}{\partial x^a}\right)^b F(x, \theta_1, \dots, \theta_n) \\
&= (2\pi)^m \int_{\mathbb{G}^k} \mathrm{d}\theta_1 \mathrm{d}\theta_2 \dots \mathrm{d}\theta_k \int_{\mathbb{G}^{n-k}} \mathrm{d}\theta_{k+1} \dots \mathrm{d}\theta_n \sum_{b=0}^{+\infty} \frac{1}{b!} \left(M^{aij}\theta_i\theta_j \frac{\partial}{\partial x^a}\right)^b F(x, \theta_1, \dots, \theta_n)\Big|_{x=0} \\
&= (2\pi)^m \int_{\mathbb{G}^k} \mathrm{d}\theta_1 \mathrm{d}\theta_2 \dots \mathrm{d}\theta_k \int_{\mathbb{G}^{n-k}} \mathrm{d}\theta_{k+1} .. \mathrm{d}\theta_n \, F\left(M^{aij}\theta_i\theta_j, \theta_1, \dots, \theta_n\right) \\
&= (2\pi)^m \int_{\mathbb{R}^m} \mathrm{d}^m x \int_{\mathbb{G}^k} \mathrm{d}\theta_1 \mathrm{d}\theta_2 \dots \mathrm{d}\theta_k \int_{\mathbb{G}^{n-k}} \mathrm{d}\theta_{k+1} \dots \mathrm{d}\theta_n \prod_{a=1}^{m} \delta\left(x^a - M^{aij}\theta_i\theta_j\right) F(x, \theta_1, \dots, \theta_n) \,. \qquad \text{(A3)}
\end{aligned}
$$

Appendix B. Component form of the notation

Let us illustrate some details regarding the notation introduced in section 2.1. Given a Lie 2-crossed module (5),

$$\left(L \xrightarrow{\delta} H \xrightarrow{\partial} G \,,\, \triangleright \,,\, \{\_,\_\}_{\mathrm{pf}} \right) , \qquad \text{(B1)}$$

the Lie groups G, H and L have their corresponding Lie algebras $\mathfrak{g}$, $\mathfrak{h}$ and $\mathfrak{l}$, which allow us to introduce a corresponding linear structure, called a differential Lie 2-crossed module, as

$$\left(\mathfrak{l} \xrightarrow{\delta} \mathfrak{h} \xrightarrow{\partial} \mathfrak{g} \,,\, \triangleright \,,\, \{\_,\_\}_{\mathrm{pf}} \right) , \qquad \text{(B2)}$$

with the maps

$$\partial : \mathfrak{h} \to \mathfrak{g} \,, \qquad \delta : \mathfrak{l} \to \mathfrak{h} \,, \qquad \text{(B3)}$$

$$\triangleright : \mathfrak{g} \times \mathfrak{a} \to \mathfrak{a} \,, \qquad \mathfrak{a} = \mathfrak{g}, \mathfrak{h}, \mathfrak{l} \,, \qquad \text{(B4)}$$

$$\{\_,\_\}_{\mathrm{pf}} : \mathfrak{h} \times \mathfrak{h} \to \mathfrak{l} \,. \qquad \text{(B5)}$$

These maps are linearized versions of (2), (3) and (4), and are subject to axioms which are naturally induced by the axioms of a Lie 2-crossed module.

One can introduce sets of basis vectors for the three algebras $\mathfrak{g}$, $\mathfrak{h}$ and $\mathfrak{l}$, denoted respectively as

$$\{ \tau_\alpha \in \mathfrak{g} \,|\, \alpha = 1, \dots, \dim \mathfrak{g} \} \,, \quad \{ t_a \in \mathfrak{h} \,|\, a = 1, \dots, \dim \mathfrak{h} \} \,, \quad \{ T_A \in \mathfrak{l} \,|\, A = 1, \dots, \dim \mathfrak{l} \} \,. \qquad \text{(B6)}$$

This allows us to introduce the components of the above maps as:

$$\partial t_a = \partial_a{}^\alpha \tau_\alpha \,, \qquad \delta T_A = \delta_A{}^a t_a \,, \qquad \{ t_a , t_b \}_{\mathrm{pf}} = X_{ab}{}^A T_A \,, \qquad \text{(B7)}$$

$$\tau_\alpha \triangleright \tau_\beta = \triangleright_{\alpha\beta}{}^{\gamma}\tau_\gamma\,, \qquad \tau_\alpha \triangleright t_a = \triangleright_{\alpha a}{}^{b} t_b\,, \qquad \tau_\alpha \triangleright T_A = \triangleright_{\alpha A}{}^{B} T_B\,. \tag{B8}$$

Next, given a 4-dimensional manifold $\mathcal{M}$, one can denote the space of differential p-forms over $\mathcal{M}$ as $\Lambda_p(\mathcal{M})$, and its natural basis as $\mathrm{d}x^{\mu_1}\wedge\cdots\wedge\mathrm{d}x^{\mu_p}$. One can then introduce a principal 3-group bundle over a 4-dimensional manifold $\mathcal{M}$. As a section of this bundle one can introduce the 3-connection (α,β,γ), which consists of a $\mathfrak{g}$-valued 1-form α, an $\mathfrak{h}$-valued 2-form β and an $\mathfrak{l}$-valued 3-form γ:

$$\alpha \in \Lambda_1(\mathcal{M})\otimes\mathfrak{g}\,, \qquad \beta\in\Lambda_2(\mathcal{M})\otimes\mathfrak{h}\,, \qquad \gamma\in\Lambda_3(\mathcal{M})\otimes\mathfrak{l}\,. \tag{B9}$$

These can be expanded into components using appropriate basis vectors as:

$$\alpha = \alpha^{\alpha}{}_{\mu}(x)\ \mathrm{d}x^\mu\otimes\tau_\alpha\,, \quad \beta = \frac{1}{2}\beta^{a}{}_{\mu\nu}(x)\ \mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\otimes t_a\,,$$
$$\gamma = \frac{1}{3!}\gamma^{A}{}_{\mu\nu\lambda}(x)\ \mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\wedge\mathrm{d}x^\lambda\otimes T_A\,. \tag{B10}$$

The field $\alpha^{\alpha}{}_{\mu}(x)$ is precisely a traditional connection of a principal G-bundle, while $\beta^{a}{}_{\mu\nu}(x)$ and $\gamma^{A}{}_{\mu\nu\lambda}(x)$ are additional fields, native to the framework based on a principal 3-group bundle.

Given a 3-connection (α,β,γ), one can introduce a so-called fake 3-curvature $(\mathcal{F},\mathcal{G},\mathcal{H})$, consisting of a $\mathfrak{g}$-valued 2-form $\mathcal{F}$, an $\mathfrak{h}$-valued 3-form $\mathcal{G}$ and an $\mathfrak{l}$-valued 4-form $\mathcal{H}$:

$$\mathcal{F}\in\Lambda_2(\mathcal{M})\otimes\mathfrak{g}\,, \qquad \mathcal{G}\in\Lambda_3(\mathcal{M})\otimes\mathfrak{h}\,, \qquad \mathcal{H}\in\Lambda_4(\mathcal{M})\otimes\mathfrak{l}\,. \tag{B11}$$

These are defined as (see equation (7) in the main text):

$$\mathcal{F} = \mathrm{d}\alpha+\alpha\wedge\alpha-\partial\beta\,, \qquad \mathcal{G} = \mathrm{d}\beta+\alpha\wedge^{\triangleright}\beta-\delta\gamma\,, \qquad \mathcal{H} = \mathrm{d}\gamma+\alpha\wedge^{\triangleright}\gamma+\{\beta\wedge\beta\}_{\mathrm{pf}}\,. \tag{B12}$$

The notation $\wedge^{\triangleright}$ means that one should apply the wedge-product $\wedge$ in the subspace of differential forms, while simultaneously apply the action $\triangleright$ in the algebra subspace. One can use the above equations to work out the explicit components of the fake 3-curvature. For example, the components of the 3-form $\mathcal{G}$ are obtained as follows. The $\mathrm{d}\beta$ term is:

$$\begin{aligned}\mathrm{d}\beta &= \mathrm{d}\left(\frac{1}{2}\beta^{b}{}_{\mu\nu}\ \mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\otimes t_b\right)\\ &= \frac{1}{2}\partial_\lambda\beta^{b}{}_{\mu\nu}\ \mathrm{d}x^\lambda\wedge\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\otimes t_b\\ &= \frac{1}{6}\left[\partial_\lambda\beta^{b}{}_{\mu\nu}+\partial_\mu\beta^{b}{}_{\nu\lambda}+\partial_\nu\beta^{b}{}_{\lambda\mu}\right]\ \mathrm{d}x^\lambda\wedge\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\otimes t_b\,.\end{aligned} \tag{B13}$$

The $\alpha\wedge^{\triangleright}\beta$ term is:

$$\begin{aligned}\alpha\wedge^{\triangleright}\beta &= \left(\alpha^{\alpha}{}_{\lambda}\,\mathrm{d}x^\lambda\otimes\tau_\alpha\right)\wedge^{\triangleright}\left(\frac{1}{2}\beta^{a}{}_{\mu\nu}\ \mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\otimes t_a\right)\\ &= \frac{1}{2}\alpha^{\alpha}{}_{\lambda}\beta^{a}{}_{\mu\nu}\ \left(\mathrm{d}x^\lambda\wedge\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\right)\otimes\left(\tau_\alpha\triangleright t_a\right)\\ &= \frac{1}{2}\triangleright_{\alpha a}{}^{b}\alpha^{\alpha}{}_{\lambda}\beta^{a}{}_{\mu\nu}\ \mathrm{d}x^\lambda\wedge\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\otimes t_b\\ &= \frac{1}{6}\left[\triangleright_{\alpha a}{}^{b}\alpha^{\alpha}{}_{\lambda}\beta^{a}{}_{\mu\nu}+\triangleright_{\alpha a}{}^{b}\alpha^{\alpha}{}_{\mu}\beta^{a}{}_{\nu\lambda}+\triangleright_{\alpha a}{}^{b}\alpha^{\alpha}{}_{\nu}\beta^{a}{}_{\lambda\mu}\right]\ \mathrm{d}x^\lambda\wedge\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\otimes t_b.\end{aligned} \tag{B14}$$

The $\delta\gamma$ term is:

$$\begin{aligned}\delta\gamma &= \delta\left(\frac{1}{3!}\gamma^A{}_{\mu\nu\lambda}\,\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\wedge\mathrm{d}x^\lambda\otimes T_A\right)\\ &= \frac{1}{6}\gamma^A{}_{\mu\nu\lambda}\,\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\wedge\mathrm{d}x^\lambda\otimes\delta T_A\\ &= \frac{1}{6}\left[\delta_A{}^b\gamma^A{}_{\mu\nu\lambda}\right]\,\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\wedge\mathrm{d}x^\lambda\otimes t_b.\end{aligned} \tag{B15}$$

Putting all three terms together, and comparing to the expansion of $\mathcal{G}$ into components,

$$\mathcal{G} = \frac{1}{3!}\mathcal{G}^b{}_{\lambda\mu\nu}(x)\,\mathrm{d}x^\lambda\wedge\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\otimes t_b\,, \tag{B16}$$

one obtains:

$$\begin{aligned}\mathcal{G}^b{}_{\lambda\mu\nu} &= \partial_\lambda\beta^b{}_{\mu\nu}+\partial_\mu\beta^b{}_{\nu\lambda}+\partial_\nu\beta^b{}_{\lambda\mu}+\rhd_{\alpha a}{}^b\alpha^\alpha{}_\lambda\beta^a{}_{\mu\nu}+\rhd_{\alpha a}{}^b\alpha^\alpha{}_\mu\beta^a{}_{\nu\lambda}\\ &\quad+\rhd_{\alpha a}{}^b\alpha^\alpha{}_\nu\beta^a{}_{\lambda\mu}-\delta_A{}^b\gamma^A{}_{\mu\nu\lambda}\,.\end{aligned} \tag{B17}$$

In a similar fashion, one can derive the components for $\mathcal{F}$ and $\mathcal{H}$ as well. The result is

$$\mathcal{F}^\alpha{}_{\mu\nu} = \partial_\mu\alpha^\alpha{}_\nu-\partial_\mu\alpha^\alpha{}_\nu+\rhd_{\beta\gamma}{}^\alpha\alpha^\beta{}_\mu\alpha^\gamma{}_\nu-\partial_a{}^\alpha\beta^a{}_{\mu\nu}\,, \tag{B18}$$

and

$$\begin{aligned}\mathcal{H}^A{}_{\mu\nu\rho\sigma} &= \partial_\mu\gamma^A{}_{\nu\rho\sigma}-\partial_\nu\gamma^A{}_{\rho\sigma\mu}+\partial_\rho\gamma^A{}_{\sigma\mu\nu}-\partial_\sigma\gamma^A{}_{\mu\nu\rho}\\ &\quad+\rhd_{\alpha B}{}^A\alpha^\alpha{}_\mu\gamma^B{}_{\nu\rho\sigma}-\rhd_{\alpha B}{}^A\alpha^\alpha{}_\nu\gamma^B{}_{\rho\sigma\mu}+\rhd_{\alpha B}{}^A\alpha^\alpha{}_\rho\gamma^B{}_{\sigma\mu\nu}-\rhd_{\alpha B}{}^A\alpha^\alpha{}_\sigma\gamma^B{}_{\mu\nu\rho}\\ &\quad+2X_{ab}{}^A\beta^a{}_{\mu\nu}\beta^b{}_{\rho\sigma}-2X_{ab}{}^A\beta^a{}_{\mu\rho}\beta^b{}_{\nu\sigma}+2X_{ab}{}^A\beta^a{}_{\mu\sigma}\beta^b{}_{\nu\rho}\,.\end{aligned} \tag{B19}$$

Looking at the expressions for $\mathcal{G}$ and $\mathcal{H}$, one can note that it is always possible to combine the derivative term with the term containing a triangle into a covariant derivative term, as:

$$\nabla_\lambda\beta^b{}_{\mu\nu} = \partial_\lambda\beta^b{}_{\mu\nu}+\rhd_{\alpha a}{}^b\alpha^\alpha{}_\lambda\beta^a{}_{\mu\nu}\,,\quad \nabla_\mu\gamma^A{}_{\nu\rho\sigma} = \partial_\mu\gamma^A{}_{\nu\rho\sigma}+\rhd_{\alpha B}{}^A\alpha^\alpha{}_\mu\gamma^B{}_{\nu\rho\sigma}\,, \tag{B20}$$

where the triangle combined with the connection 1-form α serves the purpose of the connection term for the covariant derivative. This suggests to introduce a notion of covariant exterior derivative (see equation (8) in the main text) as

$$\nabla = \mathrm{d}+\alpha\wedge^\rhd \tag{B21}$$

which can act on any object in spaces $\Lambda_p(\mathcal{M})\otimes\mathfrak{g}$, $\Lambda_p(\mathcal{M})\otimes\mathfrak{h}$ and $\Lambda_p(\mathcal{M})\otimes\mathfrak{l}$. For example, given $\beta\in\Lambda_2(\mathcal{M})\otimes\mathfrak{h}$, we have:

$$\nabla\beta = \mathrm{d}\beta+\alpha\wedge^\rhd\beta = \frac{1}{2}\left[\underbrace{\partial_\lambda\beta^b{}_{\mu\nu}+\rhd_{\alpha a}{}^b\alpha^\alpha{}_\lambda\beta^a{}_{\mu\nu}}_{\nabla_\lambda\beta^b{}_{\mu\nu}}\right]\mathrm{d}x^\lambda\wedge\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\otimes t_b\,, \tag{B22}$$

using the results of (B13) and (B14). As another example, given $\phi \in \Lambda_0(\mathcal{M}) \otimes \mathfrak{l}$ (i.e. a set of scalar fields $\phi = \phi^A T_A$), we have:

$$\nabla\phi = \mathrm{d}\phi + \alpha \wedge^{\triangleright} \phi = \partial_\lambda \phi^A \, \mathrm{d}x^\lambda \otimes T_A + \alpha^\alpha{}_\lambda \phi^B \, \mathrm{d}x^\lambda \otimes \tau_\alpha \triangleright T_B$$
$$= \Bigg[\underbrace{\partial_\lambda \phi^A + \triangleright_{\alpha B}{}^A \alpha^\alpha{}_\lambda \phi^B}_{\nabla_\lambda \phi^A} \Bigg] \mathrm{d}x^\lambda \otimes T_A \,. \tag{B23}$$

The same calculation can be rewritten so that it does not expand 0-forms and 1-forms into a basis, as follows:

$$\nabla\phi = \mathrm{d}\phi + \alpha \wedge^{\triangleright} \phi = \mathrm{d}\phi^A \otimes T_A + \alpha^\alpha \wedge \phi^B \otimes \tau_\alpha \triangleright T_B = \Bigg[\underbrace{\mathrm{d}\phi^A + \triangleright_{\alpha B}{}^A \alpha^\alpha \wedge \phi^B}_{\nabla \phi^A} \Bigg] \otimes T_A \,. \tag{B24}$$

This illustrates equation (9) from the main text. As an exercise for an interested reader, one can apply (B21) to rewrite the field strenghts (B12) into a more compact form (see equation (10) in the main text):

$$\mathcal{F} = \nabla^2 - \partial\beta \,, \qquad \mathcal{G} = \nabla\beta - \delta\gamma \,, \quad \mathcal{H} = \nabla\gamma + \{\beta \wedge \beta\}_{\mathrm{pf}} \,. \tag{B25}$$

Finally, let us rewrite the topological $3BF$ action into the component form. The action is defined as (see equation (6) in the main text):

$$S_{3BF}^{\mathrm{top}} = \int_{\mathcal{M}_4} \langle B \wedge \mathcal{F} \rangle_{\mathfrak{g}} + \langle C \wedge \mathcal{G} \rangle_{\mathfrak{h}} + \langle D \wedge \mathcal{H} \rangle_{\mathfrak{l}} . \tag{B26}$$

The G-invariant nondegenerate symmetric bilinear forms $\langle \_, \_ \rangle_{\mathfrak{g}}$, $\langle \_, \_ \rangle_{\mathfrak{h}}$ and $\langle \_, \_ \rangle_{\mathfrak{l}}$ map a pair of algebra elements into a real number. Evaluating them on the basis vectors of the corresponding Lie algebras, one obtains their components:

$$\langle \tau_\alpha , \tau_\beta \rangle_{\mathfrak{g}} = g_{\alpha\beta} \,, \qquad \langle t_a , t_b \rangle_{\mathfrak{h}} = g_{ab} \,, \qquad \langle T_A , T_B \rangle_{\mathfrak{l}} = g_{AB} \,. \tag{B27}$$

Keeping in mind that the Lagrange multipliers B, C and D belong to appropriate spaces,

$$B \in \Lambda_2(\mathcal{M}) \otimes \mathfrak{g} \,, \quad C \in \Lambda_1(\mathcal{M}) \otimes \mathfrak{h} \,, \quad D \in \Lambda_0(\mathcal{M}) \otimes \mathfrak{l} \,, \tag{B28}$$

one can rewrite the action in terms of components as

$$S_{3BF}^{\mathrm{top}} = \int_{\mathcal{M}_4} \left(\frac{1}{4} g_{\alpha\beta} B^\alpha{}_{\mu\nu} \mathcal{F}^\beta{}_{\rho\sigma} + \frac{1}{3!} g_{ab} C^a{}_\mu \mathcal{G}^b{}_{\nu\rho\sigma} + \frac{1}{4!} g_{AB} D^A \mathcal{H}^B{}_{\mu\nu\rho\sigma} \right) \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \wedge \mathrm{d}x^\rho \wedge \mathrm{d}x^\sigma \,. \tag{B29}$$

Using the basic identity for differential forms

$$\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \wedge \mathrm{d}x^\rho \wedge \mathrm{d}x^\sigma = \varepsilon^{\mu\nu\rho\sigma} \, \mathrm{d}^4 x \,, \tag{B30}$$

one can finally rewrite the action into the traditional form

$$S_{3BF}^{\mathrm{top}} = \int_{\mathcal{M}_4} \mathcal{L}_{3BF}^{\mathrm{top}} \, \mathrm{d}^4 x \,, \tag{B31}$$

where the Lagrangian density for the topological 3BF theory is given in terms of component fields as:

$$\mathcal{L}^{\rm top}_{3BF} = \varepsilon^{\mu\nu\rho\sigma}\left(\frac{1}{4}g_{\alpha\beta}B^{\alpha}{}_{\mu\nu}\mathcal{F}^{\beta}{}_{\rho\sigma} + \frac{1}{3!}g_{ab}C^{a}{}_{\mu}\mathcal{G}^{b}{}_{\nu\rho\sigma} + \frac{1}{4!}g_{AB}D^{A}\mathcal{H}^{B}{}_{\mu\nu\rho\sigma}\right). \tag{B32}$$

ORCID iDs

Pavle Stipsić 0000-0001-5790-3997
Marko Vojinović 0000-0001-6977-4870

References

[1] Polchinski J 1998 *String Theory Vol. I: An Introduction to the Bosonic String* (Cambridge University Press)
[2] Polchinski J 1998 *String Theory Vol. II: Superstring Theory and Beyond* (Cambridge University Press)
[3] Rovelli C 2004 *Quantum Gravity* (Cambridge University Press)
[4] Thiemann T 2007 *Modern Canonical Quantum General Relativity* (Cambridge University Press)
[5] Rovelli C and Vidotto F 2014 *Covariant Loop Quantum Gravity* (Cambridge University Press)
[6] Miković A and Vojinović M 2023 *State-Sum Models of Piecewise Linear Quantum Gravity* (World Scientific)
[7] Baez J C and Huerta J 2011 *Gen. Relativ. Gravit.* **43** 2335
[8] Baez J and Dolan J 1995 *J. Math. Phys.* **36** 6073
[9] Baez J 2001 Higher-Dimensional Algebra and Planck-Scale Physics *Physics Meets Philosophy at the Planck Length* (Cambridge University Press) p 177
[10] Witten E 1988 *Commun. Math. Phys.* **117** 353
[11] Atiyah M F 1988 *Publications Mathématiques de l'IHÉS* **68** 175
[12] Quinn F 1995 *Lectures on axiomatic topological quantum field theory (Geometry and Quantum Field Theory)* (*IAS/Park City Mathematical Series*) vol 325, ed D S Freed and K K Uhlenbeck (American Mathematical Society, University of Texas)
[13] Sozer K and Virelizier A 2024 3D TQFTs and 3-manifold invariants *Encyclopedia of Mathematical Physics* 2nd edn (Elsevier)
[14] Baez J 1996 *Lett. Math. Phys.* **38** 129
[15] Yetter D N 1993 *Quantum Topology* (World Scientific)
[16] Porter T 1998 *J. Lond. Math. Soc.* **58** 723
[17] Miković A and Vojinović M 2012 *Class. Quantum Grav.* **29** 165003
[18] Li Z 2019 A global geometric approach to parallel transport of strings in gauge theory (arXiv:1910.14230)
[19] Sämann C and Wolf M 2014 *Commun. Math. Phys.* **328** 527
[20] Jurco B, Saemann C and Wolf M 2015 *J. High Energy Phys.* **04** JHEP04(2015)087
[21] Hidaka Y, Nitta M and Yokokura R 2020 *Phys. Lett. B* **808** 135672
[22] Hidaka Y, Nitta M and Yokokura R 2021 *J. High Energy Phys.* JHEP01(2021)173
[23] Song D, Lou K, Wu J, Yang J and Zhang F 2022 *J. Geom. Phys.* **178** 104537
[24] Song D, Lou K, Wu J and Yang J 2022 *Eur. Phys. J. C* **82** 1034
[25] Hidaka Y, Nitta M and Yokokura R 2021 *Phys. Lett. B* **823** 136762
[26] Hidaka Y, Nitta M and Yokokura R 2022 *Prog. Theor. Exper. Phys.* 04A109
[27] Jurco B, Sämann C and Wolf M 2016 *Fortsch. Phys.* **64** 674
[28] Sämann C and Wolf M 2017 *J. High Energy Phys.* JHEP07(2017)111
[29] Jurco B, Macrelli T, Raspollini L, Sämann C and Wolf M 2019 *Fortsch. Phys.* **67** 1910025
[30] Sämann C and Wolf M 2014 *Lett. Math. Phys.* **104** 1147
[31] Radenković T and Vojinović M 2019 *J. High Energy Phys.* JHEP10(2019)222
[32] Plebanski J F 1977 *J. Math. Phys.* **18** 2511
[33] Alexandrov M, Schwarz A, Zaboronsky O and Kontsevich M 1997 *Int. J. Mod. Phys. A* **12** 7

[34] Pulmann J, Ševera P and Valach F 2021 *Adv. Theor. Math. Phys.* **25** 1
[35] Cattaneo A S, Menger L and Schiavina M 2024 *Class. Quantum Grav.* **41** 155001
[36] Baez J C 2000 *Lect. Notes Phys.* **543** 25
[37] Celada M, González D and Montesinos M 2016 *Class. Quantum Grav.* **33** 213001
[38] Girelli F, Pfeiffer H and Popescu E M 2008 *J. Math. Phys.* **49** 032503
[39] Martins J F and Miković A 2011 *Adv. Theor. Math. Phys.* **15** 1059
[40] Miković A and Oliveira M A 2015 *Gen. Relativ. Gravit.* **47** 58
[41] Miković A, Oliveira M A and Vojinović M 2016 *Class. Quantum Grav.* **33** 065007
[42] Miković A, Oliveira M Â and Vojinović M 2022 *Adv. Theor. Math. Phys.* **26** 3783
[43] Miković A, Oliveira M A and Vojinović M 2019 *Class. Quantum Grav.* **36** 015005
[44] Asante S K, Dittrich B, Girelli F, Riello A and Tsimiklis P 2020 *Class. Quantum Grav.* **37** 205001
[45] Girelli F, Laudonio M and Tsimiklis P 2021 Polyhedron phase space using 2-groups: kappa-Poincare as a poisson 2-group (arXiv:2105.10616)
[46] Miković A and Vojinović M 2021 *Europhys. Lett.* **133** 61001
[47] Radenković T and Vojinović M 2020 *Ann. Univ. Craiova Phys.* **30** 74
[48] Radenković T and Vojinović M 2020 *Symmetry* **12** 620
[49] Radenković T and Vojinović M 2022 *Class. Quantum Grav.* **39** 135009
[50] Đorđević M, Radenković T, Stipsić P and Vojinović M 2023 *Universe* **9** 281
[51] Stipsić P and Vojinović M 2025 *J. Phys. A: Math. Theor.* **58** 215201
[52] Radenković T and Vojinović M 2022 *J. High Energy Phys.* JHEP07(2022)105
[53] Radenković T and Vojinović M 2025 *J. Phys. A: Math. Theor.* **58** 345203
[54] Martins J F and Picken R 2011 *Differ. Geom. Appl. J.* **29** 179–206
[55] Wang W 2014 *J. Math. Phys.* **55** 043506
[56] Blagojević M 2002 *Gravitation and Gauge Symmetries* (Institute of Physics Publishing)
[57] Vasilić M and Vojinović M 2008 *Phys. Rev. D* **78** 104002
[58] York J W 1972 *Phys. Rev. Lett.* **28** 1082
[59] Gibbons G W and Hawking S W 1977 *Phys. Rev. D* **15** 2752
[60] Erdmenger J, Heß B, Matthaiakakis I and Meyer R 2023 *SciPost Phys.* **14** 099
[61] Abbott B P *et al* 2016 *Phys. Rev. Lett.* **116** 061102
[62] Surya S 2019 *Living Rev. Relativ.* **22** 5

universe

MDPI

Article

Henneaux–Teitelboim Gauge Symmetry and Its Applications to Higher Gauge Theories

Mihailo Đorđević †, Tijana Radenković †, Pavle Stipsić † and Marko Vojinović *,†

Institute of Physics, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia; mdjordjevic@ipb.ac.rs (M.Đ.); rtijana@ipb.ac.rs (T.R.); pstipsic@ipb.ac.rs (P.S.)
* Correspondence: vmarko@ipb.ac.rs
† These authors contributed equally to this work.

Abstract: When discussing the gauge symmetries of any theory, the Henneaux–Teitelboim transformations are often underappreciated or even completely ignored, due to their on-shell triviality. Nevertheless, these gauge transformations play an important role in understanding the structure of the full gauge symmetry group of any theory, especially regarding the subgroup of diffeomorphisms. We give a review of the Henneaux–Teitelboim transformations and the resulting gauge group in the general case and then discuss its role in the applications to the class of topological theories called *nBF* models, relevant for the constructions of higher gauge theories and quantum gravity.

Keywords: gauge symmetry; trivial gauge transformations; *nBF* theory; Chern–Simons theory; diffeomorphism symmetry

check for updates

Citation: Đorđević, M.; Radenković, T.; Stipsić, P.; Vojinović, M. Henneaux–Teitelboim Gauge Symmetry and Its Applications to Higher Gauge Theories. *Universe* **2023**, *9*, 281. https://doi.org/10.3390/universe9060281

Received: 28 April 2023
Revised: 6 June 2023
Accepted: 7 June 2023
Published: 9 June 2023

1. Introduction

In modern theoretical physics, gauge symmetries play a very prominent role. The two most-fundamental theories we have, which describe almost all observed phenomena in nature—namely Einstein's theory of general relativity and the Standard Model of elementary particle physics—are gauge theories. From Maxwell's electrodynamics to various approaches to quantum gravity, gauge theories play a central role, and gauge symmetry represents one of their most-important aspects. In light of this, there is one class of gauge transformations that is often slightly neglected in the literature, due to their specific nature and properties.

In order to introduce this particular gauge symmetry in the most-elementary way possible, let us look at the following simple example. Every action $S[\phi_1, \phi_2]$, which depends on the fields $\phi_1(x)$ and $\phi_2(x)$, is invariant under the following gauge transformation:

$$\delta_0\phi_1(x) = \epsilon(x)\frac{\delta S}{\delta\phi_2(x)}\,, \qquad \delta_0\phi_2(x) = -\epsilon(x)\frac{\delta S}{\delta\phi_1(x)}\,, \tag{1}$$

as one can see by calculating the variation of the action:

$$\delta S[\phi_1,\phi_2] = \frac{\delta S}{\delta\phi_1}\delta_0\phi_1 + \frac{\delta S}{\delta\phi_2}\delta_0\phi_2 = 0\,. \tag{2}$$

This gauge symmetry exists for every action that is a functional of at least two fields, irrespective of any other gauge symmetry that the action may or may not have. In the literature, this symmetry is often called *trivial* gauge symmetry, since the form variations of the fields are identically zero on-shell. This is in contrast to all other gauge symmetries, which perform some nontrivial change of the fields on-shell.

It should be noted that, being trivial on-shell, the above transformations cannot play a role in obtaining any predictions about observables in a given theory, due to the intrinsic on-shell nature of the physical observables. For example, in practical situations of scattering experiments and measurements of cross-sections, this trivial symmetry is irrelevant. Nevertheless, when constructing a new theory, in general, the off-shell properties of the theory are important. As a typical example, path integral quantization prescription depends not only on the classical equations of motion, but on the whole action of the theory. In this sense, while these trivial transformations are not relevant for making predictions, they do have methodological relevance and value in theory construction, despite their on-shell triviality.

For example, these transformations in fact represent a very important part of the gauge symmetry for any theory and play a crucial role in various contexts, such as in the Batalin–Vilkovisky formalism (see [1] for a review and also the original papers [2–6]), or when discussing the diffeomorphism symmetry of the BF-like class of theories [7–11]. Furthermore, in general, a commutator of two ordinary gauge transformations will remain an ordinary gauge transformation only up to the above trivial transformations, meaning that the latter are important for the algebraic closure of all gauge transformations into a group.

To the best of our knowledge, the most-complete treatment and discussion of the above gauge transformations can be found in the book [12] by Marc Henneaux and Claudio Teitelboim. Therefore, in this paper, we opted to call them Henneaux–Teitelboim (HT) transformations. This naming can also be justified with the paper [7] by Gary Horowitz (published two years before the book [12]), where the author attributes these transformations to Henneaux and Teitelboim in a footnote and thanks them "for explaining this to me".

Regarding terminology, we should also note that we use the terms "gauge symmetry" and "gauge transformations" with a certain level of charity. Namely, one could argue that there are two distinct types of local symmetries—those that are obtained by a localization procedure from a corresponding global symmetry group (the procedure of "gauging" a global symmetry) and those that are intrinsically local, not obtained by any such localization procedure. It is not known whether HT symmetry belongs to the former or the latter class, since a global symmetry whose localization would give rise to HT transformations has not yet been shown to exist. Either way, in the literature, there is no established terminology that distinguishes the two classes of symmetries, and most often, both are called "gauge symmetries". Therefore, in what follows, for a lack of better terminology, we will adhere to this practice and describe HT transformations as a gauge symmetry.

In some of the modern approaches to the problem of quantum gravity based on the spinfoam formalism of loop quantum gravity [13,14], as well as in other applications of the so-called higher gauge theory (see [15] for a review and [16] for an application to quantum gravity), the description of gauge symmetry is being extended from the notion of a Lie group to different algebraic structures, called 2-groups, 3-groups, and in general, n-groups [17–27]. In this context, it is important to revisit and study the specific class of HT gauge symmetries, since they provide a nontrivial insight into the properties of these more general algebraic structures, as well as the physics behind the symmetries they describe.

The purpose of this paper is to provide a review of HT transformations in general and then discuss their properties and applications in two concrete models—the Chern–Simons theory and the $3BF$ theory. The Chern–Simons case is simple enough to serve as an illustrative toy example, while the $3BF$ theory represents a basis for the construction of a realistic theory of quantum gravity with matter within the context of the spinfoam formalism (see also [16,28–32]), discussing that its HT symmetry represents an important stepping stone towards the goal of a more realistic theory. The main result of this work represents a clarification of the structure of the gauge symmetry of a pure topological $3BF$ action, as well as the corresponding symmetry for the constrained $2BF$ action, which is classically equivalent to Einstein's general relativity. We also discuss in detail the relationship between diffeomorphism symmetry and the HT symmetry for the Chern–Simons and $3BF$ theories and offer some conceptual suggestions regarding the notion of gauge symmetry as it is being used in the literature.

The layout of the paper is as follows. In Section 2, we give a review of the general theory of HT transformations and their main properties. Section 3 is devoted to the example of HT symmetry in Chern–Simons theory, which is convenient due to its simplicity. In Section 4, we discuss the main case of HT symmetry in the $3BF$ and $2BF$ theories, which are important for applications in quantum gravity models. Finally, Section 5 contains an overview of the results, future research directions, and some concluding remarks.

The notation and conventions in the paper are as follows. When important, we assume the $(-,+,+,+)$ signature of the spacetime metric. The Greek indices from the middle of the alphabet, $\lambda, \mu, \nu, \dots$, represent spacetime indices and take values $0, 1, \dots, D-1$, where D is the dimension of the spacetime manifold $\mathcal{M}_D$ under consideration. The Greek indices from the beginning of the alphabet, $\alpha, \beta, \gamma, \dots$, represent group indices, as well as Latin indices $a, b, c, \dots$ and uppercase Latin indices $A, B, C, \dots$ and $I, J, K, \dots$. All these indices will be assigned to various Lie groups under consideration. Lowercase Latin indices from the middle of the alphabet, $i, j, k, \dots$, are generic and will be used to count all fields in a given theory or for some other purpose depending on the context. Throughout the paper, we denote the space of algebra-valued differential p-forms as

$$\mathcal{A}^p(\mathcal{M}, \mathfrak{a}) \equiv \Lambda^p(\mathcal{M}) \otimes \mathfrak{a} \,,$$

where $\Lambda^p(\mathcal{M})$ is the ordinary space of differential p-forms over the manifold $\mathcal{M}$, while $\mathfrak{a}$ is some Lie algebra.

2. Review of HT Symmetry

We begin by studying some basic general properties of HT transformations. After the definition, we demonstrate that the group of HT transformations represents a normal subgroup of the *total* gauge group of a given theory, and we discuss the triviality of HT transformations and that they exhaust all possible trivial transformations. Finally, before moving on to concrete theories, we study the subtleties of the dependence of HT symmetry on the choice of the action.

2.1. Definition of HT Transformations

Given an action $S[\phi^i]$ as a functional of fields $\phi^i(x)$ ($i \in \{1, \dots, N\}$ where we assume $N \geqslant 2$), the infinitesimal HT transformation is defined as

$$\phi^i(x) \to \phi'^i(x) = \phi^i(x) + \delta_0\phi^i(x) \,, \tag{3}$$

where the form variations of the fields are defined as

$$\delta_0\phi^i(x) = \epsilon^{ij}(x) \frac{\delta S}{\delta\phi^j(x)} \,. \tag{4}$$

The variation of the action under HT transformations then gives

$$\delta S = \frac{\delta S}{\delta\phi^i}\delta_0\phi^i = \frac{\delta S}{\delta\phi^i}\frac{\delta S}{\delta\phi^j}\epsilon^{ij} \,. \tag{5}$$

If the HT parameters are chosen to be antisymmetric,

$$\epsilon^{ij}(x) = -\epsilon^{ji}(x) \,, \tag{6}$$

the variation of the action (5) is identically zero, and HT transformations (4) represent a gauge symmetry of the theory.

The most-striking thing in the above definition is the fact that we did not specify the action in any way. Aside from the assumption $N \geqslant 2$, which excludes only actions describing a single real scalar field, every action is invariant with respect to the HT transformations. In other words, *HT transformations are a gauge symmetry of essentially every theory*.

The second striking property of the definition is that the form variations of fields become zero on-shell, according to (4). In this sense, the HT symmetry is sometimes called *trivial symmetry*, in contrast to ordinary gauge symmetries that a theory may have, which transform the fields in a nontrivial way on-shell. Triviality is also the reason why HT gauge symmetry does not feature in any way in the Hamiltonian analysis of a theory, so only the presence of ordinary gauge symmetries can be deduced from the Hamiltonian formalism.

2.2. HT Symmetry Group and Its Properties

There are two general properties that can be formulated for HT transformations. The first is that HT transformations form a normal subgroup within the full group of gauge symmetries, while the second is that HT transformations exhaust the set of all possible trivial transformations. The consequence of these properties is that one can always write the total symmetry group of any theory as

$$\mathcal{G}_{\text{total}} = \mathcal{G}_{\text{nontrivial}} \ltimes \mathcal{G}_{HT} \,, \tag{7}$$

where $\mathcal{G}_{\text{nontrivial}}$ is the symmetry group of ordinary gauge transformations (if there are any), $\mathcal{G}_{HT}$ is the HT symmetry group, and the symbol $\ltimes$ stands for a semidirect product. One can also reformulate (7) as

$$\mathcal{G}_{\text{nontrivial}} = \mathcal{G}_{\text{total}} / \mathcal{G}_{HT} \,, \tag{8}$$

so that the group of ordinary gauge symmetries is represented as a quotient group.

The easiest way to demonstrate (7) is to prove that the Lie algebra corresponding to $\mathcal{G}_{HT}$ represents an ideal within the Lie algebra corresponding to $\mathcal{G}_{\text{total}}$. To that end, pick an arbitrary form variation of fields that represents a symmetry of the action and write it in the form

$$\hat{\delta}_0 \phi^i(x) = F^i(x) \,, \qquad \text{such that} \qquad \hat{\delta} S = \frac{\delta S}{\delta \phi^i} F^i \equiv 0 \,. \tag{9}$$

Then, using (4), we can take concatenated variations of this form variation and the HT form variation as

$$\delta_0 \hat{\delta}_0 \phi^i = \frac{\delta F^i}{\delta \phi^j} \frac{\delta S}{\delta \phi^k} \epsilon^{jk} \,,$$

and

$$\hat{\delta}_0 \delta_o \phi^i = \frac{\delta}{\delta \phi^k} \left(\epsilon^{ij} \frac{\delta S}{\delta \phi^j} \right) F^k = \frac{\delta \epsilon^{ij}}{\delta \phi^k} \frac{\delta S}{\delta \phi^j} F^k + \epsilon^{ij} \frac{\delta}{\delta \phi^j} \left(\frac{\delta S}{\delta \phi^k} F^k \right) - \epsilon^{ij} \frac{\delta S}{\delta \phi^k} \frac{\delta F^k}{\delta \phi^j} \,.$$

The term in the second parentheses is zero by (9), so the commutator of two-form variations becomes

$$[\delta_0 \,, \hat{\delta}_0] \phi^i = \left(\epsilon^{jk} \frac{\delta F^i}{\delta \phi^j} - \epsilon^{ji} \frac{\delta F^k}{\delta \phi^j} - \frac{\delta \epsilon^{ik}}{\delta \phi^j} F^j \right) \frac{\delta S}{\delta \phi^k} \,, \tag{10}$$

which is again an HT transformation, since the expression in the parentheses is antisymmetric with respect to indices i, k. Therefore, the commutator is always an element of HT algebra, which means that HT algebra itself is an ideal of the total symmetry algebra. At the Lie group level, this translates into (7).

The second general property is the statement that there are no other trivial transformations beside the HT transformations. Assuming that some transformation described by the form variation $\bar{\delta}_0 \phi^i$ is a gauge symmetry of the action that vanishes on-shell, i.e., that it satisfies

$$\frac{\delta S}{\delta \phi^i} \bar{\delta}_0 \phi^i = 0 \,, \qquad \text{and} \qquad \bar{\delta}_0 \phi^i \approx 0 \,,$$

then one can prove that this transformation is an HT transformation, i.e., there exists a choice of antisymmetric HT parameters ϵ^{ij} such that the form variation $\bar{\delta}_0 \phi^i$ is of type (4):

$$\bar{\delta}_0 \phi^i = \epsilon^{ij} \frac{\delta S}{\delta \phi^j} \,. \tag{11}$$

Provided certain suitable regularity conditions for the action S, this statement can be rigorously formulated as a theorem. However, we omitted the proof since it is technical and off topic for the purposes of this paper. The interested reader can find the details of both the theorem and the proof in [12], Appendix 10.A.2.

To sum up, the first property (10) tells us that one can always factorize the total gauge symmetry group into the form (7), while the second property (11) guarantees that the quotient group (8) contains only nontrivial gauge transformations. This factorization of the total symmetry group is a key result that lays the groundwork for any subsequent analysis of HT transformations in particular and gauge symmetry in general.

2.3. Dependence of HT Symmetry on the Action

The final property of HT transformations that needs to be discussed is their dependence on the choice of the action. Suppose we are given some action $S_{\text{old}}[\phi^i]$, where $i \in \{1, \dots, N\}$, which has the corresponding HT transformation described as in (4):

$$\delta_0^{\text{old}} \phi^i = \epsilon^{ij} \frac{\delta S_{\text{old}}}{\delta \phi^j} . \tag{12}$$

Now, suppose that we modify that action into another one, $S_{\text{new}}[\phi^i, \chi^k]$, where $k \in \{N + 1, \dots, N + M\}$, by adding an extra term to the old action:

$$S_{\text{new}}[\phi^i, \chi^k] = S_{\text{old}}[\phi^i] + S_{\text{extra}}[\phi^i, \chi^k] . \tag{13}$$

Here, χ^j are additional fields that may be introduced into the new action. The HT transformation corresponding to the new action can be written in the block-matrix form, made of blocks of sizes N and M, as follows:

$$\begin{pmatrix} \delta_0^{\text{new}} \phi^i \\ \delta_0^{\text{new}} \chi^k \end{pmatrix} = \left(\begin{array}{c|c} \epsilon^{ij} & \zeta^{il} \\ \hline \theta^{kj} & \psi^{kl} \end{array}\right) \begin{pmatrix} \dfrac{\delta S_{\text{new}}}{\delta \phi^j} \\ \dfrac{\delta S_{\text{new}}}{\delta \chi^l} \end{pmatrix} , \qquad \begin{array}{l} i, j \in \{1, \dots, N\} , \\ k, l \in \{N + 1, \dots, N + M\} . \end{array} \tag{14}$$

Here, $\epsilon = -\epsilon^T$ is an antisymmetric $N \times N$ block of parameters ϵ^{ij}, ζ is a rectangular $N \times M$ block of parameters ζ^{il}, θ is a rectangular $M \times N$ block such that $\theta = -\zeta^T$, and finally, $\psi = -\psi^T$ is an antisymmetric $M \times M$ block of parameters ψ^{kl}. Overall, the total parameter matrix is antisymmetric, as required by (6).

The question one can now study is what is the relation between the two HT gauge symmetry groups $\mathcal{G}_{HT}^{\text{old}}$ and $\mathcal{G}_{HT}^{\text{new}}$ that correspond to the two actions. In practice, this question is most often relevant in cases when one introduces the piece S_{extra} as a gauge-fixing term, whose purpose is to break the ordinary gauge symmetry down to its subgroup:

$$G_{\text{nontrivial}}^{\text{new}} \subset G_{\text{nontrivial}}^{\text{old}} .$$

Naively, one might expect a similar relationship between the HT symmetry groups, $\mathcal{G}_{HT}^{\text{new}} \subset \mathcal{G}_{HT}^{\text{old}}$. However, looking at (12) and (14), this is obviously wrong. Namely, if $M \geqslant 1$, the HT symmetry of the new action is *larger* than the HT symmetry of the old action. Counting the number of independent parameters of both, one easily sees that

$$\dim(\mathcal{G}_{HT}^{\text{old}}) = \frac{N(N-1)}{2} , \qquad \dim(\mathcal{G}_{HT}^{\text{new}}) = \frac{(N+M)(N+M-1)}{2} ,$$

so that the only possible relationship would be the opposite, $\mathcal{G}_{HT}^{\text{old}} \subset \mathcal{G}_{HT}^{\text{new}}$. However, in fact, this can also be shown to be wrong. Namely, one can choose the extra parameters ζ, θ and ψ to be zero in (14), reducing it to the form that is formally similar to (12):

$$\delta_0^{\text{new}} \phi^i = \epsilon^{ij} \frac{\delta S_{\text{new}}}{\delta \phi^j} .$$

However, taking into account the relationship (13) between the two actions, the HT transformation takes the form

$$\delta_0^{\text{new}} \phi^i = \epsilon^{ij} \frac{\delta S_{\text{old}}}{\delta \phi^j} + \epsilon^{ij} \frac{\delta S_{\text{extra}}}{\delta \phi^j} \,,$$

which is explicitly different from (12), due to the presence of the term S_{extra} in the action. Therefore, the gauge group $\mathcal{G}_{HT}^{\text{old}}$ is not a subgroup of $\mathcal{G}_{HT}^{\text{new}}$ either.

The overall conclusion is that introducing additional terms to the action changes the total gauge symmetry in a nontrivial way. On the one hand, the ordinary gauge symmetry group typically becomes *smaller* due to explicit symmetry breaking by the extra term. On the other hand, the HT gauge symmetry group may become *larger* if the extra term contains additional fields, but either way becomes *different*, as a consequence of the very presence of the extra term. Given this, one can conclude that the *total* symmetry groups for the two actions will always be mutually different:

$$\mathcal{G}_{\text{total}}^{\text{new}} = \mathcal{G}_{\text{nontrivial}}^{\text{new}} \ltimes \mathcal{G}_{HT}^{\text{new}} \qquad \neq \qquad \mathcal{G}_{\text{total}}^{\text{old}} = \mathcal{G}_{\text{nontrivial}}^{\text{old}} \ltimes \mathcal{G}_{HT}^{\text{old}} \,.$$

Specifically, one cannot claim that the group $\mathcal{G}_{\text{total}}^{\text{old}}$ is being broken down into $\mathcal{G}_{\text{total}}^{\text{new}}$ as its subgroup; such a relationship may hold exclusively for the quotient groups of ordinary gauge transformations.

In the next two sections, we will turn to explicit examples of all general properties and features of the HT symmetry that have been discussed above. Moreover, we will also discuss some additional particular properties, such as the fact that some nontrivial gauge subgroups of $\mathcal{G}_{\text{total}}$ are not simultaneously subgroups of $\mathcal{G}_{\text{nontrivial}}$, which is a consequence of the semidirect product in (7). One such example will be the diffeomorphism symmetry in the Chern–Simons and $3BF$ actions.

Let us conclude this section with one conceptual comment. Throughout the literature, the typical practice is to always take the quotient between the total and HT symmetry groups as in (8), in order to isolate the nontrivial gauge transformations, and call the latter simply as the "gauge symmetry" of a theory. This approach is in fact advocated for in [12]. However, we believe that this practice can be misleading and that one should instead describe the group $\mathcal{G}_{\text{total}}$ as "the gauge symmetry" of a theory, explicitly including the HT subgroup as a legitimate gauge symmetry group. Namely, despite the fact that it is often called "trivial", the consequences of its presence in $\mathcal{G}_{\text{total}}$ are far from trivial. Granted, it may often be enough to discuss the gauge symmetry on-shell, and then, one can indeed calculate all symmetry transformations only "up to equations of motion", with no mention of the HT subgroup. However, whenever one needs to discuss the gauge transformations off-shell, the HT subgroup simply cannot be ignored anymore. Typical situations include the Batalin–Vilkovisky formalism [1], various generalizations of gauge symmetry in the context of higher gauge theories and quantum gravity [33], and even the traditional contexts such as the Coleman–Mandula theorem [34]. The situations in which HT transformations play a significant role may be rare, but nevertheless, they tend to be important. Thus, in our opinion, it would be prudent to always be aware that, for any given theory, its total gauge symmetry group is in fact bigger, and more feature-rich, than just the group of ordinary gauge transformations that are typically discussed in the literature.

3. HT Symmetry in Chern–Simons Theory

As an illustrative example of the general properties of HT symmetry from the previous section, let us discuss the HT transformations for the simple case of the Chern–Simons theory. The Chern–Simons theory represents an excellent toy example since it is well known in the literature and most readers should be familiar with it.

Given any Lie group G, its corresponding Lie algebra $\mathfrak{g}$, and a three-dimensional manifold $\mathcal{M}_3$, the Chern–Simons theory can be defined as a topological field theory over a trivial principal bundle $G \to \mathcal{M}_3$, given by the action:

$$S_{CS} = \int_{\mathcal{M}_3} \langle A \wedge \mathrm{d}A \rangle_{\mathfrak{g}} + \frac{1}{3} \langle A \wedge [A \wedge A] \rangle_{\mathfrak{g}} \,. \tag{15}$$

Here, $A \in \mathcal{A}^1(\mathcal{M}_3, \mathfrak{g})$ is a $\mathfrak{g}$-valued connection one-form over a manifold $\mathcal{M}_3$, and $\langle \_ , \_ \rangle_{\mathfrak{g}}$ is a G-invariant symmetric nondegenerate bilinear form on $\mathfrak{g}$. One often rewrites the Chern–Simons action within the framework of the enveloping algebra of $\mathfrak{g}$, introducing the notion of a *trace* as

$$\mathrm{Tr}(XY) \equiv \langle X , Y \rangle_{\mathfrak{g}} \,,$$

for every $X, Y \in \mathfrak{g}$. Then, the Chern–Simons action can be rewritten as

$$S_{CS} = \int_{\mathcal{M}_3} \mathrm{Tr}\left(A \wedge \mathrm{d}A + \frac{2}{3} A \wedge A \wedge A \right) , \tag{16}$$

where, for the second term, one employs the identity $\mathrm{Tr}(X[Y, Z]) = \mathrm{Tr}(XYZ) - \mathrm{Tr}(XZY)$ for every $X, Y, Z \in \mathfrak{g}$.

The gauge symmetry of the Chern–Simons action consists of G-gauge transformations, determined with the parameters ${\epsilon_{\mathfrak{g}}}^I(x)$. Using the basis of generators T_I to expand the connection A into components as

$$A = A^I{}_\mu(x)\, \mathrm{d}x^\mu \otimes T_I \,,$$

the form variation of the connection components $A^I{}_\mu$ corresponding to gauge transformations can then be written as

$$\delta_0 A^I{}_\mu = \partial_\mu {\epsilon_{\mathfrak{g}}}^I - f_{JK}{}^I {\epsilon_{\mathfrak{g}}}^J A^K{}_\mu \,, \tag{17}$$

where $f_{JK}{}^I$ are the structure constants corresponding to the generators T_I. Therefore, the gauge symmetry of the Chern–Simons theory is usually quoted as the initially chosen Lie group G:

$$\mathcal{G}_{CS} = G \,. \tag{18}$$

However, as we have seen in the previous section, this is not the complete set of gauge transformations, and the *total* gauge group should in fact be

$$\mathcal{G}_{\text{total}} = \mathcal{G}_{CS} \ltimes \mathcal{G}_{HT} \,. \tag{19}$$

Let us define the HT transformations for the Chern–Simons action (15). If we denote the dimension of the Lie algebra $\mathfrak{g}$ as $\dim(\mathfrak{g}) = p$, the number of independent field components $A^I{}_\mu$ is $N = 3p$. The HT transformation is then defined with the HT parameters $\epsilon^{IJ}{}_{\mu\nu}(x)$ as

$$\delta_0 A^I{}_\mu = \epsilon^{IJ}{}_{\mu\nu} \frac{\delta S}{\delta A^J{}_\nu} \,. \tag{20}$$

The requirement that the variation of the action vanishes:

$$\delta S = \frac{\delta S}{\delta A^I{}_\mu} \frac{\delta S}{\delta A^J{}_\nu} \epsilon^{IJ}{}_{\mu\nu} = 0 \,,$$

enforces the antisymmetry restriction on the HT parameters:

$$\epsilon^{IJ}{}_{\mu\nu} = -\epsilon^{JI}{}_{\nu\mu} \,.$$

Note that this equation can be satisfied in two different ways—the parameters can be either antisymmetric with respect to group indices IJ and symmetric with respect to spacetime indices $\mu\nu$, or vice versa. We, therefore, have two possible choices for their symmetry properties. The first possibility is defined as

$$\epsilon^{IJ}{}_{\mu\nu} = \epsilon^{IJ}{}_{\nu\mu} = -\epsilon^{JI}{}_{\mu\nu} = -\epsilon^{JI}{}_{\nu\mu}\,, \tag{21}$$

while the second possibility is defined as

$$\epsilon^{IJ}{}_{\mu\nu} = \epsilon^{JI}{}_{\mu\nu} = -\epsilon^{IJ}{}_{\nu\mu} = -\epsilon^{JI}{}_{\nu\mu}\,. \tag{22}$$

Varying the action, one obtains an explicit form of the HT transformation:

$$\delta_0 A^I{}_\mu = \epsilon^{IJ}{}_{\mu\nu}\varepsilon^{\nu\rho\sigma}\Big(\partial_\rho A_{J\sigma} - \partial_\sigma A_{J\rho} + f_{KLJ}\,A^K{}_\rho A^L{}_\sigma\Big)\,. \tag{23}$$

In order to demonstrate that HT transformations have highly nontrivial implications, despite being trivial on-shell, it is instructive to discuss diffeomorphisms. Namely, looking at the action (15), one expects that the theory has diffeomorphism symmetry, since it is formulated in a manifestly covariant way using differential forms. However, one can check that diffeomorphisms are not a subgroup of the ordinary gauge symmetry group $\mathcal{G}_{CS}$ given by (18), but nevertheless can be obtained as a subgroup of the total gauge group (19). In other words, one can demonstrate that

$$Diff(\mathcal{M}_3) \not\subset \mathcal{G}_{CS}\,, \qquad \text{but} \qquad Diff(\mathcal{M}_3) \subset \mathcal{G}_{\text{total}} = \mathcal{G}_{CS} \ltimes \mathcal{G}_{HT}\,.$$

Let us examine this in detail. The diffeomorphism transformation

$$x^\mu \to x'^\mu = x^\mu + \xi^\mu(x)\,, \tag{24}$$

determined by the parameter $\xi^\mu(x)$ represents a subgroup $Diff(\mathcal{M})$ of the full gauge symmetry of some given action, if for every field $\phi(x)$ in the theory and every choice of diffeomorphism parameters $\xi^\mu(x)$, there exists a choice of the gauge parameters $\epsilon^{\text{gauge}}(x)$ and the HT parameters $\epsilon^{\text{HT}}(x)$, such that:

$$\delta_0{}^{\text{diff}}\,\phi = \delta_0{}^{\text{gauge}}\phi + \delta_0{}^{\text{HT}}\phi\,. \tag{25}$$

In other words, if a theory has diffeomorphism symmetry, the diffeomorphism form variations of all the fields in the theory should be expressible in terms of their ordinary gauge and HT form variations.

In the case of Chern–Simons theory, this can be demonstrated explicitly. If one chooses the gauge parameters $\epsilon_{\mathfrak{g}}{}^I$ and the HT parameters $\epsilon^{IJ}{}_{\mu\nu}$ as

$$\epsilon_{\mathfrak{g}}{}^I = -\xi^\lambda A^I{}_\lambda\,, \qquad \epsilon^{IJ}{}_{\mu\nu} = -\frac{1}{2}\xi^\lambda\varepsilon_{\lambda\mu\nu}g^{IJ}\,, \tag{26}$$

where g^{IJ} is the inverse of $g_{IJ} \equiv \langle T_I\,, T_J\rangle_{\mathfrak{g}}$, one can apply Equations (25) using (17) and (23) to reproduce precisely the well-known diffeomorphism form variation of the connection $A^I{}_\mu$:

$$\delta_0{}^{\text{diff}}A^I{}_\mu = -A^I{}_\lambda\partial_\mu\xi^\lambda - \xi^\lambda\partial_\lambda A^I{}_\mu\,. \tag{27}$$

Therefore, as expected, despite the fact that $Diff(\mathcal{M}_3) \not\subset \mathcal{G}_{CS}$, one obtains that $Diff(\mathcal{M}_3) \subset \mathcal{G}_{\text{total}} = \mathcal{G}_{CS} \ltimes G_{HT}$. Note that the choice of HT parameters in (26) is nontrivial, which emphasizes the role of HT transformations and the fact that the full group of gauge symmetries is $\mathcal{G}_{\text{total}}$ rather than $\mathcal{G}_{CS}$. As we shall see in the next section, this property is not specific only to the Chern–Simons theory.

4. HT Symmetry in 3*BF* Theory

After discussing the Chern–Simons theory as a toy example, we move to the more important case of the 3*BF* theory. This theory is relevant for building models of quantum gravity; see [8,20,21,33,35]. Therefore, it is important to study its gauge symmetry and, in particular, the role of HT transformations.

4.1. Review of the 3BF Theory

Analogous to the fact that Chern–Simons theory is a topological theory based on a Lie group and a 3-dimensional manifold, the $3BF$ theory is also a topological theory based on a notion of a three-group and a 4-dimensional manifold. The notion of a three-group represents a categorical generalization of the notion of a group, in the context of higher gauge theory (HGT); see [15] for a review and motivation. For the purpose of defining the $3BF$ theory, we are interested in particular in a strict Lie three-group, which is known to be isomorphic to a so-called Lie two-crossed module; see [17–19] for details.

A Lie two-crossed module, denoted as $(L \xrightarrow{\delta} H \xrightarrow{\partial} G\,, \triangleright\,, \{\_\,,\_\}_{\mathrm{pf}})$, is an algebraic structure specified by three Lie groups G, H, and L, together with the homomorphisms $\delta : L \to H$ and $\partial : H \to G$, an action $\triangleright$ of the group G on all three groups, and a G-equivariant map, called the Peiffer lifting:

$$\{\_\,,\_\}_{\mathrm{pf}} : H \times H \to L\,.$$

In order for this structure to form a two-crossed module, the structure constants of algebras $\mathfrak{g}$, $\mathfrak{h}$, and $\mathfrak{l}$ (the Lie algebras corresponding to the Lie groups G, H, and L, respectively), as well as the maps ∂ and δ, the action $\triangleright$, and the Peiffer lifting, must satisfy certain axioms; see [20] for details.

Given a two-crossed module and a four-dimensional compact and orientable spacetime manifold $\mathcal{M}_4$, one can introduce the notion of a trivial principal three-bundle, in analogy with the notion of a trivial principal bundle constructed from an ordinary Lie group and a manifold; see [15]. Then, one can introduce the notion of a three-connection, an ordered triple (α, β, γ), where α, β, and γ are algebra-valued differential forms, $\alpha \in \mathcal{A}^1(\mathcal{M}_4, \mathfrak{g})$, $\beta \in \mathcal{A}^2(\mathcal{M}_4, \mathfrak{h})$, and $\gamma \in \mathcal{A}^3(\mathcal{M}_4, \mathfrak{l})$; see [17–19]. The corresponding fake hree-curvature $(\mathcal{F}, \mathcal{G}, \mathcal{H})$ is defined as:

$$\begin{gathered} \mathcal{F} = \mathrm{d}\alpha + \alpha \wedge \alpha - \partial\beta\,, \qquad \mathcal{G} = \mathrm{d}\beta + \alpha \wedge^{\triangleright} \beta - \delta\gamma\,, \\ \mathcal{H} = \mathrm{d}\gamma + \alpha \wedge^{\triangleright} \gamma + \{\beta \wedge \beta\}_{\mathrm{pf}}\,. \end{gathered} \tag{28}$$

Then, for a four-dimensional manifold $\mathcal{M}_4$, one can define the gauge-invariant topological $3BF$ action, based on the structure of a two-crossed module $(L \xrightarrow{\delta} H \xrightarrow{\partial} G\,, \triangleright\,, \{\_\,,\_\}_{\mathrm{pf}})$, by the action

$$S_{3BF} = \int_{\mathcal{M}_4} \langle B \wedge \mathcal{F} \rangle_{\mathfrak{g}} + \langle C \wedge \mathcal{G} \rangle_{\mathfrak{h}} + \langle D \wedge \mathcal{H} \rangle_{\mathfrak{l}}\,, \tag{29}$$

where $B \in \mathcal{A}^2(\mathcal{M}_4, \mathfrak{g})$, $C \in \mathcal{A}^1(\mathcal{M}_4, \mathfrak{h})$, and $D \in \mathcal{A}^0(\mathcal{M}_4, \mathfrak{l})$ are Lagrange multipliers and $\mathcal{F} \in \mathcal{A}^2(\mathcal{M}_4, \mathfrak{g})$, $\mathcal{G} \in \mathcal{A}^3(\mathcal{M}_4, \mathfrak{h})$, and $\mathcal{H} \in \mathcal{A}^4(\mathcal{M}_4, \mathfrak{l})$ represent the fake three-curvature given by Equation (28). The forms $\langle\_\,,\_\rangle_{\mathfrak{g}}$, $\langle\_\,,\_\rangle_{\mathfrak{h}}$, and $\langle\_\,,\_\rangle_{\mathfrak{l}}$ are G-invariant symmetric nondegenerate bilinear forms on $\mathfrak{g}$, $\mathfrak{h}$, and $\mathfrak{l}$, respectively. The action (29) is an example of the so-called higher gauge theory.

By choosing the three bases of generators $\tau_\alpha \in \mathfrak{g}$, $t_a \in \mathfrak{h}$, and $T_A \in \mathfrak{l}$ of the three respective Lie algebras, one can expand all fields in the theory into components as

$$\begin{aligned} B &= \frac{1}{2} B^\alpha{}_{\mu\nu}(x)\, \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \otimes \tau_\alpha\,, & \alpha &= \alpha^\alpha{}_\mu(x)\, \mathrm{d}x^\mu \otimes \tau_\alpha\,, \\ C &= C^a{}_\mu(x)\, \mathrm{d}x^\mu \otimes t_a\,, & \beta &= \frac{1}{2}\beta^a{}_{\mu\nu}(x)\, \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \otimes t_a\,, \\ D &= D^A(x) T_A\,, & \gamma &= \frac{1}{3!}\gamma^A{}_{\mu\nu\rho}(x)\, \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \wedge \mathrm{d}x^\rho \otimes T_A\,. \end{aligned}$$

One can also make use of the following notation for the components of all maps present in the theory, in the same three bases:

$$
\begin{array}{llll}
[\tau_\alpha , \tau_\beta] = f_{\alpha\beta}{}^{\gamma}\tau_\gamma , & g_{\alpha\beta} = \langle \tau_\alpha , \tau_\beta \rangle_{\mathfrak{g}} , & \tau_\alpha \triangleright \tau_\beta = \triangleright_{\alpha\beta}{}^{\gamma}\tau_\gamma , & \delta T_A = \delta_A{}^a t_a , \\
[t_a , t_b] = f_{ab}{}^{c} t_c , & g_{ab} = \langle t_a , t_b \rangle_{\mathfrak{h}} , & \tau_\alpha \triangleright t_a = \triangleright_{\alpha a}{}^{b} t_b , & \partial t_a = \partial_a{}^{\alpha}\tau_\alpha , \\
[T_A , T_B] = f_{AB}{}^{C} T_C , & g_{AB} = \langle T_A , T_B \rangle_{\mathfrak{l}} , & \tau_\alpha \triangleright T_A = \triangleright_{\alpha A}{}^{B} T_B , & \{t_a , t_b\}_{\mathrm{pf}} = X_{ab}{}^{A} T_A .
\end{array}
$$

The complete gauge symmetry of the $3BF$ action was studied in [8] using the techniques of Hamiltonian analysis. It consists of five types of gauge transformations, G-, H-, L-, M-, and N-gauge transformations, determined with the independent parameters $\epsilon_{\mathfrak{g}}{}^{\alpha}(x)$, $\epsilon_{\mathfrak{h}}{}^{a}{}_{\mu}(x)$, $\epsilon_{\mathfrak{l}}{}^{A}{}_{\mu\nu}(x)$, $\epsilon_{\mathfrak{m}}{}^{\alpha}{}_{\mu}(x)$, and $\epsilon_{\mathfrak{n}}{}^{a}(x)$, respectively. The form variations of the fields B, C, D, α, β, and γ, obtained in [8] are given as follows:

$$
\begin{array}{lcl}
\delta_0 B^{\alpha}{}_{\mu\nu} & = & f_{\beta\gamma}{}^{\alpha}\epsilon_{\mathfrak{g}}{}^{\beta} B^{\gamma}{}_{\mu\nu} + 2C_{a[\mu|}\epsilon_{\mathfrak{h}}{}^{b}{}_{|\nu]} \triangleright_{\beta b}{}^{a} g^{\alpha\beta} - D_A \triangleright_{\beta B}{}^{A}\epsilon_{\mathfrak{l}}{}^{B}{}_{\mu\nu} g^{\alpha\beta} - 2\nabla_{[\mu|}\epsilon_{\mathfrak{m}}{}^{\alpha}{}_{|\nu]} \\
 & & +\beta_{b\mu\nu} \triangleright_{\beta a}{}^{b}\epsilon_{\mathfrak{n}}{}^{a} g^{\alpha\beta} , \\
\delta_0 C^{a}{}_{\mu} & = & \triangleright_{\alpha b}{}^{a}\epsilon_{\mathfrak{g}}{}^{\alpha} C^{b}{}_{\mu} + 2D_A X_{(ab)}{}^{A}\epsilon_{\mathfrak{h}}{}^{b}{}_{\mu} - \partial^{a}{}_{\alpha}\epsilon_{\mathfrak{m}}{}^{\alpha}{}_{\mu} - \nabla_{\mu}\epsilon_{\mathfrak{n}}{}^{a} , \\
\delta_0 D^{A} & = & \triangleright_{\alpha B}{}^{A}\epsilon_{\mathfrak{g}}{}^{\alpha} D^{B} + \delta^{A}{}_{a}\epsilon_{\mathfrak{n}}{}^{a} , \\
\delta_0 \alpha^{\alpha}{}_{\mu} & = & -\partial_{\mu}\epsilon_{\mathfrak{g}}{}^{\alpha} - f_{\beta\gamma}{}^{\alpha}\alpha^{\beta}{}_{\mu}\epsilon_{\mathfrak{g}}{}^{\gamma} - \partial_{a}{}^{\alpha}\epsilon_{\mathfrak{h}}{}^{a}{}_{\mu} , \\
\delta_0 \beta^{a}{}_{\mu\nu} & = & \triangleright_{\alpha b}{}^{a}\epsilon_{\mathfrak{g}}{}^{\alpha}\beta^{b}{}_{\mu\nu} - 2\nabla_{[\mu|}\epsilon_{\mathfrak{h}}{}^{a}{}_{|\nu]} + \delta_{A}{}^{a}\epsilon_{\mathfrak{l}}{}^{A}{}_{\mu\nu} , \\
\delta_0 \gamma^{A}{}_{\mu\nu\rho} & = & \triangleright_{\alpha B}{}^{A}\epsilon_{\mathfrak{g}}{}^{\alpha}\gamma^{B}{}_{\mu\nu\rho} + 3!\beta^{a}{}_{[\mu\nu}\epsilon_{\mathfrak{h}}{}^{b}{}_{\rho]} X_{(ab)}{}^{A} + \nabla_{\mu}\epsilon_{\mathfrak{l}}{}^{A}{}_{\nu\rho} - \nabla_{\nu}\epsilon_{\mathfrak{l}}{}^{A}{}_{\mu\rho} + \nabla_{\rho}\epsilon_{\mathfrak{l}}{}^{A}{}_{\mu\nu} .
\end{array}
\tag{30}
$$

The gauge transformations (30) form a group $\mathcal{G}_{3BF}$:

$$
\mathcal{G}_{3BF} = \tilde{G} \ltimes (\tilde{H}_L \ltimes (\tilde{N} \times \tilde{M})) , \tag{31}
$$

where $\tilde{G}$ denotes the group of G-gauge transformations, the H-gauge transformations together with the L-gauge transformations form the group $\tilde{H}_L$, while $\tilde{M}$ and $\tilde{N}$ are the groups of M- and N-gauge transformations, respectively. All these groups are determined from the structure of the initial chosen two-crossed module that defines the theory; see [8] for details.

However, as we have seen in the general theory in Section 2 and in the example of the Chern–Simons theory in Section 3, the symmetry group $\mathcal{G}_{3BF}$ determined by the Hamiltonian analysis does not include HT transformations, and therefore, the *total* gauge group should in fact be

$$
\mathcal{G}_{\mathrm{total}} = \mathcal{G}_{3BF} \ltimes \mathcal{G}_{HT} . \tag{32}
$$

4.2. Explicit HT Transformations

Let us explicitly define the HT transformations for the $3BF$ action (29). If we denote the dimensions of the Lie algebras $\mathfrak{g}, \mathfrak{h}, \mathfrak{l}$ as

$$
\dim(\mathfrak{g}) = p , \qquad \dim(\mathfrak{h}) = q , \qquad \dim(\mathfrak{l}) = r ,
$$

the number of independent field components in the theory can be counted according to the following table:

$B^{\alpha}{}_{\mu\nu}$	$C^{a}{}_{\mu}$	D^{A}	$\alpha^{\alpha}{}_{\mu}$	$\beta^{a}{}_{\mu\nu}$	$\gamma^{A}{}_{\mu\nu\rho}$
$6p$	$4q$	r	$4p$	$6q$	$4r$

The total number of independent field components is, therefore,

$$
N = 6p + 4q + r + 4p + 6q + 4r = 10p + 10q + 5r .
$$

Let ϕ^i denote all field components, where $i = 1, 2, \ldots, N$. We can write the fields schematically as a column-matrix with six blocks:

$$\phi^i = \begin{pmatrix} B^{\alpha}{}_{\mu\nu} \\ C^{a}{}_{\mu} \\ D^{A} \\ \alpha^{\alpha}{}_{\mu} \\ \beta^{a}{}_{\mu\nu} \\ \gamma^{A}{}_{\mu\nu\rho} \end{pmatrix}.$$

The HT transformation is then defined via the parameters $\epsilon^{ij}(x)$ as

$$\delta_0 \phi^i = \epsilon^{ij} \frac{\delta S}{\delta \phi^j}.$$

The requirement that the variation of the action vanishes enforces the antisymmetry restriction on the parameters, $\epsilon^{ij} = -\epsilon^{ji}$, for all $i, j \in \{1, \ldots, N\}$. These transformations can be represented more explicitly as a tensorial 6×6 block-matrix equation, in the following form:

$$\begin{pmatrix} \delta_0 B^{\alpha}{}_{\mu\nu} \\ \delta_0 C^{a}{}_{\mu} \\ \delta_0 D^{A} \\ \delta_0 \alpha^{\alpha}{}_{\mu} \\ \delta_0 \beta^{a}{}_{\mu\nu} \\ \delta_0 \gamma^{A}{}_{\mu\nu\rho} \end{pmatrix} = \left(\begin{array}{c|c|c|c|c|c} \epsilon^{\alpha\beta}{}_{\mu\nu\sigma\lambda} & \epsilon^{\alpha b}{}_{\mu\nu\sigma} & \epsilon^{\alpha B}{}_{\mu\nu} & \epsilon^{\alpha\beta}{}_{\mu\nu\sigma} & \epsilon^{\alpha b}{}_{\mu\nu\sigma\lambda} & \epsilon^{\alpha B}{}_{\mu\nu\sigma\lambda\xi} \\ \hline \mu^{a\beta}{}_{\mu\sigma\lambda} & \epsilon^{ab}{}_{\mu\sigma} & \epsilon^{aB}{}_{\mu} & \epsilon^{a\beta}{}_{\mu\sigma} & \epsilon^{ab}{}_{\mu\sigma\lambda} & \epsilon^{aB}{}_{\mu\sigma\lambda\xi} \\ \hline \mu^{A\beta}{}_{\sigma\lambda} & \mu^{Ab}{}_{\sigma} & \epsilon^{AB} & \epsilon^{A\beta}{}_{\sigma} & \epsilon^{Ab}{}_{\sigma\lambda} & \epsilon^{AB}{}_{\sigma\lambda\xi} \\ \hline \mu^{\alpha\beta}{}_{\mu\sigma\lambda} & \mu^{\alpha b}{}_{\mu\sigma} & \mu^{\alpha B}{}_{\mu} & \epsilon^{\alpha\beta}{}_{\mu\sigma} & \epsilon^{\alpha b}{}_{\mu\sigma\lambda} & \epsilon^{\alpha B}{}_{\mu\sigma\lambda\xi} \\ \hline \mu^{a\beta}{}_{\mu\nu\sigma\lambda} & \mu^{ab}{}_{\mu\nu\sigma} & \mu^{aB}{}_{\mu\nu} & \mu^{a\beta}{}_{\mu\nu\sigma} & \epsilon^{ab}{}_{\mu\nu\sigma\lambda} & \epsilon^{aB}{}_{\mu\nu\sigma\lambda\xi} \\ \hline \mu^{A\beta}{}_{\mu\nu\rho\sigma\lambda} & \mu^{Ab}{}_{\mu\nu\rho\sigma} & \mu^{AB}{}_{\mu\nu\rho} & \mu^{A\beta}{}_{\mu\nu\rho\sigma} & \mu^{Ab}{}_{\mu\nu\rho\sigma\lambda} & \epsilon^{AB}{}_{\mu\nu\rho\sigma\lambda\xi} \end{array}\right) \begin{pmatrix} \frac{1}{2} \frac{\delta S}{\delta B^{\beta}{}_{\sigma\lambda}} \\ \frac{\delta S}{\delta C^{b}{}_{\sigma}} \\ \frac{\delta S}{\delta D^{B}} \\ \frac{\delta S}{\delta \alpha^{\beta}{}_{\sigma}} \\ \frac{1}{2} \frac{\delta S}{\delta \beta^{b}{}_{\sigma\lambda}} \\ \frac{1}{3!} \frac{\delta S}{\delta \gamma^{B}{}_{\sigma\lambda\xi}} \end{pmatrix}. \quad (33)$$

The coefficients multiplying the variations of the action in the column on the right-hand side are there to compensate the overcounting of the independent field components. Due to the antisymmetry of HT parameters, all μ blocks (below the diagonal) are determined in terms of the ϵ blocks (above the diagonal), as follows. For the first column of the parameter matrix in (33), we have:

$$\begin{gathered} \mu^{b\alpha}{}_{\sigma\mu\nu} = -\epsilon^{\alpha b}{}_{\mu\nu\sigma}, \qquad \mu^{B\alpha}{}_{\mu\nu} = -\epsilon^{\alpha B}{}_{\mu\nu}, \qquad \mu^{\beta\alpha}{}_{\sigma\mu\nu} = -\epsilon^{\alpha\beta}{}_{\mu\nu\sigma}, \\ \mu^{b\alpha}{}_{\sigma\lambda\mu\nu} = -\epsilon^{\alpha b}{}_{\mu\nu\sigma\lambda}, \qquad \mu^{B\alpha}{}_{\sigma\lambda\xi\mu\nu} = -\epsilon^{\alpha B}{}_{\mu\nu\sigma\lambda\xi}. \end{gathered} \quad (34)$$

For the second column, we have:

$$\begin{gathered} \mu^{Ba}{}_{\mu} = -\epsilon^{aB}{}_{\mu}, \qquad \mu^{\beta a}{}_{\sigma\mu} = -\epsilon^{a\beta}{}_{\mu\sigma}, \\ \mu^{ba}{}_{\sigma\lambda\mu} = -\epsilon^{ab}{}_{\mu\sigma\lambda}, \qquad \mu^{Ba}{}_{\sigma\lambda\xi\mu} = -\epsilon^{aB}{}_{\mu\sigma\lambda\xi}. \end{gathered} \quad (35)$$

The μ parameters in the third column are determined via:

$$\mu^{\beta A}{}_{\sigma} = -\epsilon^{A\beta}{}_{\sigma}, \qquad \mu^{bA}{}_{\sigma\lambda} = -\epsilon^{Ab}{}_{\sigma\lambda}, \qquad \mu^{BA}{}_{\sigma\lambda\xi} = -\epsilon^{AB}{}_{\sigma\lambda\xi}, \quad (36)$$

while the remaining μ parameters in the fourth and fifth columns are determined as:

$$\mu^{b\alpha}{}_{\sigma\lambda\mu} = -\epsilon^{\alpha b}{}_{\mu\sigma\lambda}, \qquad \mu^{B\alpha}{}_{\sigma\lambda\xi\mu} = -\epsilon^{\alpha B}{}_{\mu\sigma\lambda\xi}, \qquad \mu^{Ba}{}_{\sigma\lambda\xi\mu\nu} = -\epsilon^{aB}{}_{\mu\nu\sigma\lambda\xi}. \quad (37)$$

Finally, in addition to all these, the parameters in the blocks on the diagonal also have to satisfy certain antisymmetry relations, specifically:

$$
\begin{gathered}
\epsilon^{\alpha\beta}{}_{\mu\nu\sigma\lambda} = -\epsilon^{\beta\alpha}{}_{\sigma\lambda\mu\nu}\,, \qquad \epsilon^{ab}{}_{\mu\sigma} = -\epsilon^{ba}{}_{\sigma\mu}\,, \qquad \epsilon^{AB} = -\epsilon^{BA}\,, \\
\epsilon^{\alpha\beta}{}_{\mu\sigma} = -\epsilon^{\beta\alpha}{}_{\sigma\mu}\,, \qquad \epsilon^{ab}{}_{\mu\nu\sigma\lambda} = -\epsilon^{ba}{}_{\sigma\lambda\mu\nu}\,, \qquad \epsilon^{AB}{}_{\mu\nu\rho\sigma\lambda\xi} = -\epsilon^{BA}{}_{\sigma\lambda\xi\mu\nu\rho}\,.
\end{gathered}
\tag{38}
$$

Like in the example of the Chern–Simons theory from the previous section, these antisymmetry relations can be satisfied in various multiple ways. All those possibilities are allowed, as long as the identities (38) are satisfied. The final ingredient in (33) is the expressions for the variation of the action with respect to the fields, and these are given as follows:

$$
\begin{aligned}
\frac{\delta S}{\delta B^{\beta}{}_{\nu\rho}} &= \frac{1}{2}\varepsilon^{\nu\rho\sigma\tau}\mathcal{F}_{\beta\sigma\tau}\,, \\
\frac{\delta S}{\delta C^{b}{}_{\rho}} &= \frac{1}{3!}\varepsilon^{\rho\sigma\tau\lambda}\mathcal{G}_{b\sigma\tau\lambda}\,, \\
\frac{\delta S}{\delta D^{B}} &= \frac{1}{4!}\varepsilon^{\sigma\tau\lambda\xi}\mathcal{H}_{B\sigma\tau\lambda\xi}\,, \\
\frac{\delta S}{\delta \alpha^{\beta}{}_{\rho}} &= \frac{1}{2}\varepsilon^{\rho\tau\lambda\xi}\left(\nabla_{\tau}B_{\beta\lambda\xi} - \rhd_{\beta a}{}^{b}C_{b\tau}\beta^{a}{}_{\lambda\xi} + \frac{1}{3}\rhd_{\beta B}{}^{A}D_{A}\gamma^{B}{}_{\tau\lambda\xi}\right), \\
\frac{\delta S}{\delta \beta^{b}{}_{\nu\rho}} &= \varepsilon^{\nu\rho\sigma\tau}\left(\nabla_{\sigma}C_{b\tau} - \frac{1}{2}\partial_{b}{}^{\alpha}B_{\alpha\sigma\tau} + X_{(ab)}{}^{A}D_{A}\beta^{b}{}_{\sigma\tau}\right), \\
\frac{\delta S}{\delta \gamma^{B}{}_{\mu\nu\rho}} &= \varepsilon^{\mu\nu\rho\sigma}(\nabla_{\sigma}D_{B} + \delta_{B}{}^{a}C_{a\sigma})\,.
\end{aligned}
\tag{39}
$$

4.3. Diffeomorphisms

As in the case of the Chern–Simons theory, it is instructive to discuss diffeomorphism symmetry. The $3BF$ action (29) obviously is diffeomorphism invariant, since it is formulated in a manifestly covariant way, using differential forms. However, one can check that the diffeomorphisms are not a subgroup of the gauge symmetry group $\mathcal{G}_{3BF}$ given by Equation (31), but nevertheless can be obtained as a subgroup of the total gauge group (32):

$$
Diff(\mathcal{M}_4) \not\subset \mathcal{G}_{3BF}\,, \qquad \text{but} \qquad Diff(\mathcal{M}_4) \subset \mathcal{G}_{\text{total}} = \mathcal{G}_{3BF} \ltimes \mathcal{G}_{HT}\,. \tag{40}
$$

Let us demonstrate this. Like in the Chern–Simons case, we want to demonstrate that the form variation of all fields corresponding to diffeomorphisms can be obtained as a suitable combination of the form variations for the ordinary gauge transformations (30) and the HT transformations (33). In other words, for an arbitrary choice of the diffeomorphism parameters $\xi^{\mu}(x)$ from (24), Equation (25) should hold in the case of the $3BF$ theory as well:

$$
\delta_0{}^{\text{diff}}\,\phi = \delta_0{}^{\text{gauge}}\phi + \delta_0{}^{\text{HT}}\phi\,. \tag{41}
$$

Indeed, this can be shown by a suitable choice of parameters. Regarding the parameters of the gauge transformations (30), the appropriate choice is given as:

$$
\begin{gathered}
\epsilon_{\mathfrak{g}}{}^{\alpha} = \xi^{\lambda}\alpha^{\alpha}{}_{\lambda}\,, \qquad \epsilon_{\mathfrak{h}}{}^{a}{}_{\mu} = -\xi^{\lambda}\beta^{a}{}_{\mu\lambda}\,, \qquad \epsilon_{\mathfrak{l}}{}^{A}{}_{\mu\nu} = -\xi^{\lambda}\gamma^{A}{}_{\mu\nu\lambda}\,, \\
\epsilon_{\mathfrak{m}}{}^{\alpha}{}_{\mu} = -\xi^{\lambda}B^{\alpha}{}_{\mu\lambda}\,, \qquad \epsilon_{\mathfrak{n}}{}^{a} = \xi^{\lambda}C^{a}{}_{\lambda}\,.
\end{gathered}
\tag{42}
$$

Regarding the parameters of the HT transformations (33), we chose the following special case, with the majority of the parameters equated to zero:

$$\begin{pmatrix} \delta_0 B^{\alpha}{}_{\mu\nu} \\ \delta_0 C^{a}{}_{\mu} \\ \delta_0 D^{A} \\ \delta_0 \alpha^{\alpha}{}_{\mu} \\ \delta_0 \beta^{a}{}_{\mu\nu} \\ \delta_0 \gamma^{A}{}_{\mu\nu\rho} \end{pmatrix} = \left(\begin{array}{c|c|c|c|c|c} 0 & 0 & 0 & \epsilon^{\alpha\beta}{}_{\mu\nu\sigma} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \epsilon^{ab}{}_{\mu\sigma\lambda} & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & \epsilon^{AB}{}_{\sigma\lambda\xi} \\ \hline \mu^{\alpha\beta}{}_{\mu\sigma\lambda} & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & \mu^{ab}{}_{\mu\nu\sigma} & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & \mu^{AB}{}_{\mu\nu\rho} & 0 & 0 & 0 \end{array}\right) \begin{pmatrix} \frac{1}{2}\frac{\delta S}{\delta B^{\beta}{}_{\sigma\lambda}} \\ \frac{\delta S}{\delta C^{b}{}_{\sigma}} \\ \frac{\delta S}{\delta D^{B}} \\ \frac{\delta S}{\delta \alpha^{\beta}{}_{\sigma}} \\ \frac{1}{2}\frac{\delta S}{\delta \beta^{b}{}_{\sigma\lambda}} \\ \frac{1}{3!}\frac{\delta S}{\delta \gamma^{B}{}_{\sigma\lambda\xi}} \end{pmatrix} . \tag{43}$$

Of course, due to antisymmetry, the nonzero μ blocks take negative values of the corresponding ϵ blocks, in accordance with (34), (35), and (36). The three independent nonzero ϵ blocks are chosen as

$$\epsilon^{\alpha\beta}{}_{\mu\nu\sigma} = \xi^{\rho} g^{\alpha\beta} \varepsilon_{\mu\nu\sigma\rho} \,, \qquad \epsilon^{ab}{}_{\mu\sigma\lambda} = \xi^{\rho} g^{ab} \varepsilon_{\rho\mu\sigma\lambda} \,, \qquad \epsilon^{AB}{}_{\sigma\lambda\xi} = \xi^{\rho} g^{AB} \varepsilon_{\sigma\lambda\xi\rho} \,. \tag{44}$$

Finally, substituting (42) and (44) into (30) and (43), respectively, and then substituting all those results into (41), after a certain amount of work, one obtains precisely the standard form variations corresponding to diffeomorphisms:

$$\begin{aligned} \delta_0^{\text{diff}} B^{\alpha}{}_{\mu\nu} &= -B^{\alpha}{}_{\lambda\nu}\partial_{\mu}\xi^{\lambda} - B^{\alpha}{}_{\mu\lambda}\partial_{\nu}\xi^{\lambda} - \xi^{\lambda}\partial_{\lambda}B^{\alpha}{}_{\mu\nu} \,, \\ \delta_0^{\text{diff}} C^{a}{}_{\mu} &= -C^{a}{}_{\lambda}\partial_{\mu}\xi^{\lambda} - \xi^{\lambda}\partial_{\lambda}C^{a}{}_{\mu} \,, \\ \delta_0^{\text{diff}} D^{A} &= -\xi^{\lambda}\partial_{\lambda}D^{A} \,, \\ \delta_0^{\text{diff}} \alpha^{\alpha}{}_{\mu} &= -\alpha^{\alpha}{}_{\lambda}\partial_{\mu}\xi^{\lambda} - \xi^{\lambda}\partial_{\lambda}\alpha^{\alpha}{}_{\mu} \,, \\ \delta_0^{\text{diff}} \beta^{a}{}_{\mu\nu} &= -\beta^{a}{}_{\lambda\nu}\partial_{\mu}\xi^{\lambda} - \beta^{a}{}_{\mu\lambda}\partial_{\nu}\xi^{\lambda} - \xi^{\lambda}\partial_{\lambda}\beta^{a}{}_{\mu\nu} \,, \\ \delta_0^{\text{diff}} \gamma^{A}{}_{\mu\nu\rho} &= -\gamma^{A}{}_{\lambda\nu\rho}\partial_{\mu}\xi^{\lambda} - \gamma^{A}{}_{\mu\lambda\rho}\partial_{\nu}\xi^{\lambda} - \gamma^{A}{}_{\mu\nu\lambda}\partial_{\rho}\xi^{\lambda} - \xi^{\lambda}\partial_{\lambda}\gamma^{A}{}_{\mu\nu\rho} \,. \end{aligned} \tag{45}$$

This establishes both relations (40), as we set out to demonstrate. We note again that the HT transformations play a crucial role in obtaining the result, since we had to choose the parameters (44) in a nontrivial manner.

4.4. Symmetry Breaking in 2BF Theory

Let us now turn to the topic of symmetry breaking and the way it influences HT transformations. To that end, we studied the topological $2BF$ action, which is a special case of the $3BF$ action (29) without the last term:

$$S_{2BF} = \int_{\mathcal{M}_4} \langle B \wedge \mathcal{F} \rangle_{\mathfrak{g}} + \langle C \wedge \mathcal{G} \rangle_{\mathfrak{h}} \,. \tag{46}$$

In order to be even more concrete, let us fix a two-crossed module structure with the following choice of groups:

$$G = SO(3,1) \,, \qquad H = \mathbb{R}^4 \,, \qquad L = \{e\} \,.$$

In other words, we interpret group G as the Lorentz group, group H as the spacetime translations group, while group L is trivial, for simplicity. This choice corresponds to the so-called Poincaré two-group; see [16] for details. Since the generators of the Lorentz group can be conveniently counted using the antisymmetric combinations of indices from the group of translations, instead of the G-group indices α, we shall systematically write $[ab] \in \{01, 02, 03, 12, 13, 23\}$, where $a, b \in \{0, 1, 2, 3\}$ are H-group indices, and the brackets denote antisymmetrization. With a further change in notation from the connection 1-form α to the spin-connection 1-form ω, the curvature 2-form $\mathcal{F}(\alpha)$ to $R(\omega)$, and interpreting the Lagrange multiplier 1-form C as the tetrad 1-form e, the $2BF$ action can be rewritten in new notation as

$$S_{2BF} = \int_{\mathcal{M}_4} B^{[ab]} \wedge R_{[ab]} + e^a \wedge \mathcal{G}_a \,. \tag{47}$$

The ordinary gauge symmetry group for this action has a form similar to (31):

$$\mathcal{G}_{2BF} = \tilde{G} \ltimes (\tilde{H} \ltimes (\tilde{N} \times \tilde{M}))\,, \tag{48}$$

while the total group of gauge symmetries is extended by the HT transformations, so that

$$\mathcal{G}_{\text{total}} = \mathcal{G}_{2BF} \ltimes \mathcal{G}_{HT} \,. \tag{49}$$

The explicit HT transformations are written as a tensorial 4×4 block-matrix equation, in the form

$$\begin{pmatrix} \delta_0 B^{[ab]}{}_{\mu\nu} \\ \delta_0 e^a{}_\mu \\ \delta_0 \omega^{[ab]}{}_\mu \\ \delta_0 \beta^a{}_{\mu\nu} \end{pmatrix} = \left(\begin{array}{c|c|c|c} \epsilon^{[ab][cd]}{}_{\mu\nu\sigma\lambda} & \epsilon^{[ab]c}{}_{\mu\nu\sigma} & \epsilon^{[ab][cd]}{}_{\mu\nu\sigma} & \epsilon^{[ab]c}{}_{\mu\nu\sigma\lambda} \\ \hline \mu^{a[cd]}{}_{\mu\sigma\lambda} & \epsilon^{ac}{}_{\mu\sigma} & \epsilon^{a[cd]}{}_{\mu\sigma} & \epsilon^{ac}{}_{\mu\sigma\lambda} \\ \hline \mu^{[ab][cd]}{}_{\mu\sigma\lambda} & \mu^{[ab]c}{}_{\mu\sigma} & \epsilon^{[ab][cd]}{}_{\mu\sigma} & \epsilon^{[ab]c}{}_{\mu\sigma\lambda} \\ \hline \mu^{a[cd]}{}_{\mu\nu\sigma\lambda} & \mu^{ac}{}_{\mu\nu\sigma} & \mu^{a[cd]}{}_{\mu\nu\sigma} & \epsilon^{ac}{}_{\mu\nu\sigma\lambda} \end{array}\right) \begin{pmatrix} \frac{1}{4}\frac{\delta S}{\delta B^{[cd]}{}_{\sigma\lambda}} \\ \frac{\delta S}{\delta e^c{}_\sigma} \\ \frac{1}{2}\frac{\delta S}{\delta \omega^{[cd]}{}_\sigma} \\ \frac{1}{2}\frac{\delta S}{\delta \beta^c{}_{\sigma\lambda}} \end{pmatrix}, \tag{50}$$

where the usual antisymmetry rules apply. Here, we have

$$\begin{aligned} \frac{\delta S}{\delta B^{[cd]}{}_{\sigma\lambda}} &= \varepsilon^{\mu\nu\sigma\lambda} R_{[cd]\mu\nu}\,, \\ \frac{\delta S}{\delta \omega^{[cd]}{}_\sigma} &= \varepsilon^{\sigma\mu\nu\rho}\left(\nabla_\mu B_{[cd]\nu\rho} - e_{[c|\mu}\beta_{|d]\nu\rho}\right), \\ \frac{\delta S}{\delta e^c{}_\sigma} &= \frac{1}{2}\varepsilon^{\sigma\mu\nu\rho}\nabla_\mu \beta_{c\nu\rho}\,, \\ \frac{\delta S}{\delta \beta^c{}_{\sigma\lambda}} &= \varepsilon^{\mu\nu\sigma\lambda}\nabla_\mu e_{c\nu}\,. \end{aligned} \tag{51}$$

The $2BF$ action (46) is topological, in the sense that it has no local propagating degrees of freedom. In this sense, it does not represent a theory of any realistic physics. In order to construct a more realistic theory, one proceeds by introducing the so-called *simplicity constraint* term into the action, which changes the equations of motion of the theory so that it does have nontrivial degrees of freedom. An example is the action

$$S_{GR} = \int_{\mathcal{M}_4} B^{[ab]} \wedge R_{[ab]} + e^a \wedge \nabla \beta_a - \lambda_{[ab]} \wedge \left(B^{[ab]} - \frac{1}{16\pi l_p^2}\varepsilon^{abcd} e_c \wedge e_d \right), \tag{52}$$

where the new constraint term features another Lagrange multiplier two-form $\lambda_{[ab]}$. By virtue of the simplicity constraint, the theory becomes equivalent to general relativity, in the sense that the corresponding equations of motion reduce to vacuum Einstein field equations (see [16] for the analysis and proof). In this sense, constraint terms of various types are important when building more realistic theories; see [20] for more examples.

However, adding the simplicity constraint term also changes the gauge symmetry of the theory. In particular, it breaks the gauge group $\mathcal{G}_{2BF}$ from (48) down to one of its subgroups, so that the symmetry group of the action S_{GR} is

$$\mathcal{G}_{GR} \subset \mathcal{G}_{2BF} \,. \tag{53}$$

This is expected and unsurprising. What is less obvious, however, is that the group of HT transformations $\tilde{\mathcal{G}}_{HT}$ of the action S_{GR} *is not a subgroup* of the HT group $\mathcal{G}_{HT}$ of the original action S_{2BF}:

$$\tilde{\mathcal{G}}_{HT} \not\subset \mathcal{G}_{HT}\,, \tag{54}$$

which implies that

$$\mathcal{G}^{GR}_{\text{total}} \not\subset \mathcal{G}^{2BF}_{\text{total}} , \tag{55}$$

despite (53).

Let us demonstrate this. Since the action (52) features an additional field $\lambda^{[ab]}{}_{\mu\nu}(x)$, the HT transformations (50) have to be modified to take this into account and obtain the following 5×5 block-matrix form:

$$\begin{pmatrix} \delta_0 B^{[ab]}{}_{\mu\nu} \\ \delta_0 e^a{}_\mu \\ \delta_0 \omega^{[ab]}{}_\mu \\ \delta_0 \beta^a{}_{\mu\nu} \\ \delta_0 \lambda^{[ab]}{}_{\mu\nu} \end{pmatrix} = \left(\begin{array}{c|c|c|c|c} \epsilon^{[ab][cd]}{}_{\mu\nu\sigma\lambda} & \epsilon^{[ab]c}{}_{\mu\nu\sigma} & \epsilon^{[ab][cd]}{}_{\mu\nu\sigma} & \epsilon^{[ab]c}{}_{\mu\nu\sigma\lambda} & \zeta^{[ab][cd]}{}_{\mu\nu\sigma\xi} \\ \hline \mu^{a[cd]}{}_{\mu\sigma\lambda} & \epsilon^{ac}{}_{\mu\sigma} & \epsilon^{a[cd]}{}_{\mu\sigma} & \epsilon^{ac}{}_{\mu\sigma\lambda} & \zeta^{a[cd]}{}_{\mu\sigma\xi} \\ \hline \mu^{[ab][cd]}{}_{\mu\sigma\lambda} & \mu^{[ab]c}{}_{\mu\sigma} & \epsilon^{[ab][cd]}{}_{\mu\sigma} & \epsilon^{[ab]c}{}_{\mu\sigma\lambda} & \zeta^{[ab][cd]}{}_{\mu\sigma\xi} \\ \hline \mu^{a[cd]}{}_{\mu\nu\sigma\lambda} & \mu^{ac}{}_{\mu\nu\sigma} & \mu^{a[cd]}{}_{\mu\nu\sigma} & \epsilon^{ac}{}_{\mu\nu\sigma\lambda} & \zeta^{a[cd]}{}_{\mu\nu\sigma\xi} \\ \hline \theta^{[ab][cd]}{}_{\mu\nu\sigma\lambda} & \theta^{[ab]c}{}_{\mu\nu\sigma} & \theta^{[ab][cd]}{}_{\mu\nu\sigma} & \theta^{[ab]c}{}_{\mu\nu\sigma\lambda} & \psi^{[ab][cd]}{}_{\mu\nu\sigma\xi} \end{array}\right) \begin{pmatrix} \frac{1}{4}\frac{\delta S_{GR}}{\delta B^{[cd]}{}_{\sigma\lambda}} \\ \frac{\delta S_{GR}}{\delta e^c{}_\sigma} \\ \frac{1}{2}\frac{\delta S_{GR}}{\delta \omega^{[cd]}{}_\sigma} \\ \frac{1}{2}\frac{\delta S_{GR}}{\delta \beta^c{}_{\sigma\lambda}} \\ \frac{1}{4}\frac{\delta S_{GR}}{\delta \lambda^{[cd]}{}_{\sigma\xi}} \end{pmatrix}, \tag{56}$$

where

$$\begin{aligned} \frac{\delta S_{GR}}{\delta B^{[cd]}{}_{\sigma\lambda}} &= \varepsilon^{\mu\nu\sigma\lambda}\left(R_{[cd]\mu\nu} - \lambda_{[cd]\mu\nu}\right), \\ \frac{\delta S_{GR}}{\delta \omega^{[cd]}{}_\sigma} &= \varepsilon^{\sigma\mu\nu\rho}\left(\nabla_\mu B_{[cd]\nu\rho} - e_{[c|\mu}\beta_{|d]\nu\rho}\right), \\ \frac{\delta S_{GR}}{\delta e^c{}_\sigma} &= \frac{1}{2}\varepsilon^{\sigma\mu\nu\rho}\left(\nabla_\mu \beta_{c\nu\rho} + \frac{1}{8\pi l_p^2}\varepsilon_{abcd}\lambda^{[ab]}{}_{\mu\nu} e^d{}_\rho\right), \\ \frac{\delta S_{GR}}{\delta \beta^c{}_{\sigma\lambda}} &= \varepsilon^{\mu\nu\sigma\lambda}\nabla_\mu e_{c\nu}, \\ \frac{\delta S_{GR}}{\delta \lambda^{[cd]}{}_{\sigma\xi}} &= -\varepsilon^{\sigma\xi\mu\nu}\left(B_{[cd]\mu\nu} - \frac{1}{8\pi l_p^2}\varepsilon_{abcd} e^a{}_\mu e^b{}_\nu\right). \end{aligned} \tag{57}$$

We can now investigate the differences in the form of HT transformations for the topological and constrained theory. First, comparing (56) to (50), we see that the HT transformations in the constrained theory feature *more gauge parameters* than are present in the topological theory. Namely, compared to S_{2BF}, the action S_{GR} features an extra Lagrange multiplier two-form $\lambda^{[ab]}$, which extends the matrix of HT parameters from 4×4 blocks to 5×5 blocks, and, therefore, introduces the new parameters ζ and ψ (and θ, which are the negative of ζ due to antisymmetry). This means that the group $\tilde{\mathcal{G}}_{HT}$ for the constrained theory is *larger* than the group $\mathcal{G}_{HT}$ for the topological theory. On the one hand, this immediately proves (54) and, consequently, (55). On the other hand, one can ask the opposite question—given that $\tilde{\mathcal{G}}_{HT}$ is larger than $\mathcal{G}_{HT}$, is the latter maybe a subgroup of the former?

The answer to this question is negative:

$$\mathcal{G}_{HT} \not\subset \tilde{\mathcal{G}}_{HT} , \tag{58}$$

which together with (54) implies our final conclusion:

$$\mathcal{G}_{HT} \neq \tilde{\mathcal{G}}_{HT} . \tag{59}$$

In order to demonstrate (58), we can try to set all extra parameters ζ, ψ, and θ to zero in (56), reducing it to the same form as (50). This would naively suggest that $\mathcal{G}_{HT}$ indeed is a subgroup of $\tilde{\mathcal{G}}_{HT}$. However, upon closer inspection, we can observe that this is not true, since the functional derivatives (57) are different from (51). Namely, even taking into account that the choice $\zeta = \psi = \theta = 0$ eliminates the fifth equation from (57), the first four equations are still different from their counterparts (51) because of the presence of the Lagrange multiplier $\lambda^{[ab]}$ in the action. The Lagrange multiplier is a field in the theory, and generically, it is not zero, since it is determined by the equation of motion:

$$\lambda^{[ab]}{}_{\mu\nu} = R^{[ab]}{}_{\mu\nu} .$$

Therefore, the HT transformations (56) in fact cannot be reduced to the HT transformations (50) by setting the extra parameters equal to zero, which proves (58) and (59).

The overall consequences from the above analysis are as follows. The topological action S_{2BF} has a large ordinary gauge group $\mathcal{G}_{2BF}$ and a small HT symmetry group $\mathcal{G}_{HT}$. When one changes the action to S_{GR} by adding a simplicity constraint term, two things happen—the ordinary gauge group breaks down to its subgroup $\mathcal{G}_{GR}$, so that it becomes smaller, while the HT symmetry group *grows larger* to a completely different group $\tilde{\mathcal{G}}_{HT}$. In effect, the *total* gauge groups for the two actions are intrinsically different:

$$\mathcal{G}^{2BF}_{\text{total}} = \mathcal{G}_{2BF} \ltimes \mathcal{G}_{HT} \qquad \neq \qquad \mathcal{G}^{GR}_{\text{total}} = \mathcal{G}_{GR} \ltimes \tilde{\mathcal{G}}_{HT} \,,$$

in the sense that neither is a subgroup of the other. This conclusion is often overlooked in the literature, which mostly puts emphasis on the symmetry breaking of the ordinary gauge group down to its subgroup.

Let us state here, without proof, that the action (52) represents an example of a non-topological action, for which one can also demonstrate a property analogous to (40), that diffeomorphisms are not a subgroup of its ordinary gauge group, but are a subgroup of the total gauge group. Simply put, given that the simplicity constraint term in (52) breaks the ordinary gauge symmetry group $\mathcal{G}_{2BF}$ into its subgroup $\mathcal{G}_{GR}$ (see (53)), one can expect that diffeomorphisms are not a subgroup of $\mathcal{G}_{GR}$, since they are not a subgroup of the larger group $\mathcal{G}_{2BF}$ of the topological action (46). Nevertheless, since the action (52) is written in a manifestly covariant form, diffeomorphisms are certainly a symmetry of the action and, thus, must be a subgroup of the total gauge group $\mathcal{G}^{GR}_{\text{total}} = \mathcal{G}_{GR} \ltimes \tilde{\mathcal{G}}_{HT}$, in line with the statement analogous to (40). We leave the details of the proof as an exercise for the reader. The point of this analysis was to demonstrate that the interplay (40) between diffeomorphisms and the HT symmetry is a generic property of a large class of actions, including the physically relevant ones, and not limited to examples of topological theories such as the Chern–Simons or nBF models.

As the last comment, let us remark that, in fact, almost all conclusions discussed for the cases of the Chern–Simons, $3BF$, and $2BF$ theories are not really specific to these concrete cases. One can easily generalize our analysis to any other theory, and the conclusions should remain unchanged, except maybe in some corner cases.

5. Conclusions

Let us review the results. In Section 2, we gave a short overview of HT gauge symmetry and discussed its most-important general properties. First, the HT group is a normal subgroup of the total group of gauge symmetries of any given action. Second, HT transformations exhaust all "trivial" (i.e., vanishing on-shell) symmetries, in the sense that there are no trivial symmetries that are not of the HT type. Finally, adding additional terms into the action substantially changes the HT group, often enlarging it. This may be considered a counterintuitive result, since usually adding additional terms in the action serves the purpose of fixing the gauge and, thus, is meant to reduce the gauge symmetry, rather than to enlarge it.

After these general results, in Section 3, we discussed the HT symmetry of the Chern–Simons action, which is a convenient toy example that neatly displays the general features from Section 2. Special attention was given to the issue of diffeomorphisms, and it was shown that, while they are not a subgroup of the ordinary gauge group of the Chern–Simons action, they nevertheless do represent a proper subgroup of the total gauge symmetry, and the HT subgroup plays a nontrivial role in demonstrating this.

Section 4 was devoted to the study of HT symmetry in the $2BF$ and $3BF$ theories, which are relevant for the constructions of realistic quantum gravity models within the generalized spinfoam approach and higher gauge theory. After a brief review and introduction to the notion of three-groups and the $3BF$ theory, appropriate HT transformations were explicitly constructed, complementing the ordinary group of gauge symmetries of the $3BF$ action based on a given three-group. This gave us the total gauge symmetry group for this class of theories. We again discussed the issue of diffeomorphisms and demonstrated again that they are a subgroup of the total gauge group, without being a subgroup of the ordinary gauge group, just like in the case of the Chern–Simons theory. Finally, we introduced a completely concrete example of the $2BF$ theory based on the Poincaré two-group, which becomes classically equivalent to Einstein's general relativity when one introduces the additional term into the action, called the simplicity constraint. As argued in general in Section 2, the presence of this constraint breaks the ordinary gauge group down into its subgroup, while simultaneously enlarging the HT group, since it introduces an additional Lagrange multiplier field into the action. This represents an explicit example of the general statement from Section 2 that the total gauge symmetry group changes nontrivially, as opposed to simply breaking down to its subgroup.

It should be noted that the analysis and results discussed here do not cover everything that can be said about HT symmetry. Among the topics not covered, one can mention the question of an explicit form of finite HT transformations, as opposed to infinitesimal ones. Can one write down finite HT transformations in closed form, either for some conveniently chosen action or maybe even in general? A related topic is the explicit evaluation of the commutator of two HT transformations, or equivalently, the structure constants of the HT Lie algebra, or in yet other words, the multiplication rule in the group $\mathcal{G}_{HT}$. Is the group Abelian or not and for which choices of the action? Finally, one would also like to know the topological properties of the group $\mathcal{G}_{HT}$, i.e., its global structure. All these are potentially interesting topics for future research.

As a particularly interesting topic for future research, we should mention the nontrivial change of the HT symmetry group when additional terms are being added to the action. In Section 4.4, we briefly demonstrated that HT symmetry does change in a nontrivial way, on the example action (52). Nevertheless, the precise properties and the physical interpretation of this change are yet to be studied in full and for a general choice of the action. This topic is the subject of ongoing research.

Finally, we would like to reiterate the differences in two possible approaches to the notion of "the gauge symmetry" of a theory. The overwhelmingly common approach throughout the literature is to factor out the HT group and work only with the ordinary, nontrivial gauge group as the relevant symmetry. Admittedly, this approach does feature a certain level of appeal due to its simplicity and economy, since it does not have to deal with HT symmetry at all. Nevertheless, there are important situations where this is not enough, and one really needs to take into account the *total* gauge symmetry group, which includes HT transformations. As a rule, these situations always involve the gauge symmetry off-shell, either for the purpose of quantization or otherwise. A typical example is the Batalin–Vilkovisky formalism, where one needs to explicitly keep track of HT transformations throughout the whole analysis. Another situation, which was discussed here in more detail, is the question of diffeomorphism symmetry, where HT transformations are required in order to prove that diffeomorphisms are a symmetry of the theory even off-shell. This is especially relevant for building quantum gravity models. Finally, the third scenario would be the discussion of the Coleman–Mandula theorem. One of the main assumptions of the theorem is that the Poincaré group is a subgroup of the full symmetry group of the theory. Given this assumption, and a number of other assumptions, the theorem implies that the full symmetry group must be a direct product of the Poincaré subgroup and the internal symmetry subgroup. In certain cases of theories (such as the $3BF$ action), the full symmetry group is not explicitly expressed as such a direct product, and moreover, it is not obvious that the Poincaré group is a subgroup of the full symmetry group to begin with. Therefore, in order to verify whether the above assumption of the theorem is satisfied, one needs to inspect if the Poincaré group is or is not a subgroup of the full symmetry group. At this point, one may run into a scenario similar to diffeomorphisms: the Poincaré group may fail to be a subgroup of the ordinary gauge group, but still be a subgroup of the total gauge group, once the HT symmetry is taken into account. In this sense, HT symmetry may become relevant for the proper analysis and application of the Coleman–Mandula theorem in certain contexts. This topic is the subject of ongoing research [34].

All of the above arguments suggest that it may be prudent to abandon the common approach of factoring out the HT group and instead adopt the description of the symmetry with the total gauge group, which includes HT transformations on equal footing as the ordinary gauge transformations. In the long run, this may be a conceptually cleaner approach. However, either way, we believe that HT symmetry is relevant for the overall symmetry structure of a theory and that better understanding of its properties can add value to and benefit research.

Author Contributions: Conceptualization, M.V.; investigation, M.Đ., T.R., P.S. and M.V.; writing—original draft preparation, M.Đ., T.R., P.S. and M.V.; writing—review and editing, M.Đ., T.R., P.S. and M.V. All authors have read and agreed to the published version of the manuscript.

Funding: All authors were supported by the Ministry of Science, Technological development and Innovations of the Republic of Serbia. In addition, T.R. and M.V. were supported by the Science Fund of the Republic of Serbia, Grant 7745968, "Quantum Gravity from Higher Gauge Theory 2021"—QGHG-2021. The contents of this publication are the sole responsibility of the authors and can in no way be taken to reflect the views of the Science Fund of the Republic of Serbia.

Data Availability Statement: No new data were created nor analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: The authors would like to thank Igor Prlina for discussions and suggestions when writing this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gomis, J.; Paris, J.; Samuel, S. Antibracket, antifields and gauge-theory quantization. *Phys. Rep.* **1995**, *259*, 1. [CrossRef]
2. Batalin, I.A.; Vilkovisky, G.A. Gauge Algebra and Quantization. *Phys. Lett. B* **1981**, *102*, 27. [CrossRef]
3. Batalin, I.A.; Vilkovisky, G.A. Feynman Rules for Reducible Gauge Theories. *Phys. Lett. B* **1983**, *120*, 166. [CrossRef]
4. Batalin, I.A.; Vilkovisky, G.A. Quantization of Gauge Theories with Linearly Dependent Generators. *Phys. Rev. D* **1983**, *28*, 2567; Erratum in *Phys. Rev. D* **1984**, *30*, 508. [CrossRef]
5. Batalin, I.A.; Vilkovisky, G.A. Closure of the Gauge Algebra, Generalized Lie Algebra Equations and Feynman Rules. *Nucl. Phys.* **1984**, *B234*, 106. [CrossRef]
6. Batalin, I.A.; Vilkovisky, G.A. Existence Theorem for Gauge Algebra. *J. Math. Phys.* **1985**, *26*, 172. [CrossRef]
7. Horowitz, G.T. Exactly Soluble Diffeomorphism Invariant Theories. *Commun. Math. Phys.* **1989**, *125*, 417. [CrossRef]
8. Radenković, T.; Vojinović, M. Gauge symmetry of the 3*BF* theory for a generic semistrict Lie 3-group. *Class. Quant. Grav.* **2022**, *39*, 135009. [CrossRef]
9. Celada, M.; González, D.; Montesinos, M. BF gravity. *Class. Quant. Grav.* **2016**, *33*, 213001. [CrossRef]
10. Girelli, F.; Pfeiffer, H.; Popescu, E.M. Topological Higher Gauge Theory—From BF to BFCG theory. *J. Math. Phys.* **2008**, *49*, 032503. [CrossRef]
11. Martins, J.F.; Miković, A. Lie crossed modules and gauge-invariant actions for 2-BF theories. *Adv. Theor. Math. Phys.* **2011**, *15*, 1059. [CrossRef]
12. Henneaux, M.; Teitelboim, C. *Quantization of Gauge Systems*; Princeton University Press: Princeton, NJ, USA, 1991.
13. Rovelli, C. *Quantum Gravity*; Cambridge University Press: Cambridge, UK, 2004.
14. Rovelli, C.; Vidotto, F. *Covariant Loop Quantum Gravity*; Cambridge University Press: Cambridge, UK, 2014.
15. Baez, J.C.; Huerta, J. An Invitation to Higher Gauge Theory. *Gen. Relativ. Gravit.* **2011**, *43*, 2335. [CrossRef]
16. Miković, A.; Vojinović, M. Poincaré 2-group and quantum gravity. *Class. Quant. Grav.* **2012**, *29*, 165003. [CrossRef]
17. Martins, J.F.; Picken, R. The fundamental Gray 3-groupoid of a smooth manifold and local 3-dimensional holonomy based on a 2-crossed module. *Differ. Geom. Appl. J.* **2011**, *29*, 179. [CrossRef]
18. Wang, W. On 3-gauge transformations, 3-curvatures and Gray-categories. *J. Math. Phys.* **2014**, *55*, 043506. [CrossRef]
19. Saemann, C.; Wolf, M. Six-Dimensional Superconformal Field Theories from Principal 3-Bundles over Twistor Space. *Lett. Math. Phys.* **2014**, *104*, 1147. [CrossRef]
20. Radenković, T.; Vojinović, M. Higher Gauge Theories Based on 3-groups. *J. High Energy Phys.* **2019**, *10*, 222. [CrossRef]
21. Miković, A.; Vojinović, M. Standard Model and 4-groups. *Europhys. Lett.* **2021**, *133*, 61001. [CrossRef]
22. Hidaka, Y.; Nitta, M.; Yokokura, R. Higher-form symmetries and 3-group in axion electrodynamics. *Phys. Lett. B* **2020**, *808*, 135672. [CrossRef]

23. Hidaka, Y.; Nitta, M.; Yokokura, R. Global 3-group symmetry and 't Hooft anomalies in axion electrodynamics. *J. High Energy Phys.* **2021**, *1*, 173. [CrossRef]
24. Song, D.; Wu, K.; Yang, J. 3-form Yang-Mills based on 2-crossed modules. *J. Geom. Phys.* **2022**, *178*, 104537. [CrossRef]
25. Song, D.; Wu, K.; Yang, J. Higher form Yang-Mills as higher BFYM theories. *Eur. Phys. J. C* **2022**, *82*, 1034. [CrossRef]
26. Hidaka, Y.; Nitta, M.; Yokokura, R. Topological axion electrodynamics and 4-group symmetry. *Phys. Lett. B* **2021**, *823*, 136762. [CrossRef]
27. Hidaka, Y.; Nitta, M.; Yokokura, R. Global 4-group symmetry and 't Hooft anomalies in topological axion electrodynamics. *Prog. Theor. Exp. Phys.* **2022**, *2022*, 04A109. [CrossRef]
28. Miković, A.; Vojinović, M. A finiteness bound for the EPRL/FK spin foam model. *Class. Quant. Grav.* **2013**, *30*, 035001. [CrossRef]
29. Baez, J.C. An Introduction to Spin Foam Models of Quantum Gravity and BF Theory. *Lect. Notes Phys.* **2000**, *543*, 25.
30. Baratin, A.; Freidel, L. A 2-categorical state sum model. *J. Math. Phys.* **2015**, *56*, 011705. [CrossRef]
31. Miković, A.; Vojinović, M. Solution to the cosmological constant problem in a Regge quantum gravity model. *Europhys. Lett.* **2015**, *110*, 40008. [CrossRef]
32. Asante, S.K.; Dittrich, B.; Girelli, F.; Riello, A.; Tsimiklis, P. Quantum geometry from higher gauge theory. *Class. Quant. Grav.* **2020**, *37*, 205001. [CrossRef]
33. Radenković, T.; Vojinović, M. Topological invariant of 4-manifolds based on a 3-group. *J. High Energy Phys.* **2022**, *07*, 105. [CrossRef]
34. Đorđević, M.; Vojinović, M. Higher gauge theory and Coleman-Mandula theorem for 3-groups. 2023, *in preparation*.
35. Radenković, T.; Vojinović, M. Hamiltonian Analysis for the Scalar Electrodynamics as 3BF Theory. *Symmetry* **2020**, *12*, 620. [CrossRef]

Метод моделовања резонатора са две блиске фреквенције применом нарушења симетрије

Павле Стипсић
Група за гравитацију, честице и поља
Институт за физику, Универзитет у Београду
Прегревица 118, 11080, Београд, Србија
pstipsic@ipb.ac.rs

***Апстракт*—У овом раду ћемо изложити поступак моделовања полигоналних резонатора са деформисаном C_{nv} симетријом. Резонатори су облика једнакостраничног троугла, квадрата и круга, благо издужених дуж одређеног правца. Нарушење симетрије смањује димензију неких иредуцибилних подпростора са два на један и доводи до раздвајања фреквенција двоструких модова за осциловање у спектру. Метод даје предвиђање резонантних фреквенција и цепање спектра у зависности од параметра деформације.**

***Кључне речи*—резонатор, симетрија, теорија репрезеентација, нарушење симетрије**

I Увод

Резонатори су основни елементи за конструкцију кола у електротехници. Коришћење резонатора са више резонантних учестаности смањује број елемената потребних за реализацију кола и самим тим смањује физичке димензије кола. Једна од могућих реализација резонатора јесте полигон или полигонална линија изведена у проводном слоју са једне стране плочице кола. У секцији II ћемо размотрити утицај симетрије полигона на резонантни спектар, а затим ћемо у секцији III разложити простор стања резонатора на иредуцибилне компоненте. У секцији IV ћемо наметнути граничне услове на модове осциловања у појединачним иредуцибилним подпросторима, а у секцијама V и VI ћемо редом моделовати утицај термалних губитака и нарушења симетрије на резонантни спектар. У целом раду је подразумеван природни систем јединица, што значи да су све физичке величионе изражене у метрима на одговарајући степен. Метрика у простору Минковског је са сигнатуром $\eta_{\mu\nu} = \mathrm{diag}(+1, -1, -1, -1)$.

II Утицај симетрије на спектар резонатора

Посматраћемо резонаторе облика правилног n-тоугла. Група симетрије коју поседује правилан n-тоугао је C_{nv}, односно операције које чувају симетрију су ротације у равни за целобројни умножак угла од $\frac{2\pi}{n}$ радијана, што чини C_n подгрупу, као и рефлексије у односу на свих n оса симетрије. Групу рефлексије у односу на вертикалну осу симетрије означавамо са σ_v, а цела група симетрије онда има структуру

$$C_{nv} = C_n \rtimes \sigma_v \,. \tag{1}$$

Група C_n је инваријантна подгрупа укупне групе симетрије C_{nv} и коњугација елемента c_k групе C_n рефлексијом σ_v даје инверзни елемент c_{-k}. То значи да C_{nv} група није Абелова и ако C_n и σ_v јесу:

$$\sigma c_k \sigma^{-1} = \sigma c_k \sigma = (\sigma c_1 \sigma)^k = (c_{-1})^k = c_{-k} \,. \tag{2}$$

Простор стања резонатора је Хилбертов простор функција које описују расподелу електромагнетног поља унутар резонатора. Групу симетрије репрезентујемо скупом оператора у том простору и тај скуп називамо репрезентацијом. Репрезентацију сваког елемента g групе G означавамо са $D(g)$. Избор репрезентације групе је у великој мери произвољан докле год задовољава правило множења $D(gg') = D(g)D(g')$, али се испоставља да се у случају наше групе симетрије, као и велоког броја група од значаја за физику, свака репрезентација може разложити по скупу иредуцибилних репрезентација. Иредуцибилне репрезентације групе представљају репрезентације групе у иредуцибилним подпросторима укупног простора стања, при чему је дефиниција иредуцибилног подпростора та да је то подпростор у ком се сваки вектор из подпростора деловањем свих елемената групе засебно пресликава у скуп вектора који разапињу цео подпростор. Другим речима, унутар иредуцибилног подпростора не постоји вектор помоћу ког се деловањем свих елемената групе не може реконструисати цео иредуцибилни подпростор. Разлагање репрезентације на иредуцибилне компоненте је онда ортогонални збир репрезентација по иредуцибилним подпросторима дат са:

$$D(G) = \bigoplus_{\mu=1}^{r} f_\mu D^{(\mu)}(G) \,, \tag{3}$$

где је f_μ фреквенција појављивања μ-те иредуцибилне репрезентације у разлагању, а r укупан број иредуцибилних репрезентација.

Иредуцибилне репрезентације поседују скуп особина које ће нам бити од значаја приликом разматрања како симетријских особина резонатора, тако и приликом

ISBN 978-86-6200-032-3 ©2025 AM & ETRAN
https://doi.org/10.69994/69E25021

одређивања разлагања простора стања на његове иредуцибилне компоненте. Неке од значајних особина су дате следећим исказима:

- Прва Шурова лема: Оператор који комутира са свим елементима групе у једној иредуцибилној репрезентацији делује као скаларни оператор у тој репрезентацији. То значи да је оператор у базису иредуцибилних репрезентација дијагоналан и да су му све вредности на дијагонали исте унутар појединачних иредуцибилних подпростора.

- Друга Шурова лема: Једини оператор A за који једнакост $AD^{(\mu)}(g) = D^{(\nu)}(g)A$ важи за сваки елемент g групе G, при чему су $D^{(\mu)}$ и $D^{(\nu)}$ нееквивалентне иредуцибилне репрезентације, тј. једна није добијена од друге променом базиса, је оператор $A = 0$.

- Теорема о ортогоналности компонената иредуцибилних репрезентација:
$$\frac{1}{|G|}\sum_{g\in G} D^{(\mu)}_{ij}(g^{-1})D^{(\nu)}_{kl}(g) = \frac{1}{|\nu|}A_{il}\left(A^{-1}\right)_{kj}\delta^{\mu\nu}\,, \tag{4}$$
где је $|G|$ број елемената групе G, $|\nu|$ димензија ν-тог иредуцибилног подпростора, а A оператор промене базиса између μ-те и ν-те репрезентације у случају када су оне међусобно еквивалентне.

- Карактер репрезентације $\chi(G)$, ондносно траг матричне репрезентације, као функција на групи, такође задовољава сличну релацију ортогоналности:
$$\frac{1}{|G|}\sum_{g\in G}\chi^{(\mu)}(g^{-1})\chi^{(\nu)}(g) = \delta^{\mu\nu}\,. \tag{5}$$
Ова релација нам омогућава да одредимо коефицијенте у разлагању репрезентације на нееквивалентне иредуцибилне компоненте као:
$$f^{\mu} = \frac{1}{|G|}\sum_{g\in G}\chi^{(\mu)}(g^{-1})\chi(g)\,. \tag{6}$$

- Број иредуцибилних репрезентација групе је једнак броју класа коњугације у групи.

- Може се увести појам регуларне репрезентације чији је карактер функција на групи која јединичном елементу додељује број елемената групе, а свим осталим елементима нулу. Ова функција је аналогон Дираквој односно Кронекеровој делти над елементима групе. На основу претходне особине се добија да је фреквенција појављивања сваке нееквивалентне иредуцибилне репрезентације у разлагању регуларне репрезентације једнак димензији иредуцибилне репрезентације.

- Број елемената групе једнак је збиру квадрата димензија свих нееквивалентних иредуцибилних репрезентација групе.

На основу прве Шурове леме можемо да закључимо да су све иредуцибилне репрезентације Абелових група једнодимензионалне. Такође, како се оператор фреквенције модова за осциловање не трансформише приликом дејства елемената групе симетрије, закључујемо да он комутира са репрезентацијама свих елемената групе, па су му поново, према првој Шуровој леми, својствене вредности дегенерисане унутар иредуцибилних подпростора. То значи да ће сви модови за осциловање унутар једног иредуцибилног подпростора бити са истом фреквенцијом, а дегенерација фреквенције је једнака димензији иредуцибилног подпростора.

У даљем раду ћемо из скупа нееквивалентних иредуцибилних репрезентација одабрати унитарне представнике, па ће ротације у равни бити репрезентоване множењем комплексним бројем са јединичне кружнице, $e^{i\mu\frac{2\pi}{n}}$, а рефлексија ће бити репрезентована комплексном коњугацијом. Бројач репрезентације μ узима све вредности природних бројева од 1 до n.

Поступак добијања скупа нееквивалентних иредуцибилних репрезентација произвољне групе не постоји. Ипак, у случају C_{nv} групе постоји алгоритам индукције са инваријантне C_n подгрупе индекса 2, који даје комплетан скуп иредуцибилних репрезентација које се могу пописати табеларно:

ТАБЕЛА I Иредуцибилне репрезентације C_{nv} група

$D^{(\mu)}(C_{nv})$	μ	c_k	σc_k
$A_0/B_0 = D^{(0\,\pm)}(C_{nv})$	0	1	± 1
$A_{n/2}/B_{n/2} = D^{(n/2\,\pm)}(C_{nv})$	$\frac{n}{2}$	$(-1)^k$	$\pm(-1)^k$
$E_\mu = D^{(\mu)}(C_{nv})$	$1\ldots\lfloor\frac{n-1}{2}\rfloor$	$\begin{pmatrix} e^{i\mu\frac{2\pi}{n}k} & 0 \\ 0 & e^{-i\mu\frac{2\pi}{n}k}\end{pmatrix}$	$\begin{pmatrix} 0 & e^{-i\mu\frac{2\pi}{n}k} \\ e^{i\mu\frac{2\pi}{n}k} & 0\end{pmatrix}$

Очекивано, како је инваријантна подгрупа Абелова и индекса два, димензије иредуцибилних репрезентација не могу бити веће од два. На основу табеле такође закључујемо да симетрија обезбеђује постојање парова модова за осциловање са истом фреквенцијом. Детаљан преглед особина иредуцибилних репрезентација се може пронаћи у [1].

III Разлагање координатне репрезентације

Скуп једначина стања резонатора чине диференцијална једначина кретања електромагнетног поља у резонатору и једначина за проводност средине σ:
$$\Box A^{\mu} - \partial^{\mu}\partial_{\nu}A^{\nu} = j^{\mu}\,, \qquad j^{i} = \sigma^{i}{}_{j}\left(\partial^{j}A^{0} - \partial^{0}A^{j}\right)\,. \tag{7}$$

Прва једначина се може појеодноставити фиксирањем Лоренцовог калибрационог услова за електромагнетни потенцијал A^μ, $\partial_\mu A^\mu = 0$, док за струју j^μ већ важи идентична релација која описује једначину континуитета $\partial_\mu j^\mu = 0$. У свим једначинама подразумевамо Ајнштајнову сумациону конвенцију по поновљеним индексима, као и да су малим грчким словима означене све просторвременске компоненте поља, а малим латинским само просторне компоненте поља. Нулта координата је резервисана за временску компоненту.

Обе једначине су записане у координатној репрезентацији у координатном базису који није својствен за диференцијалне операторе у њима, па је први корак свакако прелазак у базис равних таласа, импулсни базис, који је својствен овим једначинама, а потом је потребно конструисати модификоване групне пројекторе

$$P(D^{(\mu)*} \otimes D) = \frac{1}{|G|} \sum_{g \in G} D^{(\mu)}(g^{-1}) \otimes D(g)\,, \tag{8}$$

који ће комплетан простор стања у својственом базису пројектовати на вишеструки збир иредуцибилних подпростора групе симетрије, такође у својственом базису. Вектори фиксне тачке овог пројектора су облика:

$$|\mu\, t_\mu\rangle = \sum_{i=1}^{|\mu|} |\mu\, i\rangle \otimes |\mu\, t_\mu\, i\rangle\,, \tag{9}$$

где индекс t_μ пребројава појављивање μ-те иредуцибилне репрезентације у разлагању. Тако ћемо добити базис C_{nv}-симетричних таласа. Како је простор стања несепарабилан, очекујемо да се бар једна иредуцибилна репрезентација у разлагању појављује бесконачан број пута и да се њено појављивање пребројава реалним параметром. Након разлагања репрезентације ћемо видети да су појављивања свих иредуцибилних компонената пребројана реалним параметрима који ће имати смисао квадрадта импулса електромагнетног таласа. Овај базис је у случају једнодимензионалних репрезентација дат са:

$$\begin{aligned}
f^{A_0}_{k_\mu}(x^\mu) &= \frac{1}{n} e^{i(k_z z - \omega t)} \sum_{m=1}^{n} e^{ik_x\left(x\cos\left(\frac{2\pi}{n}m\right) - y\sin\left(\frac{2\pi}{n}m\right)\right)} \\
&\quad \times \cos\left(k_y\left(y\cos\left(\tfrac{2\pi}{n}m\right) + x\sin\left(\tfrac{2\pi}{n}m\right)\right)\right), \\
f^{B_0}_{k_\mu}(x^\mu) &= \frac{1}{n} e^{i(k_z z - \omega t)} \sum_{m=1}^{n} e^{ik_x\left(x\cos\left(\frac{2\pi}{n}m\right) - y\sin\left(\frac{2\pi}{n}m\right)\right)} \\
&\quad \times \sin\left(k_y\left(y\cos\left(\tfrac{2\pi}{n}m\right) + x\sin\left(\tfrac{2\pi}{n}m\right)\right)\right), \\
f^{A_{\frac{n}{2}}}_{k_\mu}(x^\mu) &= \frac{1}{n} e^{i(k_z z - \omega t)} \sum_{m=1}^{n} e^{ik_x\left(x\cos\left(\frac{2\pi}{n}m\right) - y\sin\left(\frac{2\pi}{n}m\right)\right)} \\
&\quad \times \cos\left(k_y\left(y\cos\left(\tfrac{2\pi}{n}m\right) + x\sin\left(\tfrac{2\pi}{n}m\right)\right)\right) e^{-im\pi}, \\
f^{B_{\frac{n}{2}}}_{k_\mu}(x^\mu) &= \frac{1}{n} e^{i(k_z z - \omega t)} \sum_{m=1}^{n} e^{ik_x\left(x\cos\left(\frac{2\pi}{n}m\right) - y\sin\left(\frac{2\pi}{n}m\right)\right)} \\
&\quad \times \sin\left(k_y\left(y\cos\left(\tfrac{2\pi}{n}m\right) + x\sin\left(\tfrac{2\pi}{n}m\right)\right)\right) e^{-im\pi},
\end{aligned} \tag{10}$$

док је у случају дводимензионалних репрезентација:

$$\begin{aligned}
f^{E_j}_{k^\mu}(x^\mu) &= \frac{1}{n} e^{i(\omega t - k_z z))} \sum_{m=1}^{n} e^{-ik_x\left(x\cos\left(\frac{2\pi}{n}m\right) - y\sin\left(\frac{2\pi}{n}m\right)\right)} \\
&\quad \left(\cos\left(k_y\left(y\cos\left(\tfrac{2\pi}{n}m\right) + x\sin\left(\tfrac{2\pi}{n}m\right)\right) - \tfrac{2\pi}{n}mj\right)\right. \\
&\quad \left.\oplus \cos\left(k_y\left(y\cos\left(\tfrac{2\pi}{n}m\right) + x\sin\left(\tfrac{2\pi}{n}m\right)\right) + \tfrac{2\pi}{n}mj\right)\right).
\end{aligned} \tag{11}$$

У следећој секцији ћемо наметањем граничних услова из овог несепарабилног простора издвојити сепарабилни подпростор који ће бити дефинисан дискретним скупом дозвољених вредности индекса k_μ. Сепарабилност простора стања обезбеђује дискретну природу спектра резонатних фреквенција.

IV Модови за осциловање идеалног резонатора

Да бисмо из укупног несепарабилног простора стања издвојили сепарабилни подпростор који одговара модовима за осциловање, потребно је да наметнемо одређене услове на сама стања. Ови услови су заправо гранични услови који су садржани у тензору проводности:

$$\begin{aligned}
\sigma^i{}_j &= \left(\tilde\delta(z+d) + \tilde\delta(z)\Pi(x,y)\right)\frac{1}{\rho}\left(\delta^i_x\delta^x_j + \delta^i_y\delta^y_j\right) \\
&\quad + \theta(z+d)\theta(-z)\sigma\delta^i_j\,.
\end{aligned} \tag{12}$$

Функција $\Pi(x,y)$ је прозорска функција полигоналног облика која има вредност 1 на домену који покрива проводни слој резонатора, а вредност 0 иначе, док је $\tilde\delta(z)$ јединични прозор дебљине проводног слоја. Функција $\theta(z)$ је стандардна Хевисајдова тета функсија.

Први гранични услов који намећемо је по z координати, па параметар базиса k_z узима целобаројне умношке од $\frac{2\pi}{d}$, где је d дебљина резонатора. Прва апроксимација коју ћемо направити је та да ћемо сматрати да је резонатор превише танак да би се побудили модови за осциловање у z правцу, па ћемо се ограничити само на подпростор $k_z = 0$.

Наметање граничних услова у равни резонатора се не може увек тако лако извршити. Наиме, како је базис у простору стања састављен од периодичних функција, сви наметнути гранични услови ће се периодично понављати. То значи да се на појединачни базисни вектор могу наметнути гранични услови само у случају оних полигоналних облика којим се раван може поплочати без празнина, а то су троуглови, квадрати и шестоуглови. Додатно, како шестоугаона решетка има елементарну Вигнер-Зајцову ћелију састављену од два типа чворова, ово наметање граничних услова неће бити могуће у свим иредуцибилним репрезентацијама, већ само у оним у којима су ова два чвора еквивалентна. У осталим случајевима се гранични услови намећу линеарним комбинацијама вектора вишеструких иредуцибилних компонената, што је случај са кружним резонатором код којег је базис у поларним координатама дат преко цилиндричних Беселових функција прве врсте. Ове функције су, иако осцилаторне, непериодичне, па ће такве бити и таласне функције са симетријом C_{nv} за $n > 4$. Због тога ћемо се у даљем раду фокусирати специјално на троугаони, квадратни и кружни резонатор.

У случају троугаоног резонатора, висине троугла једнаке h, наметање граничних услова даје по два базисна вектора из сваке једнодимензионалне иредуцибилне репрезентације:

$$
\begin{aligned}
f^{A_{0,s}}_{k^{1,2}}(x^{1,2}) &= \tfrac{1}{3}\Big(\sin\Big(x\tfrac{(2p+3q)\pi}{h}\Big)\cos\Big(y\tfrac{q\sqrt{3}\pi}{h}\Big)\\
&- \sin\big(x\tfrac{p\pi}{h}\big)\cos\Big(y\tfrac{(p+2q)\sqrt{3}\pi}{h}\Big)\\
&- \sin\Big(x\tfrac{(p+3q)\pi}{h}\Big)\cos\Big(y\tfrac{(p+q)\sqrt{3}\pi}{h}\Big)\Big),\\
f^{A_{0,c}}_{k^{1,2}}(x^{1,2}) &= \tfrac{1}{3}\Big(\cos\Big(x\tfrac{(2p+3q)\pi}{h}\Big)\cos\Big(y\tfrac{q\sqrt{3}\pi}{h}\Big)\\
&+ \cos\big(x\tfrac{p\pi}{h}\big)\cos\Big(y\tfrac{(p+2q)\sqrt{3}\pi}{h}\Big)\\
&+ \cos\Big(x\tfrac{(p+3q)\pi}{h}\Big)\cos\Big(y\tfrac{(p+q)\sqrt{3}\pi}{h}\Big)\Big),\\
f^{B_{0,s}}_{k^{1,2}}(x^{1,2}) &= \tfrac{1}{3}\Big(\sin\Big(x\tfrac{(2p+3q)\pi}{h}\Big)\sin\Big(y\tfrac{q\sqrt{3}\pi}{h}\Big)\\
&+ \sin\big(x\tfrac{p\pi}{h}\big)\sin\Big(y\tfrac{(p+2q)\sqrt{3}\pi}{h}\Big)\\
&- \sin\Big(x\tfrac{(p+3q)\pi}{h}\Big)\sin\Big(y\tfrac{(p+q)\sqrt{3}\pi}{h}\Big)\Big),\\
f^{B_{0,c}}_{k^{1,2}}(x^{1,2}) &= \tfrac{1}{3}\Big(\cos\Big(x\tfrac{(2p+3q)\pi}{h}\Big)\sin\Big(y\tfrac{q\sqrt{3}\pi}{h}\Big)\\
&- \cos\big(x\tfrac{p\pi}{h}\big)\sin\Big(y\tfrac{(p+2q)\sqrt{3}\pi}{h}\Big)\\
&+ \cos\Big(x\tfrac{(p+3q)\pi}{h}\Big)\sin\Big(y\tfrac{(p+q)\sqrt{3}\pi}{h}\Big)\Big),
\end{aligned}
\tag{13}
$$

док се у случају дводимензионалне репрезентације услови могу наметнути на два начина за сваки вектор из репрезентације. У случају E_1^+ репрезентације се добија:

$$
\begin{aligned}
f^{E_1^+,s_1}_{k^{1,2}}(x^{1,2}) &= \tfrac{1}{3}\Big(\sin\Big(x\tfrac{(2p+3q+1)\pi}{h}\Big)\cos\Big(y\tfrac{(3q+1)\pi}{h\sqrt{3}}\Big)\\
&- \sin\Big(x\tfrac{(p+3q+1)\pi}{h}\Big)\cos\Big(y\tfrac{(3p+3q+1)\pi}{h\sqrt{3}}\Big)\\
&- \sin\big(x\tfrac{p\pi}{h}\big)\cos\Big(y\tfrac{(3p+6q+2)\pi}{h\sqrt{3}}\Big)\Big),\\
f^{E_1^+,c_1}_{k^{1,2}}(x^{1,2}) &= \tfrac{1}{3}\Big(\cos\Big(x\tfrac{(2p+3q+1)\pi}{h}\Big)\cos\Big(y\tfrac{(3q+1)\pi}{h\sqrt{3}}\Big)\\
&+ \cos\Big(x\tfrac{(p+3q+1)\pi}{h}\Big)\cos\Big(y\tfrac{(3p+3q+1)\pi}{h\sqrt{3}}\Big)\\
&+ \cos\big(x\tfrac{p\pi}{h}\big)\cos\Big(y\tfrac{(3p+6q+2)\pi}{h\sqrt{3}}\Big)\Big),\\
f^{E_1^+,s_2}_{k^{1,2}}(x^{1,2}) &= \tfrac{1}{3}\Big(\sin\Big(x\tfrac{(2p+3q+2)\pi}{h}\Big)\cos\Big(y\tfrac{(3q+2)\pi}{h\sqrt{3}}\Big)\\
&- \sin\Big(x\tfrac{(p+3q+2)\pi}{h}\Big)\cos\Big(y\tfrac{(3p+3q+2)\pi}{h\sqrt{3}}\Big)\\
&- \sin\big(x\tfrac{p\pi}{h}\big)\cos\Big(y\tfrac{(3p+6q+4)\pi}{h\sqrt{3}}\Big)\Big),\\
f^{E_1^+,c_2}_{k^{1,2}}(x^{1,2}) &= \tfrac{1}{3}\Big(\cos\Big(x\tfrac{(2p+3q+2)\pi}{h}\Big)\cos\Big(y\tfrac{(3q+2)\pi}{h\sqrt{3}}\Big)\\
&+ \cos\Big(x\tfrac{(p+3q+2)\pi}{h}\Big)\cos\Big(y\tfrac{(3p+3q+2)\pi}{h\sqrt{3}}\Big)\\
&+ \cos\big(x\tfrac{p\pi}{h}\big)\cos\Big(y\tfrac{(3p+6q+4)\pi}{h\sqrt{3}}\Big)\Big),
\end{aligned}
\tag{14}
$$

а у случају E_1^-:

$$
\begin{aligned}
f^{E_1^-,s_1}_{k^{1,2}}(x^{1,2}) &= \tfrac{1}{3}\Big(\sin\Big(x\tfrac{(2p+3q+1)\pi}{h}\Big)\sin\Big(y\tfrac{(3q+1)\pi}{h\sqrt{3}}\Big) \qquad (15)\\
&- \sin\Big(x\tfrac{(p+3q+1)\pi}{h}\Big)\sin\Big(y\tfrac{(3p+3q+1)\pi}{h\sqrt{3}}\Big)\\
&+ \sin\big(x\tfrac{p\pi}{h}\big)\sin\Big(y\tfrac{(3p+6q+2)\pi}{h\sqrt{3}}\Big)\Big),\\
f^{E_1^-,c_1}_{k^{1,2}}(x^{1,2}) &= \tfrac{1}{3}\Big(\cos\Big(x\tfrac{(2p+3q+1)\pi}{h}\Big)\sin\Big(y\tfrac{(3q+1)\pi}{h\sqrt{3}}\Big)\\
&+ \cos\Big(x\tfrac{(p+3q+1)\pi}{h}\Big)\sin\Big(y\tfrac{(3p+3q+1)\pi}{h\sqrt{3}}\Big)\\
&- \cos\big(x\tfrac{p\pi}{h}\big)\sin\Big(y\tfrac{(3p+6q+2)\pi}{h\sqrt{3}}\Big)\Big),\\
f^{E_1^-,s_2}_{k^{1,2}}(x^{1,2}) &= \tfrac{1}{3}\Big(\sin\Big(x\tfrac{(2p+3q+2)\pi}{h}\Big)\sin\Big(y\tfrac{(3q+2)\pi}{h\sqrt{3}}\Big)\\
&- \sin\Big(x\tfrac{(p+3q+2)\pi}{h}\Big)\sin\Big(y\tfrac{(3p+3q+2)\pi}{h\sqrt{3}}\Big)\\
&+ \sin\big(x\tfrac{p\pi}{h}\big)\sin\Big(y\tfrac{(3p+6q+4)\pi}{h\sqrt{3}}\Big)\Big),\\
f^{E_1^-,c_2}_{k^{1,2}}(x^{1,2}) &= \tfrac{1}{3}\Big(\cos\Big(x\tfrac{(2p+3q+2)\pi}{h}\Big)\sin\Big(y\tfrac{(3q+2)\pi}{h\sqrt{3}}\Big)\\
&+ \cos\Big(x\tfrac{(p+3q+2)\pi}{h}\Big)\sin\Big(y\tfrac{(3p+3q+2)\pi}{h\sqrt{3}}\Big)\\
&- \cos\big(x\tfrac{p\pi}{h}\big)\sin\Big(y\tfrac{(3p+6q+4)\pi}{h\sqrt{3}}\Big)\Big). \qquad (16)
\end{aligned}
$$

Бројеви p и q су природни бројеви или нула.

У случају C_{4v} симетрије се на сличан начин добијају базисни вектори. У случају једнодимензионалних репрезентација вектори су једнаки:

$$
\begin{aligned}
f^{A_{0,s}}_{k^{1,2}}(x^{1,2}) &= \tfrac{1}{2}\Big(\cos\Big(x\tfrac{(2q+1)\pi}{a}\Big)\cos\Big(y\tfrac{(2(p+q)+1)\pi}{a}\Big)\\
&+ \cos\Big(x\tfrac{(2(p+q)+1)\pi}{a}\Big)\cos\Big(y\tfrac{(2q+1)\pi}{a}\Big)\Big),\\
f^{A_{0,c}}_{k^{1,2}}(x^{1,2}) &= \tfrac{1}{2}\Big(\cos\big(x\tfrac{2q\pi}{a}\big)\cos\Big(y\tfrac{2(p+q)\pi}{a}\Big)\\
&+ \cos\Big(x\tfrac{2(p+q)\pi}{a}\Big)\cos\big(y\tfrac{2q\pi}{a}\big)\Big),\\
f^{B_{0,s}}_{k^{1,2}}(x^{1,2}) &= \tfrac{1}{2}\Big(\sin\big(x\tfrac{2q\pi}{a}\big)\sin\Big(y\tfrac{2(p+q)\pi}{a}\Big)\\
&- \sin\Big(x\tfrac{2(p+q)\pi}{a}\Big)\sin\big(y\tfrac{2q\pi}{a}\big)\Big),\\
f^{B_{0,c}}_{k^{1,2}}(x^{1,2}) &= \tfrac{1}{2}\Big(\sin\Big(x\tfrac{(2q+1)\pi}{a}\Big)\sin\Big(y\tfrac{(2(p+q)+1)\pi}{a}\Big)\\
&- \sin\Big(x\tfrac{(2(p+q)+1)\pi}{a}\Big)\sin\Big(y\tfrac{(2q+1)\pi}{a}\Big)\Big),\\
f^{A_{2,s}}_{k^{1,2}}(x^{1,2}) &= \tfrac{1}{2}\Big(\cos\Big(x\tfrac{(2q+1)\pi}{a}\Big)\cos\Big(y\tfrac{(2(p+q)+1)\pi}{a}\Big)\\
&- \cos\Big(x\tfrac{(2(p+q)+1)\pi}{a}\Big)\cos\Big(y\tfrac{(2q+1)\pi}{a}\Big)\Big),\\
f^{A_{2,c}}_{k^{1,2}}(x^{1,2}) &= \tfrac{1}{2}\Big(\cos\big(x\tfrac{2q\pi}{a}\big)\cos\Big(y\tfrac{2(p+q)\pi}{a}\Big)\\
&- \cos\Big(x\tfrac{2(p+q)\pi}{a}\Big)\cos\big(y\tfrac{2q\pi}{a}\big)\Big),\\
f^{B_{2,s}}_{k^{1,2}}(x^{1,2}) &= \tfrac{1}{2}\Big(\sin\big(x\tfrac{2q\pi}{a}\big)\sin\Big(y\tfrac{2(p+q)\pi}{a}\Big)\\
&+ \sin\Big(x\tfrac{2(p+q)\pi}{a}\Big)\sin\big(y\tfrac{2q\pi}{a}\big)\Big),\\
f^{B_{2,c}}_{k^{1,2}}(x^{1,2}) &= \tfrac{1}{2}\Big(\sin\Big(x\tfrac{(2q+1)\pi}{a}\Big)\sin\Big(y\tfrac{(2(p+q)+1)\pi}{a}\Big)\\
&+ \sin\Big(x\tfrac{(2(p+q)+1)\pi}{a}\Big)\sin\Big(y\tfrac{(2q+1)\pi}{a}\Big)\Big),
\end{aligned}
\tag{17}
$$

а у случају дводимензионалне репрезентације једнаки:

$$\begin{aligned} f_{k^{1,2}}^{E_1^{\pm},s}(x^{1,2}) &= \tfrac{1}{2}\left(\sin\left(x\tfrac{2q\pi}{a}\right)\cos\left(y\tfrac{(2(p+q)+1)\pi}{a}\right)\right. \\ &\pm \left. i\cos\left(x\tfrac{(2(p+q)+1)\pi}{a}\right)\sin\left(y\tfrac{2q\pi}{a}\right)\right), \\ f_{k^{1,2}}^{E_1^{\pm},c}(x^{1,2}) &= \tfrac{1}{2}\left(\sin\left(x\tfrac{(2q+1)\pi}{a}\right)\cos\left(y\tfrac{2(p+q)\pi}{a}\right)\right. \\ &\pm \left. i\cos\left(x\tfrac{2(p+q)\pi}{a}\right)\sin\left(y\tfrac{(2q+1)\pi}{a}\right)\right) \quad (18) \end{aligned}$$

Бројеви p и q и овде узимају вредности из скупа природних бројева и нуле, а a је дужина странице квадрата. За више детаља о наметању граничних услова погледати [2].

Ови вектори описују како компоненте електромагнетног поља, тако и компоненте струје. За анализу модова за осциловање ћемо се фокусирати на опис компонената струје. Струју ћемо поделити на две компоненте, паралелну с ивицом полигона и нормалну на ивицу полигона. Паралелна компонента струје може да се разлаже и у s и у c базису, док нормална компонента постоји само у s базису. У специјалном случају када је резонатор реализован у облику полигоналне линије, ортогонална компонента струје се може занемарити, а паралелна се редукује само на c базис.

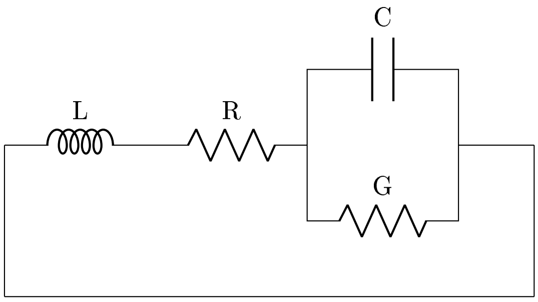
V Утицај губитака на спектар резонатора

У колико бисмо убацили базисне векторе добијене у претходном поглављу у хомогену једначину кретања за електромагнетно поље, добили бисмо услов на фреквенцију слободног електормагнетног таласа и ове фреквенције су једнаке:

$$\begin{aligned} \left(\omega_0^{C_{3v}}\right)^2 &= \frac{4\pi^2}{3h^2}\left(q^2+3pq+3p^2\right), \\ \left(\omega_0^{C_{4v}}\right)^2 &= \frac{\pi^2}{a^2}\left(\xi^2+\psi^2\right), \quad (19) \end{aligned}$$

где је $\xi = 2q$ или $\xi = 2q+1$ и $\psi = 2(p+q)$ или $\psi = 2(p+q)+1$ у зависности од репрезентације.

У присуству губитака услед отпорности проводних слојева и проводности материјала од којег је плочица израђена, резонантне фреквенције ће бити снижене. Сам резонатор ћемо апроксимирати колом са слике 1.



Слика 1: Апроксимативно коло за моделовање губитака у резонатору.

Апроксимативне вредности елемената кола су:

$$C = \frac{S}{d}, \quad G = \sigma C, \quad R = \frac{\rho}{\omega_0^2 S}, \quad L = \frac{d}{\omega_0^2 S}. \quad (20)$$

Капацитет C и проводност G зависе искључиво од геометрије резонатора и специфичне проводности диелектрика σ од ког је плочица састављена и расту с повећањем површине резонатора S. Отпорност R зависи како од геометрије резонатора и отпорности проводног слоја ρ, тако и од просторне расподеле струја, односно таласне дужине мода за осциловање. Индуктивност L такође зависи од геометрије резонатора и таласне дужине мода за осциловање и расте с повећањем дебљине диелектрика d. Анализом кола се добија да је нова резонантна фреквенција са урачунатим губицима:

$$\omega = \sqrt{\omega_0^2 - \frac{1}{2}\left(\sigma^2 + \frac{\rho^2}{d^2}\right)}. \quad (21)$$

Резонантна фреквенција је имагинарни део пола пропагатора функција стања у својственом базису тангентног простора, у овом случају простора тангентног на време.

VI Нарушење симетрије резонатора

Да бисмо контролисано уклонили дегенерацију фреквенције унутар дводимензионалних иредуцибилних подпростора деформисаћемо резонатор дуж правца под произвољним углом ϕ у равни резонатора. Постоје два главна приступа решавању једначина за деформисани резонатор. Први је да се промене гранични услови тако да одговарају деформисаном резонатору у старим координатама, а други је да се старе координате рескалирају тако да се гранични услови деформисаног резонатора поклопе са граничним условима из базиса, али да се ефективно промени једначина стања резонатора. Овај други приступ је много једноставнији за рад, будући да смо целу процедуру конструисали тако да у њу уградимо граничне услове, а додатни сабирци у једначини стања осцилатора се могу интерпретирати као мале пертурбације по параметру деформације. У ту сврху ћемо увести оператор рескалирања координата као:

$$\begin{aligned} \partial_i' &= D_\phi(\varepsilon)_i{}^j\partial_j = \left(R(\phi)D_0(\varepsilon)R(-\phi)\right)_i{}^j\partial_j \\ &= \begin{pmatrix} 1+\varepsilon\cos^2(\phi) & \varepsilon\sin(\phi)\cos(\phi) \\ \varepsilon\sin(\phi)\cos(\phi) & 1+\varepsilon\sin^2(\phi) \end{pmatrix}_i^{\,j}\partial_j. \quad (22) \end{aligned}$$

Параметар ε је параметар рескалирања координата који ефективно деформише резонатор. Нова једначина стања резонатора се разликује од старе јер сада Лапласијан у равни резонатора постаје:

$$\partial_i'\partial'^i = \partial_i\partial^i - \varepsilon(2+\varepsilon)\left(\cos(\phi)\partial_x + \sin(\phi)\partial_y\right)^2. \quad (23)$$

Додатни сабирак $\varepsilon(2+\varepsilon)\left(\cos(\phi)\partial_x + \sin(\phi)\partial_y\right)^2$ разбија дегенерацију унутар иредуцибилних подпростора, па се прве поправке фреквенција могу добити дијагонализацијом његове матричне репрезентације у иредуцибилном подпростору. Ако својствене вредности поправке у иредуцибилном подпростору означимо са $\Delta^i_{\mu,t_\mu}(\varepsilon,\phi)$, где је μ бројач репрезентације, t_μ бројач појављивања

μ-те репрезентације и i пребројава својствене вредности поправке унутар (μ, t_μ)-тог иредуцибилног подпростора, фреквенције модова постају

$$\omega = \sqrt{\omega_0^2 - \Delta^i_{\mu,t_\mu}(\varepsilon,\phi) - \frac{1}{2}\left(\sigma^2 + \frac{\rho^2}{d^2}\right)}\,. \qquad (24)$$

Када пројектујемо резонатор, димензије прилагођавамо тако да пропусни опсег одговара фреквенцијама из најниже дводимензионалне иредуцибилне репрезентације, а да остале фреквенције буду потиснуте. Због тога ћемо се фокусирати само на први дводимензионални иредуцибилни подпростор. У случају троугаоног резонатора ова репрезентација одговара паралелној компоненти струје у c базису $E_1^{\pm,1}$ и поправка је једнака:

$$\Delta_{0,1} = \frac{\varepsilon(2+\varepsilon)}{h^2}\begin{pmatrix} \frac{32\pi^2+81\cos(2\phi)}{48} & -\frac{27}{16}\sin(2\phi) \\ -\frac{27}{16}\sin(2\phi) & \frac{32\pi^2-81\cos(2\phi)}{48}\end{pmatrix},$$

$$\Longrightarrow \quad \Delta^{\pm}_{0,1} = \frac{\varepsilon(2+\varepsilon)}{h^2}\left(\frac{2\pi^2}{3} \pm \frac{27}{16}\right). \qquad (25)$$

Видимо да поправка овде не зависи од угла под којим се врши деформација.

У случају квадратног резонатора се поново најнижи дегенерисани спектар добија за паралелну компоненту струје у c базису репрезентације $E_1^{\pm}$. За компоненте поправке добијамо:

$$\Delta_{0,1} = \frac{\varepsilon(2+\varepsilon)\pi^2}{a^2}\begin{pmatrix}\cos^2(\phi) & 0 \\ 0 & \sin^2(\phi)\end{pmatrix}. \qquad (26)$$

Видимо да је поправка дијагонална и да разбија дегенерацију у првом реду када се деформација не врши по дијагонали квадрата, а да је раздвајање фреквенција најизраженије када се квадрат издужује у правоугаоник.

Кружни резонатор је специфичан по томе што се његова анализа може извршити преласком у поларне координате у којима гранични услов не зависи од угла. Његова група симетрије је лимес бесконачног n, $C_{\infty v}$, односно група цикличних ротација прелази у групу континуалних ротација $C_\infty \to SO(2)$, чија је универзално наткривајућа група $U(1)$. Таблица иредуцибилних репрезентација губи средњу врсту јер $A_{n/2}$ и $B_{n/2}$ репрезентације нису дефинисане, вектори у B_0 репрезентацији су једнаки нули, тако да остаје само једна једнодимензионална иредуцибилна репрезентација A_0, док у последњој врсти бројач дводимензионалних иредуцибилних репрезентација m броји до бесконачности, а параметар $\varphi = \frac{2\pi}{n}k$ постаје континуалан. Базисни вектори су:

$$f_k^{A_0}(r,\varphi) = J_0(kr)\,, \quad f_k^{E_m^{\pm}}(r,\varphi) = J_m(kr)e^{\pm im\varphi}\,. \qquad (27)$$

Као и до сада, постоји s и c базис, такав да је у s базису вредност Беселове финкције на граници једнака нули, а у c базису је вредност првог извода Беселове функције на граници једнака нули. Фреквенција (k,m)-тог мода идеалног резонатора је једнака $\omega_0 = k$.

Структура оператора поправке нам указује да је способан да уклони дегенерацију само унутар $E_1^{\pm}$ репрезентације, јер садржи само сабирке $e^{\pm 2i\varphi}$ и $e^{0i\varphi}$. Раније у случају квадрата и троугла нисмо приметили ову особину, јер се тамо нису појављивале дводимензионалне репрезентације реда већег од 1. Фреквенције у осталим дводимензионалним репрезентацијама остају дегенерисане и бивају само померене као последица деформације. На исти начин се померају и фреквенције унутар A_0 репрезентације. Овај померај фреквенције је једнак:

$$\Delta^{\pm}_{k,m\neq 1} = \varepsilon(2+\varepsilon)\frac{k^2}{2}\,. \qquad (28)$$

У случају репрезентације $E_1^{\pm}$ дегенерација се уклања и помераји фреквенције су једнаки:

$$\Delta^{\pm}_{k,1} = \varepsilon(2+\varepsilon)\frac{k^2}{2} \qquad (29)$$
$$\pm\frac{\varepsilon(2+\varepsilon)}{4}\left(k^2 + \frac{J_1^2(kR)}{R^2\left(J_1^2(kR) - J_0(kR)J_2(kR)\right)}\right),$$

где је R полупречник резонатора. Најниже фреквенције у спектру и овог пута припадају дводимензионалној репрезентацији E_1 у c базису, где је $kR \approx 1.8412$. Очекивано, због $SO(2)$ симетрије круга, поправке фреквенција не зависе од угла под којим се врши деформација.

За више детаља о примени симетрије и механизмима и последицама њеног нарушења погледати [3].

VII Закључак

Модови у c базису E_1 репрезентације су увек имали најнижу фреквенцију, тако да резонатору треба ограничити пропусни опсег само са горње стране. Овде можемо да извршимо поређење фреквенција десетог мода за осциловање идеалних резонатора са све три геометрије и покажемо да са повећањем симетрије (порастом n), спектар резонантних фреквенција постаје гушћи, па је тако за $h = a = 2R = 1$:

$$\omega_{10}^{\triangle} \approx \sqrt{250}\,, \quad \omega_{10}^{\square} \approx \sqrt{168}\,, \quad \omega_{10}^{\circ} \approx \sqrt{114}\,. \qquad (30)$$

Такође је однос између фреквенција у првом дегенерисаном спектру и фреквенције прве следеће моде у спектру опадајући са порастом симетрије и износи $\sqrt{3}$ за троугани, $\sqrt{2}$, за квадратни, и приближно 1.31 за кружни резонатор. Одатле следи да је троугаона геометрија најповољнија за конструкцију резонатора са две блиске резонантне учестаности, јер троугаони резонатор има најразређенији спектар.

VIII Референце

[1] М. Дамњановић "Хилберт-ови простори и групе," Физички факултет, Универзитет у Београду, 2016.

[2] П. Стипсић, М. Миливојевић "Control of a spin qubit in a lateral GaAs quantum dot based on symmetry of gating potential," Phys. Rev. B, vol. 101, issue 16, American Physical Society, pp. 165302-165311, doi 10.1103/PhysRevB.101.165302, April 2020.

[3] М. Дамњановић "О симетрији у квантној нерелативистичкој физици," Физички факултет, Универзитет у Београду, 2000.

УНИВЕРЗИТЕТ У БЕОГРАДУ
ФИЗИЧКИ ФАКУЛТЕТ

Павле Д. Стипсић

Симетрије у вишим градијентним теоријама

докторска дисертација

Београд, 2026.

UNIVERSITY OF BELGRADE
FACULTY OF PHYSICS

Pavle D. Stipsić

Symmetries in Higher Gauge Theories

Doctoral Dissertation

Belgrade, 2026.

БЕЛГРАДСКИЙ УНИВЕРСИТЕТ
ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ

Павле Д. Стипсич

Симметрии в высших калибровочных теориях

Докторская диссертация

Белград, 2026.

Ментор

др Марко Војиновић
научни саветник Института за физику у Београду

Комисија за одбрану докторске дисертације

проф. др Марија Димитријевић Ћирић
редовни професор Физичког факултета Универзитета у Београду

доц. др Драгољуб Гочанин
доцент Физичког факултета Универзитета у Београду

др Игор Прлина
научни сарадник Института за физику у Београду

Датум одбране: 27.2.2026.

Симетрије у вишим градијентним теоријама

Сажетак

У овом раду је испитана симетрија и механизми експлицитног и спонтаног нарушења симетрије класичне 3BF теорије са везама која описује Стандардни Модел спрегнут са Ајнштајн-Картановом гравитацијом. Затим је конструисана непертурбативна веза између квантне 3BF теорије и квантне теорије Ајнштајн-Картанове гравитације са Стандардним Моделом. Коначно, дата је дефиниција квантне 3BF теорије са везама на триангулацији и прелиминарно је анализиран њен семикласичан лимес.

Кључне речи: квантна гравитација, Стандардни Модел, 3BF теорија, више градијентне теорије, симетрија, Хигсов механизам

Научна област: Квантна поља, честице и гравитација

Ужа научна област: Квантна гравитација

Symmetries in Higher Gauge Theories

Abstract

In this work we study the symmetry and mechanisms of explicit and spontaneous symmetry breaking of the classical constrained 3BF theory, which describes the Standard Model coupled to Einstein-Cartan gravity. Then we construct a nonperturbative correspondence between the quantum 3BF theory and the quantum theory of Einstein-Cartan gravity with Standard Model. Finally, we provide a definition of the quantum constrained 3BF theory on a spacetime triangulation and give a preliminary analysis of its semiclassical limit.

Keywords: quantum gravity, Standard Model, 3BF theory, higher gauge theories, symmetry, Higgs mechanism

Field of Study: Quantum fields, particles and gravity

Specific Field of Study: Quantum gravity

Симметрии в высших калибровочных теориях

Аннотация

В данной работе изучаются симметрия и механизмы явного и спонтанного нарушений симметрии классической ограниченной 3BF теории, которая описывает объединение взаимодействий Стандартной Модели с гравитацией Эйнштейна-Картана. Далее строится непертурбативное соответствие между квантовой 3BF теорией и квантовой теорией гравитации Эйнштейна-Картана с полями Стандартной Модели. Наконец, мы даём определение квантовой ограниченной 3BF теории в рамках пространственно-временной триангуляции и предварительный анализ её полуклассического предела.

Ключевые слова: квантовая гравитация, Стандартная Модель, 3BF теория, высшие калибровочные теории, симметрия, механизм Хиггса

Область исследований (общая): квантовая теория поля, физика частиц и теория гравитации

Область исследований (уточнённая): квантовая теория гравитации

Садржај

1 Увод **1**

2 Преглед досадашњих резултата, дефиниције, нотација и конвенције **3**
2.1 Тополошко 3BF дејство . . . 3
2.2 3BF дејство са везама . . . 6
2.3 Ајнштајн-Картанова дејства . . . 10
2.4 Симетрије тополошког 3BF дејства . . . 11
2.5 Тополошка инваријанта за квантну 3BF теорију . . . 13

3 Механизми нарушења симетрије 3BF дејства са везама **15**
3.1 Експлицитно нарушење симетрије . . . 15
3.1.1 Веза за гравитационо поље . . . 16
3.1.2 Веза за скаларна поља . . . 16
3.1.3 Веза за Диракова поља . . . 17
3.1.4 Веза за Јанг-Милсова поља . . . 18
3.1.5 Везе за Хигсов и Јукавин потенцијал, спинска и веза са космолошком константом . . . 18
3.1.6 Преглед нарушења симетрије . . . 19
3.2 Спонтано нарушење симетрије и Хигсов механизам . . . 21
3.2.1 3BF дејство са везама за Прокино поље . . . 22
3.2.2 3BF дејство са везама за теорију електрослабих интеракција . . . 25
3.2.3 Преглед Хигсовог механизма . . . 27
3.2.4 Трансформација дејства . . . 29
3.2.5 Спонтано нарушење 3-групе симетрије и масени спектар фермиона . . . 35

4 Веза између формулација квантне Ајнштајн-Картанове и 3BF теорије са везама **38**
4.1 Квантне опсервабле . . . 38
4.1.1 Математички резултати . . . 38
4.1.2 Очекиване вредности опсервабли . . . 40
4.2 Примери . . . 47
4.2.1 Густина четворозапремине просторвремена . . . 47
4.2.2 Гравитациони таласи . . . 48

5 Конструкција квантне 3BF теорије са везама **52**
5.1 Дискретизација дејства . . . 52
5.2 Прелазак са интеграције по пољима на интеграцију по групама . . . 62
5.2.1 Делта на групи и Питер-Вејлова теорема . . . 64
5.3 Подела дејства на тополошки део и везе . . . 67
5.4 Прелиминарна анализа семикласичног лимеса . . . 70

6 Закључак **76**

A Додаци **77**
A.1 Доказ Теореме 1 77
A.2 Произвољност k-билинеарне форме 77
A.3 Докази идентитета који садрже Диракову делту 79

1 Увод

Формулација теорије квантне гравитације представља један од главних отворених проблема у савременој фундаменталној теоријској физици. Током година је развијено више различитих приступа решавању овог проблема као што су Теорија Струна (ST) [1, 2], Квантна гравитација на петљама (LQG) [3, 4] и остали приступи. Сваки од њих има своје предности и мане. Конкретно, приступ коваријантне квантизације теорије гравитације на петљама [5] се заснива на прецизној дефиницији интеграла по трајекторијама, помоћу ког се дефинише остатак квантне теорије. Једна од главних предности овог приступа је управо та да се оваква прецизна дефиниција може формулисати коришћењем процедуре квантизације модела спинске пене и да се том процедуром може успешно квантовати гравитационо поље. Са друге стране, главна мана овог приступа је да се квантизација може спровести само за чисто гравитационо поље без осталих поља која постоје у природи [6, 7, 8].

Другим речима, иако је једноставно квантовати гравитацију, није очигледно на који начин је истовремено могуће квантовати и гравитацију и осталу материју.

Последњих година је направљен напредак по питању превазилажења овог проблема. Једно од обећавајућих предлога решења је засновано на вишим градијентним теоријама [9, 10], које представљају уопштење појма симетрије коришћењем више теорије категорија. Пажња је усмерена првенствено на структуре под називом n-групе, које су одређен облик уопштења појма групе и користе се за опис градијентне симетрије теорије уместо појма групе [11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24]. Конкретно, испоставља се да структура 3−групе најбоље одговара алгебарском опису свих поља у природи (гравитационог, Јанг-Милсовог, Дираковог и скаларног) [25]. Са друге стране, структура 3−групе омогућава директно уопштење процедуре квантизације у моделу спинске пене на сва остала поља поред гравитационог [26], што омогућава конструкцију квантне теорије гравитације и материје јединственим математичким описом.

Један од главних елемената у конструкцији теорије је дејство BF теорије и његово категоријско уопштење за nBF теорије. Прве резултате формулисања опште теорије релативности и других теорија гравитације заснованих на BF теорији је дао Плебански у свом раду [27], а остале примене за конструкцију различитих модела се могу наћи у [28, 29, 30]. Формулација теорије помоћу 2-групе, названа 2BF или BFCG модел, је по први пут уведена у [31, 32] и даље проучавана у [33, 34, 35, 36, 37, 38], а класичне 3BF и 4BF теорије су формулисане у [25] и [39], редом. На квантном нивоу, nBF теорије дају класу тополошких теорија поља први пут уведених у раду Портера [40], касније у [26, 41].

Добијен је значајан број конкретних резултата у оквиру програма истраживања виших градијентних теорија заснованих на структури 3-групе. Прво је формулисан поступак конструкције класичног дејства који описује цео Стандардни Модел (SM) природно спрегнут са Ајнштајн-Картановом гравитацијом, у форми која је компатибилна са процедуром квантизације модела спинске пене [25]. Овај поступак доводи до реформулације класичне теорије у теорију описану 3BF дејством са везема. Ово дејство се састоји из два главна дела, тополошког 3BF дејства, одређеног структуром 3-групе и дела са везама које деформишу тополошку у нетополошку теорију, дајући јој пропагирајуће степене слободе. Доста истраживања је већ урађено на једноставнијим моделима заснованим на BF и 2BF теоријама [28, 29, 31, 32, 33, 34, 35, 36, 37, 38], а постоји и модел заснован на 4BF теорији [39].

Даље, процедура квантизације тополошког дела је успешно формулисана у облику интеграла по конфигурацијама поља, која одговара тополошкој квантној теорији поља (TQFT) заснованој на датој 3-групи [26]. Осим тога, детаљно су испитане симетрије тополошке 3BF теорије [43, 44], што је довело до бољег разумевања различитих особина модела. Неки од значајних математичких резултата се такође могу наћи у [17, 45, 46].

У наставку рада ће у глави 2 бити дат детаљнији преглед постојећих резултата од значаја за даљи рад, дефиниције појмова, нотација и конвенције. У глави 3 ће бити изложена анализа симетрије и механизама њеног нарушења у теорији заснованој на 3BF моделу са везама. Резултат ове анализе је потврда еквиваленције између стандардне класичне Ајнштајн-Картанове теорије гравитацје и Стандардног Модела са једне стране и класичне теорије засноване на 3BF моделу са везама са друге, у погледу симетрије дејства и Хигсовог механизма спонтаног нарушења симетрије дела који одговара електрослабој интеракцији [47]. У глави 4 је разматрана веза између квантних теорија које се добијају конзистентном квантизацијом Ајнштајн-Картанове теорије гравитације и Стандардног Модела с једне и 3BF теорије са везама са друге стране [48], док је у глави 5 дата дефиниција квантне 3BF теорије са везама.

2 Преглед досадашњих резултата, дефиниције, нотација и конвенције

У овој глави је изложен преглед четири класичне теорије које су релевантне за даљу анализу. У првом поглављу је уведено тополошко 3BF дејство засновано на структури 3-групе. У следећем поглављу је дат метод за конструкцију реалистичне нетополошке теорије додавањем веза. У трећем поглављу је дат преглед стандардног Ајнштајн-Картановог дејства спрегнутог са Стандардним Моделом на уобичајен начин, а потом и у случају постојања сабирка са спин-спин контактном интеракцијом, названом Ајнштајн-Картанова контактна теорија. У четвртом поглављу је дат преглед укупне симетрије тополошког 3BF дејства, а у петом кратка формулација тополошке 3BF квантне теорије поља. Резултати у овој глави проистичу из радова Т. Раденковић и М. Војиновић [25, 26, 43, 44].

Нотација и конвенција су следеће. Просторвременски индекси су означени словима из средине Грчког алфабета $\mu, \nu, \dots$, и подижу се и спуштају просторвременском метриком $g_{\mu\nu}$, када је дефинисана. Лоренцова метрика је означена са $\eta_{ab} = \text{diag}(-1, +1, +1, +1)$. Индекси који пребројавају генераторе Лијевих група G, H, и L су означени почетним словима Грчког алфабета $\alpha, \beta, \dots$, малим латиничним словима са почетка абецеде $a, b, c, \dots$, и великим латиничним словима са почетка абецеде $A, B, C, \dots$, редом. Сами генератори су означени као τ_α, t_a и T_A, редом. Коришћен је природни систем јединица, дефинисан са $c = \hbar = 1$ и $G = l_p^2$, где је l_p Планкова дужина.

Индекси који одговарају Лоренцовој групи су парови индекса ab и величине које зависе од њих су антисиметричне на замену њихових места. То значи да су све независне компоненте ових величина према Ајнштајновој сумационој конвенцији пребројане два пута. Због тога је резултат суме потребно поделити са два. Алтернативно, у циљу избегавања овог проблема, може се увести нотација $[ab]$ која репрезентује пар индекса као један индекс за који је увек претпостављено да важи $a > b$. Сумација по оваквим индексима пребројава сваку независну компоненту тачно једном, па суму није потребно поделити са два. На пример, за дату величину K^{ab}, добија се

$$K^{[ab]}\sigma_{[ab]} = \frac{1}{2}K^{ab}\sigma_{ab} \,. \tag{1}$$

У овом раду ће правоугаоне заграде међу индексима бити коришћене искључиво за означавање пара Лоренцових индекса, уместо уобичајене примене за антисиметризацију индекса.

2.1 Тополошко 3BF дејство

У циљу конструисања тополошког 3BF дејства, потребно је кренути од појма стриктне Лијеве 3-групе, која је уопштење појма Лијеве групе проистекло из више теорије категорија. Ова структура је еквивалентна структури Лијевог 2-укрштеног модула. Лијев 2-укрштени модул је уређена тројка три Лијеве групе, G, H и L, заједно са два хомоморфизма између њих,

$$\partial : H \to G \,, \qquad \delta : L \to H \,, \tag{2}$$

дејства групе G на све три групе,

$$\triangleright : G \times X \to X \,, \qquad X = G, H, L \,, \tag{3}$$

и пресликавања Пајферовог подизања,

$$\{\_ \,, \_\}_{\text{pf}} : H \times H \to L \,. \tag{4}$$

Сва три пресликавања су обухваћена одређеним скупом аксиома и заједно чине структуру звану Лијев 2-укрштени модул, који се означава као

$$(L \xrightarrow{\delta} H \xrightarrow{\partial} G \,,\, \triangleright \,,\, \{\_\,,\_\}_{\rm pf}) \,. \tag{5}$$

Ова структура представља појам 3-групе на најуобичајенији начин. Даље математичке аксиоме и дефиниције се могу наћи у референцама [9, 17, 25, 26, 41, 44, 45, 46, 49].

Дата математичка структура 3-групе намеће природан избор дејства звано 3BF дејство. Ово дејство је чисто тополошко, дефинисано са:

$$S_{3BF}^{\rm top} = \int_{\mathcal{M}_4} \langle B \wedge \mathcal{F} \rangle_{\mathfrak{g}} + \langle C \wedge \mathcal{G} \rangle_{\mathfrak{h}} + \langle D \wedge \mathcal{H} \rangle_{\mathfrak{l}}. \tag{6}$$

Лагранжеви множитељи B, C и D су два-, један- и нула-форме и истовремено елементи Лијевих алгебри $\mathfrak{g}$, $\mathfrak{h}$ и $\mathfrak{l}$, које одговарају Лијевим групама G, H и L, редом. Јачине поља $\mathcal{F}$, $\mathcal{G}$ и $\mathcal{H}$ су дефинисане као

$$\mathcal{F} = {\rm d}\alpha + \alpha \wedge \alpha - \partial\beta \,, \qquad \mathcal{G} = {\rm d}\beta + \alpha \wedge^{\triangleright} \beta - \delta\gamma \,, \qquad \mathcal{H} = {\rm d}\gamma + \alpha \wedge^{\triangleright} \gamma + \{\beta \wedge \beta\}_{\rm pf} \,, \tag{7}$$

и називају се лажним кривинама за конексије један-форме α, два-форме β и три-форме γ, које су такође елементи алгебра $\mathfrak{g}$, $\mathfrak{h}$ и $\mathfrak{l}$, редом. Билинеарне форме $\langle \_\,,\_ \rangle_{\mathfrak{g}}$, $\langle \_\,,\_ \rangle_{\mathfrak{h}}$ и $\langle \_\,,\_ \rangle_{\mathfrak{l}}$ треба да буду симетрична, недегенерисана и G-инваријантна пресликавања пара елемената алгебре у скуп реалних бројева. На основу структуре 3-групе, може се још увести и појам коваријантног извода дефинисаног као

$$\nabla = {\rm d} + \alpha \wedge^{\triangleright} \tag{8}$$

у смислу да када ∇ делује на, на пример, компоненте ϕ^A елемента $\phi \in \mathfrak{l}$, дејство $\triangleright$ представља дејство елемента Лијеве алгебре $\mathfrak{g}$ на елемент Лијеве алгебре $\mathfrak{l}$, на начин:

$$\nabla\phi^A = {\rm d}\phi^A + \triangleright_{\alpha B}{}^{A}\, \alpha^\alpha \wedge \phi^B \,, \tag{9}$$

и слично за елементе алгебри $\mathfrak{g}$ и $\mathfrak{h}$. Коришћењем ове нотације се лажне кривине (7) могу преписати у функцији од обичних кривина:

$$\mathcal{F} = \nabla^2 - \partial\beta \,, \qquad \mathcal{G} = \nabla\beta - \delta\gamma \,, \qquad \mathcal{H} = \nabla\gamma + \{\beta \wedge \beta\}_{\rm pf} \,. \tag{10}$$

Дејство свих пресликавања $\triangleright$, ∂, δ и $\{\_\,,\_\}_{\rm pf}$, као и билинеарних форми на базис генератора Лијевих група је:

$$\begin{gathered} \tau_\alpha \triangleright \tau_\beta = \triangleright_{\alpha\beta}{}^{\gamma}\tau_\gamma = f_{\alpha\beta}{}^{\gamma}\tau_\gamma \,, \qquad \tau_\alpha \triangleright t_a = \triangleright_{\alpha a}{}^{b} t_b \,, \qquad \tau_\alpha \triangleright T_A = \triangleright_{\alpha A}{}^{B} T_B \,, \\ \partial t_a = \partial_a{}^{\alpha}\tau_\alpha \,, \qquad \delta T_A = \delta_A{}^{a} t_a \,, \qquad \{t_a\,, t_b\} = X_{ab}{}^{A} T_A \,, \\ \langle \tau_\alpha \,, \tau_\beta \rangle_{\mathfrak{g}} = g_{\alpha\beta} \,, \qquad \langle t_a \,, t_b \rangle_{\mathfrak{h}} = g_{ab} \,, \qquad \langle T_A \,, T_B \rangle_{\mathfrak{l}} = g_{AB} \,. \end{gathered} \tag{11}$$

Пошто су билинеарне форме недегенерисане, постоје њихови инверзи и компоненте тих инверза су означене са $g^{\alpha\beta}$, g^{ab} и g^{AB}. Такође се користе за подизање и спуштање групних индекса.

Да би се помоћу ове структуре конструисало дејство које одговара Стандардном Моделу и Ајнштајн-Картановој гравитацији, потребно је изабрати одговарајућу 3-групу, звану 3-група Стандардног Модела [25, 47, 50]. Три Лијеве групе G, H и L се бирају да буду:

$$G = SO(3,1) \times SU(3) \times SU(2) \times U(1) \,, \qquad H = \mathbb{R}^4 \,, \qquad L = \mathbb{C}^4 \times \mathbb{G}^{64} \times \mathbb{G}^{64} \times \mathbb{G}^{64} \,. \tag{12}$$

Физичка интерпретација овог избора је следећа. Група G представља уобичајену групу градијентних симетрија Стандардног Модела, заједно са групом локалне Лоренцове симетрије. Група H представља просторвременске транслације, док група L одговара пољима материје. Конкретно, $\mathbb{C}^4$ одговара Хигсовом сектору, док три Грасманове групе $\mathbb{G}^{64}$ одговарају трима фамилијама фермиона.

Да би 3-група Стандардног Модела била у потпуности дефинисана, потребно је одабрати и сва пратећа пресликавања. Хомоморфизми ∂ и δ се бирају тако да буду тривијални, као и Пајферово подизање $\{\_\,,\_\}_{\mathrm{pf}}$. Дејство $\triangleright$, је дефинисано на следећи начин. Група G се може природно поделити на Лоренцов део $SO(3,1)$ (генератори су пребројани коришћењем пара индекса $[ab]$) и унутрашње градијентне симетрије $SU(3)\times SU(2)\times U(1)$ (генератори су пребројани индексима $\alpha,\beta,\dots$). Дејство групе G на саму себе је онда дато избором дејства Лоренцовог дела на самог себе и део унутрашњих симетрија, као

$$\triangleright_{[ab][cd]}{}^{[ef]} \equiv f_{[ab][cd]}{}^{[ef]} = \frac{1}{2}\left(\eta_{[a|c}\delta^{[f|}_{|b]}\delta^{|e]}_{d} - \eta_{[a|d}\delta^{[f|}_{|b]}\delta^{|e]}_{c}\right)\,, \qquad \triangleright_{[ab]\beta}{}^{\gamma} = 0\,, \tag{13}$$

док је дејство унутрашње групе симетрије на саму себе и Лоренцов део дато са

$$\triangleright_{\alpha\beta}{}^{\gamma} = f_{\alpha\beta}{}^{\gamma}\,, \qquad \triangleright_{\alpha[ab]}{}^{[cd]} = 0\,. \tag{14}$$

Даље, дејство групе G на групу H је дефинисано природно, претпостављајући да је група H интерпретирана као група четвородимензионалних транслација. Тада Лоренцов део групе G делује на стандардан начин на транслације, док унутрашњи део групе G делује тривијално:

$$\triangleright_{[cd]a}{}^{b} = \frac{1}{2}\eta_{a[d|}\delta^{b}_{|c]}\,, \qquad \triangleright_{\alpha a}{}^{b} = 0\,. \tag{15}$$

Коначно, дејство целе групе G на L је такође дато на природан начин, у складу са трансформационим особинама различитих фермиона и Хигсовог поља. На пример, дејство групе G на лептон левог изоспина је:

$$\triangleright_{[cd]A}{}^{B} = \left(\sigma_{cd}\right)_{A}{}^{B}\,, \qquad \triangleright_{\alpha A}{}^{B} = \frac{1}{2}\left(\sigma_{\alpha}\right)_{A}{}^{B}\,. \tag{16}$$

Матрице $\left(\sigma_{\alpha}\right)_{A}{}^{B}$ су Паулијеве матрице и $\left(\sigma_{ab}\right)_{A}{}^{B} = \frac{1}{4}[\gamma_a,\gamma_b]_A{}^B$, где су γ_a стандардне Диракове матрице које задовољавају антикомутациону релацију $\gamma_a\gamma_b+\gamma_b\gamma_a = -2\eta_{ab}$. Стандардно се може увести и матрица $\gamma_5 \equiv -\gamma_0\gamma_1\gamma_2\gamma_3$. На сличан начин се може дефинисати и дејство групе G на све остале фермионе и скаларна поља у групи L, зависно од њихових трансформационих особина [25].

Додатно, поред прецизирања 3-групе Стандардног Модела, дејство (6) такође зависи од избора билинеарних форми. У случају не-Абелове групе је природан избор ове форме Картан-Килингова форма (Картанов тензор), док у случају Абелових група не постоји природан избор, али је он ограничен захтевом да билинеарна форма мора да буде G-инваријантна и недегенерисана. Узимајући ово у обзир, за 3-групу Стандардног Модела и дејство (6) се бирају следеће билинеарне форме. За алгебру $\mathfrak{g}$:

$$g_{[ab][cd]} = \frac{1}{2}\eta_{d[a|}\eta_{|b]c}\,, \qquad g_{\alpha\beta} = \delta_{\alpha\beta}\,, \qquad g_{\alpha[ab]} = 0\,. \tag{17}$$

За алгебру $\mathfrak{h}$, због G-инваријантности

$$g_{ab} = \eta_{ab}\,. \tag{18}$$

Коначно, у случају алгебре $\mathfrak{l}$ је ситуација компликованија јер Грасманови бројеви антикомутирају. Конкретно, за било који избор $A, B \in \mathfrak{l}$, важи:

$$\langle A, B\rangle_{\mathfrak{l}} = A^I B^J g_{IJ}\,, \qquad \langle B, A\rangle_{\mathfrak{l}} = B^J A^I g_{JI}\,. \tag{19}$$

Пошто билинеарне форме морају да буду симетричне, два израза морају да буду једнака. Међутим, у зависности да ли су коефицијенти A^I и B^J Грасманови бројеви или обични реални бројеви, они ће или антикомутирати или комутирати, што за последицу има да је матрица компонената билинеарне форме g_{IJ} антисиметрична или симетрична, редом. Због тога се генератори T_A алгебре $\mathfrak{l}$ могу груписати у три класе: $T_{\tilde{A}}$ која припада Хигсовом сектору, и пару $T_{\hat{A}}$, $T^{\hat{A}}$ који припада фермионском сектору. Онда се компоненте билинеарне форме могу записати као

$$g_{AB} = \left[\begin{array}{c|cc} \delta_{\tilde{A}\tilde{B}} & 0 & 0 \\ \hline 0 & 0 & \delta_{\hat{A}}{}^{\hat{B}} \\ 0 & -\delta^{\hat{B}}_{\hat{A}} & 0 \end{array}\right]. \tag{20}$$

Горњи леви блок одговара алгебри $\mathbb{C}^4$, док доњи десни блок одговара алгебри $\mathbb{G}^{64} \times \mathbb{G}^{64} \times \mathbb{G}^{64}$.

Након задавања 3-групе за Стандардни Модел, дејство за одговарајућу тополошку 3BF теорију (6) се може записати у следећем облику:

$$S^{\text{top}}_{3BF} = \int B_\alpha \wedge F^\alpha + B^{[ab]} \wedge R_{[ab]} + e_a \wedge \nabla\beta^a + \phi^A(\nabla\tilde{\gamma})_A + \bar{\psi}_A(\overrightarrow{\nabla}\gamma)^A - (\bar{\gamma}\overleftarrow{\nabla})_A\psi^A\,, \tag{21}$$

где је уведена нова нотација. Прво, $\mathcal{F}$ је подељен на јачину поља која одговара групи унутрашње симетрије F^α (која је функција унутрашње конексије α^α) и два-форму Риманове кривине $R_{[ab]}$ (која је функција спинске конексије $\omega^{[ab]}$). Лагранжев множитељ C је замењен један-формом поља тетраде e_a, а Лагранжев множитељ D је замењен компонентама Хигсовог поља и поља фермиона $(\phi^A, \psi^A, \bar{\psi}_A)$. Ова замена нотације представља физичку интерпретацију поља у дејству (6).

2.2 3BF дејство са везама

Иако дејство (21) има одговарајућу симетрију описану 3-групом Стандардног Модела и садржи сва градијентна, гравитационо и поља материје, једначине кретања које се добијају варијацијом по пољима не одговарају класичним једначинама кретања ових поља. Заправо, ово дејство је пример *тополошког* 3BF дејства, и као такво даје тривијалне једначине кретања, без пропагирајућих степени слободе. У циљу кориговања теорије се додају сабирци у дејству који представљају везе зване *симплицијалне везе*. Пажљивим избором ових веза које се додају у дејство, могуће је конструисати нетополошко дејство које даје једначине кретања за поља сагласне једначинама кретања у Стандардном Моделу спрегнутим са Ајнштајн-Картановом гравитацијом. Такво дејство се често назива 3BF дејство са везама, и има следећи облик:

$$S_{3BF} = S^{\text{top}}_{3BF} + S_{\text{grav}} + S_{\text{scal}} + S_{\text{Dirac}} + S_{\text{Yang-Mills}} + S_{\text{Higgs}} + S_{\text{Yukawa}} + S_{\text{spin}} + S_{\text{CC}}\,. \tag{22}$$

Његови појединачни делови су:

$$
\begin{aligned}
S^{\text{top}}_{3BF} &= \int B_\alpha \wedge F^\alpha + B^{[ab]} \wedge R_{[ab]} + e_a \wedge \nabla\beta^a + \phi^A(\nabla\tilde{\gamma})_A + \bar{\psi}_A(\overset{\rightarrow}{\nabla}\gamma)^A - (\bar{\gamma}\overset{\leftarrow}{\nabla})_A\psi^A \,, && (23)\\
S_{\text{grav}} &= -\int \lambda_{[ab]} \wedge \left(B^{[ab]} - \frac{1}{8\pi l_p^2}\varepsilon^{[ab]cd} e_c \wedge e_d \right) , && (24)\\
S_{\text{scal}} &= \int \tilde{\lambda}^A \wedge \left(\tilde{\gamma}_A - H_{abcA} e^a \wedge e^b \wedge e^c\right) \\
&+ \Lambda^{abA} \wedge \left(H_{abcA}\varepsilon^{cdef} e_d \wedge e_e \wedge e_f - (\nabla\phi)_A \wedge e_a \wedge e_b \right) , && (25)\\
S_{\text{Dirac}} &= \int \bar{\lambda}_A \wedge \left(\gamma^A + \frac{i}{6}\varepsilon_{abcd} e^a \wedge e^b \wedge e^c \left(\gamma^d\psi\right)^A \right) \\
&- \lambda^A \wedge \left(\bar{\gamma}_A - \frac{i}{6}\varepsilon_{abcd} e^a \wedge e^b \wedge e^c \left(\bar{\psi}\gamma^d\right)_A \right) , && (26)\\
S_{\text{Yang-Mills}} &= \int \lambda^\alpha \wedge \left(B_\alpha - 12 C_{\alpha\beta} M^\beta{}_{ab} e^a \wedge e^b \right) \\
&+ \zeta_\alpha{}^{ab} \left(M^\alpha{}_{ab}\varepsilon_{cdef} e^c \wedge e^d \wedge e^e \wedge e^f - F^\alpha \wedge e_a \wedge e_b \right) , && (27)\\
S_{\text{Higgs}} &= -\int \frac{2}{4!}\chi \left(\phi^A\phi_A - v^2\right)^2 \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d \,, && (28)\\
S_{\text{Yukawa}} &= -\int \frac{2}{4!} Y_{ABC}\bar{\psi}^A\psi^B\phi^C \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d \,, && (29)\\
S_{\text{spin}} &= \int 2\pi i l_p^2 \bar{\psi}_A \gamma_5 \gamma^a \psi^A \varepsilon_{abcd} e^b \wedge e^c \wedge \beta^d \,, && (30)\\
S_{\text{CC}} &= -\int \frac{1}{96\pi l_p^2}\Lambda \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d \,. && (31)
\end{aligned}
$$

Док би записивање целог дејства у једном изразу било компликовано и нечитко, овако издељеном се може протумачити значење и сврха сваког сабирка:

- тополошки 3BF сабирак (23) је идентичан једначини (21), која набраја сва поља присутна у теорији (онако како намеће структура 3-групе Стандардног Модела),
- веза за гравитационо поље (24) је задужена за стварање гравитационих степени слободе,
- веза за скаларна поља (25) је задужена за стварање степени слободе безмасених скаларних поља,
- веза за Диракова поља (26) је задужена за стварање степени слободе Диракових поља,
- веза за Јанг-Милсова поља (27) је задужена за стварање степени слободе градијентних поља,
- веза за Хигсов потенцијал (28) описује самоинтеракције и масу Хигсових поља,
- веза за Јукавину интеракцију (29) описује интеракцију између Хигсових и Диракових поља, као и матрице мешања фермиона,
- спинска веза (30) је неопходна за исправан опис интеракције између спина фермиона и торзије, и

- веза са космолошком константом (31) уводи сабирак пропорционалан космолошкој константи.

Слободни параметри присутни у дејству су:

- l_p је Планкова дужина, садржана у S_{grav}, S_{spin} и S_{CC},
- $C_{\alpha\beta}$ представља билинеарну форму интеракционих константи градијентних поља, садржана у $S_{\text{Yang-Mills}}$,
- χ је интеракциона константа за самоинтеракцију четвртог степена Хигсовог поља, садржана у S_{Higgs},
- v је вакуумски очекивана вредност Хигсовог поља, такође садржана у S_{Higgs},
- Y_{ABC} представља константе Јукавине интеракције и матрице мешања фермиона, садржане у S_{Yukawa}, и
- Λ је космолошка константа, садржана у S_{CC}.

Тополошки део S^{top}_{3BF} и везе S_{scal} и S_{Dirac} не садрже слободне параметре.

Једначине кретања добијене варијацијом целог дејства се могу поделити у две групе. Прву чине једначине кретања за Лагранжеве множитеље, које се могу решити тако да се Лагранжеви множитељи изразе у функцији од динамичких поља и њихових извода ([25, 47]):

$$
\begin{aligned}
M_{\alpha ab} &= -\frac{1}{48}\varepsilon_{abcd}F_{\alpha}{}^{\mu\nu}e^{c}{}_{\mu}e^{d}{}_{\nu}\,, & \zeta^{\alpha ab} &= \frac{1}{4}C_{\beta}{}^{\alpha}\varepsilon^{abcd}F^{\beta}{}_{\mu\nu}e_{c}{}^{\mu}e_{d}{}^{\nu}\,,\\
\lambda_{\alpha\mu\nu} &= -F_{\alpha\mu\nu}\,, & B_{\alpha\mu\nu} &= -\frac{e}{2}C_{\alpha}{}^{\beta}\varepsilon_{\mu\nu\rho\sigma}F_{\beta}{}^{\rho\sigma}\,,\\
\lambda_{[ab]\mu\nu} &= R_{[ab]\mu\nu}\,, & B_{[ab]\mu\nu} &= \frac{1}{8\pi l_p^2}\varepsilon_{[ab]cd}e^{c}{}_{\mu}e^{d}{}_{\nu}\,,\\
\tilde{\lambda}^{A}{}_{\mu} &= \left(\nabla_\mu\phi\right)^A\,, & \tilde{\gamma}^{A}{}_{\mu\nu\rho} &= -e\varepsilon_{\mu\nu\rho\sigma}\left(\nabla^\sigma\phi\right)^A\,,\\
H^{abcA} &= \frac{1}{6e}\varepsilon^{\mu\nu\rho\sigma}\left(\nabla_\mu\phi\right)^A e^{a}{}_{\nu}e^{b}{}_{\rho}e^{c}{}_{\sigma}\,, & \Lambda^{abA}{}_{\mu} &= \frac{1}{6e}g_{\mu\lambda}\varepsilon^{\lambda\nu\rho\sigma}\left(\nabla_\nu\phi\right)^A e^{a}{}_{\rho}e^{b}{}_{\sigma}\,,\\
\gamma^{A}{}_{\mu\nu\rho} &= -i\varepsilon_{abcd}e^{a}{}_{\mu}e^{b}{}_{\nu}e^{c}{}_{\rho}\left(\gamma^d\psi\right)^A\,, & \bar{\gamma}_{A\mu\nu\rho} &= i\varepsilon_{abcd}e^{a}{}_{\mu}e^{b}{}_{\nu}e^{c}{}_{\rho}\left(\bar{\psi}\gamma^d\right)_A\,,\\
\lambda^{A}{}_{\mu} &= \left(\overrightarrow{\nabla}_\mu\psi\right)^A\,, & \bar{\lambda}_{A\mu} &= \left(\bar{\psi}\overleftarrow{\nabla}_\mu\right)_A\,,\\
\beta^{a}{}_{\mu\nu} &= 0\,.
\end{aligned}
\tag{32}
$$

Другу групу чине једначине кретања за динамичка поља. Спинска конексија $\omega^{[ab]}{}_{\mu}$ није еквивалентна Леви-Чивита конексији, будући да фермионска поља генеришу ненулту торзију. Због тога је потребно прво поделити спинску конексију на збир Ричијевих ротационих коефицијената $\Delta^{[ab]}{}_{\mu}$ и тензор конторзије $K^{[ab]}{}_{\mu}$:

$$
\omega^{[ab]}{}_{\mu} = \Delta^{[ab]}{}_{\mu} + K^{[ab]}{}_{\mu}\,. \tag{33}
$$

Ричијеви ротациони коефицијенти су дати као линеарна комбинација комутационих коефицијената

$$
\Delta^{ab}{}_{\mu} = \frac{1}{2}\left(c^{abc} - c^{bac} - c^{cab}\right)e_{c\mu}\,, \tag{34}
$$

који су дефинисани као

$$c^{abc} = e^{b\mu} e^{c\nu} \left(\partial_\mu e^a{}_\nu - \partial_\nu e^a{}_\mu \right) . \tag{35}$$

Тензор конторзије је дат изразом:

$$K^{ab}{}_\mu = \frac{1}{2} \left(T^{cab} + T^{bac} - T^{abc} \right) e_{c\mu} . \tag{36}$$

Овде је $T^{abc} \equiv T^a{}_{\mu\nu} e^{b\mu} e^{c\nu}$, где су $T^a{}_{\mu\nu}$ компоненте 2-форме торзије, дефинисане као:

$$T^a \equiv \nabla e^a = \frac{1}{2} T^a{}_{\mu\nu} \, \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu , \qquad T^a{}_{\mu\nu} \equiv \nabla_\mu e^a{}_\nu - \nabla_\nu e^a{}_\mu . \tag{37}$$

Комбиновањем свих претходних једначина, једначина кретања за торзију се може записати у облику:

$$T^a = 2\pi i l_p^2 \, s^a , \qquad s^a \equiv \varepsilon^{abcd} \, e_b \wedge e_c \, \bar{\psi}_A \gamma_5 \gamma_d \psi^A . \tag{38}$$

Сада је очигледно да је 2-форма торзије пропорционална спинској 2-форми s^a. Као што ће се испоставити у следећем поглављу, овај резултат је исти и у Ајнштајн-Картановој теорији гравитације. Такође, коришћењем (36), компоненте 1-форме конторзије су

$$K^{ab}{}_\mu = -2\pi i l_p^2 \bar{\psi}_A \gamma_5 \gamma_d \psi^A \varepsilon^{abcd} e_{c\mu} , \tag{39}$$

па се веза између конторзије и торзије поједностављује и постаје:

$$T^a = K^{ab} \wedge e_b . \tag{40}$$

Даље, Ајнштајнова једначина има уобичајен облик:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \Lambda g_{\mu\nu} = 8\pi l_p^2 \, T_{\mu\nu} , \tag{41}$$

где је тензор енергије-импулса:

$$\begin{aligned} T_{\mu\nu} &= F^\alpha{}_{\mu\rho} C_\alpha{}^\beta F_{\beta\nu}{}^\rho - \frac{1}{4} g_{\mu\nu} F^\alpha{}_{\rho\sigma} C_\alpha{}^\beta F_\beta{}^{\rho\sigma} \\ &+ \nabla_\mu \phi^A \nabla_\nu \phi_A - \frac{1}{2} g_{\mu\nu} \left(\nabla_\rho \phi^A \nabla^\rho \phi_A + 2\chi \left(\phi^A \phi_A - v^2 \right)^2 \right) \\ &+ \frac{i}{2} \left(\bar{\psi}_A \overleftrightarrow{\nabla}_\mu \gamma_d \psi^A \right) e^d_\nu - \frac{1}{2} g_{\mu\nu} \left(i \left(\bar{\psi}_A \overleftrightarrow{\nabla}_\rho \gamma^d \psi^A \right) e_d{}^\rho - 2 Y_{ABC} \bar{\psi}^A \psi^B \phi^C \right) . \end{aligned} \tag{42}$$

Тензор садржи три дела који одговарају Јанг-Милсовом, скаларном и Дираковом пољу, редом.

Једначине кретања за фермионска и скаларна поља су:

$$\left(i\gamma^\mu \overrightarrow{\nabla}_\mu \delta^A_B - Y^A{}_{BC} \phi^C \right) \psi^B = 0 , \tag{43}$$

$$\bar{\psi}_B \left(\delta^B_A i \overleftarrow{\nabla}_\mu \gamma^\mu + Y_{BAC} \phi^C \right) = 0 , \tag{44}$$

$$\nabla_\mu \nabla^\mu \phi^A - 4\chi \left(\phi^B \phi_B - v^2 \right) \phi^A = 0 , \tag{45}$$

док је једначина кретања за Јанг-Милсово поље:

$$\nabla_\mu F_\alpha{}^{\mu\nu} + \frac{1}{2} C^{-1}{}_\alpha{}^\beta \left(\rhd_{\beta AB} \left(\phi^A \nabla^\nu \phi^B - \phi^B \nabla^\nu \phi^A \right) + i \bar{\psi}_A \psi_B \left(\rhd_{\beta C}{}^A \gamma^{\nu CB} - \gamma^{\nu AC} \rhd_{\beta C}{}^B \right) \right) = 0 . \tag{46}$$

На основу претходних једначина се може закључити да све једначине кретања тачно одговарају једначинама кретања Стандардног Модела спрегнутог са Ајнштајн-Картановом гравитацијом са космолошком константом.

Из дефиниције дејства (22) следи да цео конфигурациони простор теорије садржи и нединамичка поља-Лагранжеве множитеље

$$M_{\alpha ab}\,,\zeta^{\alpha ab}\,,\lambda_{\alpha\mu\nu}\,,B_{\alpha\mu\nu}\,,\lambda_{[ab]\mu\nu}\,,B_{[ab]\mu\nu}\,,\tilde{\lambda}^A{}_\mu\,,\tilde{\gamma}^A{}_{\mu\nu\rho}\,,H^{abcA}\,,\Lambda^{abA}{}_\mu\,,\gamma^A{}_{\mu\nu\rho}\,,\bar{\gamma}_{A\mu\nu\rho}\,,\lambda^A{}_\mu\,,\bar{\lambda}_{A\mu}\,,\beta^a{}_{\mu\nu}\,, \tag{47}$$

као и динамичка поља

$$\omega^{[ab]}{}_\mu\,,e^a{}_\mu\,,\phi^A\,,\psi^A\,,\bar{\psi}_A\,,\alpha^\alpha{}_\mu\,. \tag{48}$$

Разлика између динамичких и нединамичких поља је последица једначина кретања. Једначине кретања за Лагранжеве множитеље су алгебарске, док су једначине кретања за динамичка поља диференцијалне. Једини изузетак од овог правила је једначина кретања за торзију (38), која се може решити по спинској конексији као функцији тетрада и Диракових поља, указујући на то да је спинска конексија такође нединамичко поље. Ово је добро позната особина Ајнштајн-Картанове теорије, али је уобичајено да се спинска конексија ипак убраја у динамичка поља, јер у општијим теоријама у Риман-Картановим просторима спинска конексија може да постане право динамичко поље [42, 51].

2.3 Ајнштајн-Картанова дејства

Стандардно Ајнштајн-Картаново (ЕС) дејство је уобичајено наведено у литератури (на пример у [42]) као збир дејства за Стандардни Модел минимално спрегнутог са гравитацијом и Ајнштајн-Хилбертовог дејства за гравитационо поље (израженог у формализму првог реда). У досадашњој нотацији је то:

$$\begin{aligned} S_{EC}[e,\omega,\phi,\psi,\bar{\psi},\alpha] \;=\; & \int \frac{1}{16\pi l_p^2}\varepsilon^{abcd}\,R_{ab}\wedge e_c\wedge e_d - F^\alpha\wedge C_{\alpha\beta}\star F^\beta - (\nabla\phi)^A\wedge(\star\nabla\phi)_A \\ & -\frac{i}{6}\varepsilon_{abcd}\,e^a\wedge e^b\wedge e^c\wedge\left[(\bar{\psi}\overleftarrow{\nabla})_A\gamma^d\psi^A - \bar{\psi}_A\gamma^d(\overrightarrow{\nabla}\psi)^A\right] \\ & -\frac{1}{12}\left[\chi\left(\phi^A\phi_A - v^2\right)^2 + Y_{ABC}\bar{\psi}^A\psi^B\phi^C + \frac{\Lambda}{8\pi l_p^2}\right]\varepsilon_{abcd}\,e^a\wedge e^b\wedge e^c\wedge e^d\,. \end{aligned} \tag{49}$$

Овде $\star F$ означава Хоџов дуал 2-форме F, слично као и дуал 1-форме $\nabla\phi$:

$$\star F^\alpha = \frac{1}{4}F^\alpha{}_{cd}\varepsilon^{abcd}e_a\wedge e_b\,,\qquad (\star\nabla\phi)_A = \frac{1}{3!}\left(\nabla_d\phi\right)_A\varepsilon^{dabc}e_a\wedge e_b\wedge e_c\,. \tag{50}$$

Конфигурациони простор ове теорије је еквивалентан конфигурационом простору динамичких поља 3BF теорије (48), где спинска конексија поново задовољава алгебарску једначину

$$\omega_{ab\mu} = \Delta_{ab\mu} - 2\pi i l_p^2\bar{\psi}_A\gamma_5\gamma^d\psi^A\varepsilon_{abcd}e^c{}_\mu\,, \tag{51}$$

добијену из једначине кретања за торзију аналогну једначини (38), па не представља динамичко поље. Због тога је заменом ове једначине натраг у дејство могуће конструисати еквивалентну класичну теорију са редукованим конфигурационим простором и истим једначинама кретања. Ова еквивалентна теорија се у литератури обично назива *теорија другог* *реда* [42]. Замена се изводи раздвајањем спинске конексије на конторзију и Ричијеве ротационе коефицијенте, и раздвајањем њихових доприноса у дејству. То је еквивалентно следећој замени:

$$\nabla_{(\omega)} = \nabla_{(\Delta)} + \frac{1}{2} K^{ab} \sigma_{ab} \tag{52}$$

У Ајнштајн-Картановом дејству постоје само два сабирка која зависе од спинске конексије. То су сабирак са Римановим тензором кривине и Лагранжијан за Дираково поље. После замене ови сабирци постају:

$$\begin{aligned} R_{ab} &= \mathrm{d}\omega_{ab} + (\omega\wedge\omega)_{ab} = \mathrm{d}\Delta_{ab} + (\Delta\wedge\Delta)_{ab} + \mathrm{d}K_{ab} + (\Delta\wedge K)_{ab} + (K\wedge\Delta)_{ab} + K_a{}^c \wedge K_{cb} \\ &= R_{ab}(\Delta) + \left(\nabla_{(\Delta)} K\right)_{ab} + K_a{}^c \wedge K_{cb}\,, \end{aligned} \tag{53}$$

и

$$\begin{aligned} \left(\bar{\psi}\overleftarrow{\nabla}\right)_A \gamma^d \psi^A - \bar{\psi}_A \gamma^d \left(\nabla\psi\right)^A &= \left(\bar{\psi}\overleftarrow{\nabla}_{(\Delta)}\right)_A \gamma^d \psi^A - \bar{\psi}_A \gamma^d \left(\nabla_{(\Delta)}\psi\right)^A - \frac{1}{2}\bar{\psi}_A \left\{\sigma^{ab}, \gamma^d\right\} \psi^A K_{ab} \\ &= \left(\bar{\psi}\overleftarrow{\nabla}_{(\Delta)}\right)_A \gamma^d \psi^A - \bar{\psi}_A \gamma^d \left(\nabla_{(\Delta)}\psi\right)^A + \frac{1}{2}\varepsilon^{abcd}\bar{\psi}_A \gamma_5 \gamma_c \psi^A K_{ab}\,. \end{aligned} \tag{54}$$

Када се ови сабирци врате натраг у стандардно Ајнштајн-Картаново дејство, након сређивања се добије дејство истог облика као почетно са додатним сабирком који представља контактну спин-спин интеракцију и фиксираном спинском конексијом Δ_{ab}. Овај додатни сабирак, изражен преко конторзије, је $\frac{1}{8\pi l_p^2} K^{ab} \wedge \star K_{ab}$, па се коришћењем релације између тензора спина и конторзије из стандардног Ајнштајн-Картановог дејства може елиминисати конторзија:

$$\begin{aligned} S_{ECC}[e,\phi,\psi,\bar{\psi},\alpha] &= \int \frac{1}{16\pi l_p^2}\varepsilon^{abcd} R_{ab}\wedge e_c \wedge e_d - F^\alpha \wedge C_{\alpha\beta} \star F^\beta - (\nabla\phi)^A \wedge (\star\nabla\phi)_A \\ &\quad - \frac{i}{6}\varepsilon_{abcd}\, e^a \wedge e^b \wedge e^c \wedge \left[(\bar{\psi}\overleftarrow{\nabla})_A \gamma^d \psi^A - \bar{\psi}_A \gamma^d (\overrightarrow{\nabla}\psi)^A \right] + \frac{3\pi l_p^2}{4} s^a \wedge \star s_a \\ &\quad - \frac{1}{12}\left[\chi\left(\phi^A \phi_A - v^2\right)^2 + Y_{ABC}\bar{\psi}^A \psi^B \phi^C + \frac{\Lambda}{8\pi l_p^2} \right] \varepsilon_{abcd}\, e^a \wedge e^b \wedge e^c \wedge e^d\,. \end{aligned} \tag{55}$$

На тај начин је добијено Ајнштајн-Картаново дејство за контактну теорију (ECC), са интеракцијом четвртог степена између фермиона, садржане у сабирку $s^a \wedge \star s_a$. Једначине кретања ове теорије су еквивалентне једначинама кретања стандардне Ајнштајн-Картанове теорије и 3BF теорије са везама, али је конфигурациони простор другачији и једнак динамичком конфигурационом простору ових теорија:

$$e^a{}_\mu\,, \phi^A\,, \psi^A\,, \bar{\psi}_A\,, \alpha^\alpha{}_\mu\,. \tag{56}$$

Овим је закључена анализа 3BF дејства са везама, стандардне Ајнштајн-Картанове теорије и Ајнштајн-Картанове теорије са контактном интеракцијом.

2.4 Симетрије тополошког 3BF дејства

Укупна група симетрије тополошког $3BF$ дејства је проучена у [44, 52], и показано је да је њен облик $\mathcal{G}_{3BF} = (\tilde{G} \ltimes (\tilde{H}_L \ltimes (\tilde{N} \times \tilde{M}))) \ltimes HT$. Семидиректни и директни производ група

$\tilde{G}$, $\tilde{H}_L$, $\tilde{M}$, $\tilde{N}$ одговара уобичајеној градијентној симетрији дејства, док HT група одговара такозваној Ено-Таителбом (НТ) симетрији, која је тривијална на једначинама кретања [52].

Избор 3-групе Стандардног Модела, уведене у претходним поглављима, имплицира специјалнију структуру групе градијентне симетрије $\mathcal{G}_{3BF}$. Конкретно, у општем случају, генератори групе $\tilde{H}_L$ могу природно да се поделе на $\hat{H}$-генераторе и $\hat{L}$-генераторе, који задовољавају комутационе релације Лијеве алгебре у облику

$$[\hat{H},\hat{H}] \sim \hat{L}\,, \qquad [\hat{H},\hat{L}] \sim 0\,, \qquad [\hat{L},\hat{L}] \sim 0\,, \tag{57}$$

где су структурне константе у првом комутатору пропорционалне компонентама пресликавања Пајферовог подизања у датој 3-групи [44]. У случају избора тривијалног Пајферовог подизања, комутационе релације постају

$$[\hat{H},\hat{H}] \sim 0\,, \qquad [\hat{H},\hat{L}] \sim 0\,, \qquad [\hat{L},\hat{L}] \sim 0\,. \tag{58}$$

То значи да група $\tilde{H}_L$ има структуру директног производа

$$\tilde{H}_L = \tilde{H} \times \tilde{L} \tag{59}$$

своје две Абелове нормалне подгрупе $\tilde{H}$ и $\tilde{L}$. Додатно, будући да су $\hat{H}$-генератори одговорни за структуру семидиректног производа $\tilde{H}_L \ltimes (\tilde{N}\times\tilde{M})$ унутар $\mathcal{G}_{3BF}$, са комутационим релацијама у облику

$$[\hat{H},\hat{N}] \sim \hat{M}\,, \qquad [\hat{H},\hat{M}] \sim 0\,, \qquad [\hat{L},\hat{M}] \sim 0\,, \qquad [\hat{L},\hat{N}] \sim 0\,, \tag{60}$$

могуће је директно закључити да у случају сваке 3-групе са тривијалним Пајферовим подизањем, одговарајућа група симетрије $\mathcal{G}_{3BF}$ има облик:

$$\mathcal{G}_{3BF} = (\tilde{G} \ltimes (\tilde{L} \times (\tilde{H} \ltimes (\tilde{N} \times \tilde{M})))) \ltimes HT\,. \tag{61}$$

Дакле, (61) представља групу градијентне симетрије $3BF$ теорије засноване на 3-групи Стандардног Модела.

Дејство ове групе на поља присутна у 3BF дејству у случају произвољне 3-групе је дато инфинитезималним тренсформацијама у [44] и представљено као варијација форме, док су инфинитезималне НТ трансформације дефинисане у [52]. У специјалном случају 3-групе Стандардног Модела, обичне градијентне трансформације су дате експлицитно као варијација форме:

$$\begin{array}{rcl}
\delta_0^g \alpha^\alpha &=& -\nabla \epsilon_{\mathfrak{g}}{}^\alpha\,,\\
\delta_0^g \omega^{[ab]} &=& -\nabla \epsilon_{\mathfrak{g}}{}^{[ab]}\,,\\
\delta_0^g \beta^a &=& \rhd_{\alpha b}{}^a \epsilon_{\mathfrak{g}}{}^\alpha \beta^b - \nabla \epsilon_{\mathfrak{h}}{}^a\,,\\
\delta_0^g \gamma^A &=& \rhd_{\alpha B}{}^A \epsilon_{\mathfrak{g}}{}^\alpha \gamma^B + \nabla \epsilon_{\mathfrak{l}}{}^A\,,\\
\delta_0^g B^\alpha &=& f_{\beta\gamma}{}^\alpha \epsilon_{\mathfrak{g}}{}^\beta B^\gamma + e_a \wedge \epsilon_{\mathfrak{h}}{}^b \rhd^\alpha{}_b{}^a - D_A \rhd^\alpha{}_B{}^A \epsilon_{\mathfrak{l}}{}^B - \nabla \epsilon_{\mathfrak{m}}{}^\alpha + \beta_b \rhd^\alpha{}_a{}^b \epsilon_{\mathfrak{n}}{}^a\,,\\
\delta_0^g B^{[ab]} &=& f_{[gh][ij]}{}^{[ab]} B^{[ij]} \epsilon_{\mathfrak{g}}{}^{[gh]} - \nabla \epsilon_{\mathfrak{m}}{}^{[ab]} + e_c \wedge \epsilon_{\mathfrak{h}}{}^d \rhd^{[ab]}{}_d{}^c + \beta_d \rhd^{[ab]}{}_c{}^d \epsilon_{\mathfrak{n}}{}^c - \epsilon_{\mathfrak{l}}{}^A \rhd^{[ab]}{}_A{}^B D_B\,,\\
\delta_0^g e^a &=& -\nabla \epsilon_{\mathfrak{n}}{}^a + \epsilon_{\mathfrak{g}}{}^\alpha \rhd_{\alpha b}{}^a e^b\,,\\
\delta_0^g D^A &=& \rhd_{\alpha B}{}^A \epsilon_{\mathfrak{g}}{}^\alpha D^B\,.
\end{array} \tag{62}$$

Трансформације су задате помоћу пет слободних параметара који одговарају својим Лијевим алгебрама. Параметри $\epsilon_{\mathfrak{g}}{}^\alpha$ и $\epsilon_{\mathfrak{n}}{}^a$ су нула-форме, $\epsilon_{\mathfrak{h}}{}^a$ и $\epsilon_{\mathfrak{m}}{}^\alpha$ су један-форме, и $\epsilon_{\mathfrak{l}}{}^A$ су три-форме.

Инфинитезималне трансформације HT симетрије је најлакше изразити у пратећој матричној форми [52]:

$$
\begin{pmatrix}
\delta_0^{HT} B^{\alpha}{}_{\mu\nu} \\
\delta_0^{HT} C^{a}{}_{\mu} \\
\delta_0^{HT} D^{A} \\
\delta_0^{HT} \alpha^{\alpha}{}_{\mu} \\
\delta_0^{HT} \beta^{a}{}_{\mu\nu} \\
\delta_0^{HT} \gamma^{A}{}_{\mu\nu\rho}
\end{pmatrix}
=
\left(\begin{array}{c|c|c|c|c|c}
\epsilon^{\alpha\beta}{}_{\mu\nu\sigma\lambda} & \epsilon^{\alpha b}{}_{\mu\nu\sigma} & \epsilon^{\alpha B}{}_{\mu\nu} & \epsilon^{\alpha\beta}{}_{\mu\nu\sigma} & \epsilon^{\alpha b}{}_{\mu\nu\sigma\lambda} & \epsilon^{\alpha B}{}_{\mu\nu\sigma\lambda\xi} \\ \hline
\mu^{a\beta}{}_{\mu\sigma\lambda} & \epsilon^{ab}{}_{\mu\sigma} & \epsilon^{aB}{}_{\mu} & \epsilon^{a\beta}{}_{\mu\sigma} & \epsilon^{ab}{}_{\mu\sigma\lambda} & \epsilon^{aB}{}_{\mu\sigma\lambda\xi} \\ \hline
\mu^{A\beta}{}_{\sigma\lambda} & \mu^{Ab}{}_{\sigma} & \epsilon^{AB} & \epsilon^{A\beta}{}_{\sigma} & \epsilon^{Ab}{}_{\sigma\lambda} & \epsilon^{AB}{}_{\sigma\lambda\xi} \\ \hline
\mu^{\alpha\beta}{}_{\mu\sigma\lambda} & \mu^{\alpha b}{}_{\mu\sigma} & \mu^{\alpha B}{}_{\mu} & \epsilon^{\alpha\beta}{}_{\mu\sigma} & \epsilon^{\alpha b}{}_{\mu\sigma\lambda} & \epsilon^{\alpha B}{}_{\mu\sigma\lambda\xi} \\ \hline
\mu^{a\beta}{}_{\mu\nu\sigma\lambda} & \mu^{ab}{}_{\mu\nu\sigma} & \mu^{aB}{}_{\mu\nu} & \mu^{a\beta}{}_{\mu\nu\sigma} & \epsilon^{ab}{}_{\mu\nu\sigma\lambda} & \epsilon^{aB}{}_{\mu\nu\sigma\lambda\xi} \\ \hline
\mu^{A\beta}{}_{\mu\nu\rho\sigma\lambda} & \mu^{Ab}{}_{\mu\nu\rho\sigma} & \mu^{AB}{}_{\mu\nu\rho} & \mu^{A\beta}{}_{\mu\nu\rho\sigma} & \mu^{Ab}{}_{\mu\nu\rho\sigma\lambda} & \epsilon^{AB}{}_{\mu\nu\rho\sigma\lambda\xi}
\end{array}\right)
\begin{pmatrix}
\frac{1}{2}\frac{\delta S}{\delta B^{\beta}{}_{\sigma\lambda}} \\
\frac{\delta S}{\delta C^{b}{}_{\sigma}} \\
\frac{\delta S}{\delta D^{B}} \\
\frac{\delta S}{\delta \alpha^{\beta}{}_{\sigma}} \\
\frac{1}{2}\frac{\delta S}{\delta \beta^{b}{}_{\sigma\lambda}} \\
\frac{1}{3!}\frac{\delta S}{\delta \gamma^{B}{}_{\sigma\lambda\xi}}
\end{pmatrix} . \tag{63}
$$

Да би се осигурала антисиметричност параметара матрице, морају да буду задовољени следећи идентитети:

$$
\begin{gathered}
\mu^{b\alpha}{}_{\sigma\mu\nu} = -\epsilon^{\alpha b}{}_{\mu\nu\sigma} , \qquad \mu^{B\alpha}{}_{\mu\nu} = -\epsilon^{\alpha B}{}_{\mu\nu} , \qquad \mu^{\beta\alpha}{}_{\sigma\mu\nu} = -\epsilon^{\alpha\beta}{}_{\mu\nu\sigma} , \\
\mu^{b\alpha}{}_{\sigma\lambda\mu\nu} = -\epsilon^{\alpha b}{}_{\mu\nu\sigma\lambda} , \qquad \mu^{B\alpha}{}_{\sigma\lambda\xi\mu\nu} = -\epsilon^{\alpha B}{}_{\mu\nu\sigma\lambda\xi} . \\
\mu^{Ba}{}_{\mu} = -\epsilon^{aB}{}_{\mu} , \qquad \mu^{\beta a}{}_{\sigma\mu} = -\epsilon^{a\beta}{}_{\mu\sigma} , \\
\mu^{ba}{}_{\sigma\lambda\mu} = -\epsilon^{ab}{}_{\mu\sigma\lambda} , \qquad \mu^{Ba}{}_{\sigma\lambda\xi\mu} = -\epsilon^{aB}{}_{\mu\sigma\lambda\xi} , \\
\mu^{\beta A}{}_{\sigma} = -\epsilon^{A\beta}{}_{\sigma} , \qquad \mu^{bA}{}_{\sigma\lambda} = -\epsilon^{Ab}{}_{\sigma\lambda} , \qquad \mu^{BA}{}_{\sigma\lambda\xi} = -\epsilon^{AB}{}_{\sigma\lambda\xi} , \\
\mu^{b\alpha}{}_{\sigma\lambda\mu} = -\epsilon^{\alpha b}{}_{\mu\sigma\lambda} , \qquad \mu^{B\alpha}{}_{\sigma\lambda\xi\mu} = -\epsilon^{\alpha B}{}_{\mu\sigma\lambda\xi} , \qquad \mu^{Ba}{}_{\sigma\lambda\xi\mu\nu} = -\epsilon^{aB}{}_{\mu\nu\sigma\lambda\xi} .
\end{gathered} \tag{64}
$$

За више детаља о особинама и значају HT трансформација погледати [52].

2.5 Тополошка инваријанта за квантну 3BF теорију

У процесу конструкције тополошке квантне теорије поља засноване на 3BF дејству, потребно је дефинисати суму по стањима:

$$
Z = \int \mathcal{D}\alpha \mathcal{D}\beta \mathcal{D}\gamma \mathcal{D}B \mathcal{D}C \mathcal{D}D \exp\left(i \int_{\mathcal{M}_4} \langle B \wedge \mathcal{F} \rangle_{\mathfrak{g}} + \langle C \wedge \mathcal{G} \rangle_{\mathfrak{h}} + \langle D \wedge \mathcal{H} \rangle_{\mathfrak{l}} \right) . \tag{65}
$$

Интеграцијом по пољима B, C и D се добија структура суме по стањима у облику интеграла по конексијама:

$$
Z = \mathcal{N} \int \mathcal{D}\alpha \mathcal{D}\beta \mathcal{D}\gamma \delta\left(\mathcal{F}\right) \delta\left(\mathcal{G}\right) \delta\left(\mathcal{H}\right) . \tag{66}
$$

Израчунавање вредности ове суме по стањима је могуће тек након формалног дефинисања мере у интегралу и домена интеграције. Ова мера се дефинише тако што се многострукост издели на ћелије и потом свакој ћелији додели вредност поља које на њој живи, тако да 0-форме живе у вертексима, 1-форме на ивицама, 2-форме на површинама, 3-форме на трозапреминама и 4-форме на четворозапреминама. Особине тополошких теорија су такве да вредност суме по стањима не зависи од начина на који је многострукост издељена на ћелије, односно, сума по стањима је инваријантна на Пахнерове потезе (додавање нових елемената и ћелија).

Најелементарнија подела на ћелије је дата триангулацијом, односно у четвородимензионалном случају дељењем на 4-симплексе. Инваријантна сума по стањима на многострукости издељеној на 4-симплексе је:

$$
\begin{aligned}
Z = & |G|^{-|v|+|l|-|\Delta|}|H|^{|v|-|l|+|\Delta|-|\tau|}|L|^{-|v|+|l|-|\Delta|+|\tau|-|\sigma|} \\
& \times \left(\prod_{(jk)\in\varepsilon} \int_G dg_{jk} \right) \left(\prod_{(jkl)\in\Delta} \int_H dh_{jkl} \right) \left(\prod_{(jklm)\in\tau} \int_L dl_{jklm} \right) \\
& \times \left(\prod_{(jkl)\in\Delta} \delta_G \left(\partial(h_{jkl}) g_{kl} g_{jk} g_{jl}^{-1} \right) \right) \left(\prod_{(jklm)\in\tau} \delta_H \left(\delta(l_{jklm}) h_{jlm} (g_{lm} \triangleright h_{jkl}) h_{klm}^{-1} h_{jkm}^{-1} \right) \right) \\
& \times \prod_{(jklmn)\in\sigma} \delta_L \left(l_{jlmn}^{-1} h_{jln} \triangleright' \{h_{lmn}, (g_{mn} g_{lm}) \triangleright h_{jkl}\}_{\rm pf} l_{jkln}^{-1} (h_{jkn} \triangleright' l_{klmn}) l_{jkmn} h_{jmn} \triangleright' (g_{mn} \triangleright l_{jklm}) \right) ,
\end{aligned}
\tag{67}
$$

где су $|G|$, $|H|$ и $|L|$ број елемената група G, H и L, редом, $|v|$, $|l|$, $|\Delta|$, $|\tau|$ и $|\sigma|$ број вертекса, ивица, троуглова, тетраедара и 4-симплекса редом, и дејство $\triangleright'$ је дефинисано као:

$$
h \triangleright' l = \{h, \delta l\}_{\rm pf} \,, \qquad h \in H \,, \quad l \in L \,. \tag{68}
$$

Пресликавања δ_G, δ_H и δ_L су делте на групама G, H и L [26].

3 Механизми нарушења симетрије 3BF дејства са везама

У наредним поглављима ће бити изложен преглед механизама нарушења симетрије тополошког 3BF дејства, прво експлицитно, додавањем веза, а потом ће бити спроведен поступак спонтаног нарушења симетрије електрослабе интеракције и Хигсов механизам.

На овом месту, пре почетка дискусије о нарушењу симетрије, је важно истаћи једну везу између 3-групе и три билинеарне форме, јер ће ова веза имати значајну улогу у механизму нарушења симетрије и Хигсовом механизму. Специјалан услов који билинеарне форме морају да задовоље јесте да буду G-инваријантне. Ово намеће нетривијалне услове на избор дејства $\triangleright$. Ови услови су одређени теоремом.

Теорема 1. *За дат 2-укрштени модул* $(L \xrightarrow{\delta} H \xrightarrow{\partial} G,\ \triangleright,\ \{\_\,,\_\}_{\mathrm{pf}})$ *и симетричне, недегенерисане билинеарне форме* $\langle \_\,,\_ \rangle_{\mathfrak{g}}$, $\langle \_\,,\_ \rangle_{\mathfrak{h}}$ *и* $\langle \_\,,\_ \rangle_{\mathfrak{l}}$, *важи да ако су билинеарне форме G-инваријантне, онда компоненте дејства* $\triangleright_{\alpha\beta\gamma}$, $\triangleright_{\alpha ab}$ *и* $\triangleright_{\alpha AB}$ *морају да буду антисиметричне на замену другог и трећег индекса. Додатно, постоје избори базиса у Лијевим алгебрама* $\mathfrak{g}$, $\mathfrak{h}$ *и* $\mathfrak{l}$ *такви да* ${\triangleright_{\alpha\beta}}^{\gamma}$, ${\triangleright_{\alpha a}}^{b}$ *и* ${\triangleright_{\alpha A}}^{B}$ *имају нулте елементе на дијагонали у односу на други и трећи индекс, а билинеарне форме су у овом базису дијагоналне.*

За доказ теореме погледати Додатак A.1. Потребно је нагласити да су ове рестрикције на дејство $\triangleright$ последица искључиво G-инваријантности билинеарних форми и да не важе у општем случају.

3.1 Експлицитно нарушење симетрије

Трансформације задате једначинама (62) и (63) чувају облик тополошког 3BF дејства. Међутим, у случају 3BF дејства са везама (22), ово више није испуњено, јер свака од додатих веза може експлицитно да наруши једну или више симетрија тополошког дејства. У циљу одређивања која симетрија је очувана, а која нарушена додавањем које везе, потребно је дефинисати помоћна дејства за сваку везу посебно у облику

$$S = S_{3BF} + S_{\text{constraint}}\,, \tag{69}$$

и варирати ово дејство према једначинама (62). Тополошки део S_{3BF} ће свакако остати непромењен при овим трансформацијама, што имплицира да је услов очуваности целог дејства дат захтевом да је и веза очувана

$$\delta_0^g S_{\text{constraint}} = 0\,. \tag{70}$$

Овај захтев не мора да буде аутоматски задовољен, већ може да се деси да је потребно неке од параметара трансформација ${\epsilon_{\mathfrak{g}}}^{\alpha}$, ${\epsilon_{\mathfrak{h}}}^{a}$, ${\epsilon_{\mathfrak{l}}}^{A}$, ${\epsilon_{\mathfrak{m}}}^{\alpha}$ и ${\epsilon_{\mathfrak{n}}}^{a}$ фиксирати на нулу. Сваки параметар који је потребно фиксирати указује да је одговарајућа подгрупа симетрије нарушена посматраном везом. У наредним одељцима ће бити испитано која веза (24)-(31) нарушава коју подгрупу симетрије, на основу услова (70).

Треба истаћи да је претходни метод заснован на услову (70) употребљив само за испитивање стандардних градијентних симетрија, док се HT симетрија не може проверити на овакав начин. Конкретно, пошто дефиниција HT симетрије (63) експлицитно зависи од облика дејства, додавање веза у дејство мења HT групу симетрије на нетривијалан начин, најчешће тако што *повећава* број њених генератора и параметара. То значи да је нова HT група *већа* од

НТ групе тополошког дејства [52]. Пошто је оваква промена НТ групе у потпуној супротности у односу на промену групе градијентне симетрије, која се нарушава додавањем веза, под нарушењем симетрије додавањем веза се може сматрати само нарушење градијентне групе симетрије.

3.1.1 Веза за гравитационо поље

Као што је претходно објашњено, процедура за анализу нарушења симетрије у случају везе за гравитационо поље се своди на одређивање варијације форме једначине (24) задатим са (62) и потом испитивање испуњености захтева (70). Конкретно, добија се:

$$\delta_0^g S_{\text{grav}} = -\int \left(\delta_0^g \lambda_{[ab]} \wedge \left(B^{[ab]} - \frac{1}{16\pi l_p^2} \varepsilon^{[ab]cd} e_c \wedge e_d \right) + \lambda_{[ab]} \wedge \left(\delta_0^g B^{[ab]} + \frac{1}{8\pi l_p^2} \varepsilon^{[ab]cd} \delta_0^g e_c \wedge e_d \right) \right) . \tag{71}$$

Варијација Лагранжевог множитеља $\lambda_{[ab]}$ није у старту дефинисана, па се његова варијација може изабрати тако да симетрија буде очувана што је више могуће. Заменом (62) у (71), се добија да варијација везе за гравитацију постаје:

$$\begin{aligned} \delta_0^g S_{\text{grav}} &= \int \left(\delta_0^g \lambda_{[ij]} - \lambda_{[i|h} \epsilon_{\mathfrak{g}|j]}{}^h \right) \wedge \left(B^{[ij]} - \frac{1}{16\pi l_p^2} \varepsilon^{[ij]nm} e_n \wedge e_m \right) \\ &+ \lambda_{[ab]} \wedge e_d \wedge \left(\epsilon_{\mathfrak{h}}{}^{[a|} \eta^{|b]d} - \frac{1}{8\pi l_p^2} \varepsilon^{[ab]cd} \eta_{fc} \nabla \epsilon_{\mathfrak{n}}{}^f \right) \\ &+ \lambda_{[ab]} \wedge \left(\epsilon_{\mathfrak{n}}{}^{[a|} \beta^{|b]} - \nabla \epsilon_{\mathfrak{m}}{}^{[ab]} - \epsilon_{\mathfrak{l}}{}^A \rhd^{[ab]}{}_A{}^B D_B \right) , \end{aligned} \tag{72}$$

одакле се може изабрати варијација множитеља $\lambda_{[ab]}$ у облику:

$$\delta_0^g \lambda_{[ij]} = -\lambda_{[ab]} f_{[gh][ij]}{}^{[ab]} \epsilon_{\mathfrak{g}}{}^{[gh]} = \lambda_{[i|h} \epsilon_{\mathfrak{g}|j]}{}^h . \tag{73}$$

Овај избор уклања цео први ред у једначини (72). Међутим, преостала два реда указују да се захтев (70) може испунити само фиксирањем вредности параметара $\epsilon_{\mathfrak{h}}{}^a$, $\epsilon_{\mathfrak{n}}{}^a$, $\epsilon_{\mathfrak{m}}{}^\alpha$ и $\epsilon_{\mathfrak{l}}{}^A$. Једини параметар који остаје слободан је $\epsilon_{\mathfrak{g}}{}^\alpha$. Ово имплицира да веза нарушава све подгрупе симетрије $\tilde{M}$, $\tilde{N}$, $\tilde{L}$, и $\tilde{H}$, осим $\tilde{G}$, која је у потпуности очувана.

3.1.2 Веза за скаларна поља

Коришћењем претходне процедуре се могу испитати и све остале везе. У случају везе за скаларна поља се добија:

$$\begin{aligned} \delta_0^g S_{\text{scal}} &= \int \Big[\delta_0^g \tilde{\lambda}_A \wedge \left(\tilde{\gamma}^A - H_{abc}{}^A e^a \wedge e^b \wedge e^c \right) \\ &+ \tilde{\lambda}_A \wedge \left(\delta_0^g \tilde{\gamma}^A - \delta_0^g H_{abc}{}^A e^a \wedge e^b \wedge e^c - 3 H_{abc}{}^A \delta_0^g e^a \wedge e^b \wedge e^c \right) \\ &+ \delta_0^g \Lambda^{ab}{}_A \wedge \left(H_{abc}{}^A \varepsilon^{cdef} e_d \wedge e_e \wedge e_f - \nabla \phi^A \wedge e_a \wedge e_b \right) \\ &+ \Lambda^{ab}{}_A \wedge \Big(\delta_0^g H_{abc}{}^A \varepsilon^{cdef} e_d \wedge e_e \wedge e_f + 3 H_{abc}{}^A \varepsilon^{cdef} \delta_0^g e_d \wedge e_e \wedge e_f \\ &- \nabla \delta_0^g \phi^A \wedge e_a \wedge e_b - \delta_0^g \alpha^{[kl]} \rhd_{[kl]B}{}^A \phi^B \wedge e_a \wedge e_b - 2 \nabla \phi^A \wedge \delta_0^g e_a \wedge e_b \Big) \Big] . \end{aligned} \tag{74}$$

Замена (62) у (74) даје:

$$
\begin{aligned}
\delta_0^g S_{\rm scal} \;=\; & \int \Big[(\delta_0^g \tilde{\lambda}_A + \tilde{\lambda}_B \epsilon_{\mathfrak{g}}{}^{[ij]} \rhd_{[ij]A}{}^B) \wedge \big(\tilde{\gamma}^A - H_{abc}{}^A e^a \wedge e^b \wedge e^c\big) \\
- & \; \big(\delta_0^g H_{abc}{}^A - \epsilon_{\mathfrak{g}}{}^{[ij]} H_{abc}{}^B \rhd_{[ij]B}{}^A\big) \, (\tilde{\lambda}_A \eta^{ad}\eta^{be}\eta^{cf} - \Lambda^{ab}{}_A \varepsilon^{cdef}) \wedge e_d \wedge e_e \wedge e_f \\
+ & \; (\delta_0^g \Lambda^{ab}{}_A + \Lambda^{ab}{}_B \epsilon_{\mathfrak{g}}{}^{[cd]} \rhd_{[cd]A}{}^B) \wedge \big(H_{abc}{}^A \varepsilon^{cdef} e_d \wedge e_e \wedge e_f - \nabla\phi^A \wedge e_a \wedge e_b\big) \qquad (75) \\
+ & \; (\tilde{\lambda}_A \eta^{ad}\eta^{be}\eta^{cf} - \Lambda^{ab}{}_A \varepsilon^{cdef}) \wedge 3 H_{abc}{}^A (\nabla \epsilon_{\mathfrak{n}d}) \wedge e_e \wedge e_f \\
+ & \; 2\Lambda^{ab}{}_A \wedge \nabla\phi^A \wedge \nabla\epsilon_{\mathfrak{n}a} \wedge e_b + \tilde{\lambda}_A \wedge \nabla \epsilon_{\mathfrak{l}}{}^A \Big] \,.
\end{aligned}
$$

Очигледно је да из четвртог и петог реда следи да ова веза нарушава само подгрупе $\tilde{N}$ и $\tilde{L}$, док $\tilde{H}$ и $\tilde{M}$ симетрије остају очуване јер се њихови параметри $\epsilon_{\mathfrak{h}}{}^a$ и $\epsilon_{\mathfrak{m}}{}^\alpha$ и не појављују у једначини за варијацију (75). Коначно, у циљу очувања $\tilde{G}$ симетрије, могу се изабрати дефиниције варијација нових множитеља:

$$
\delta_0^g \tilde{\lambda}^A = \epsilon_{\mathfrak{g}}{}^{[ij]} \tilde{\lambda}^B \rhd_{[ij]B}{}^A \,, \qquad \delta_0^g H_{abc}{}^A = H_{abc}{}^B \epsilon_{\mathfrak{g}}{}^{[ij]} \rhd_{[ij]B}{}^A \,, \qquad \delta_0^g \Lambda^{abA} = \epsilon_{\mathfrak{g}}{}^{[cd]} \Lambda^{abB} \rhd_{[cd]B}{}^A \,. \qquad (76)
$$

3.1.3 Веза за Диракова поља

На исти начин, варијација везе за Диракова поља даје:

$$
\begin{aligned}
\delta_0^g S_{\rm Dirac} \;=\; & \int (\delta_0^g \bar{\lambda}_A) \wedge \left(\gamma^A + \frac{i}{6} \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \left(\gamma^d \psi\right)^A \right) \\
- & \; (\delta_0^g \lambda^A) \wedge \left(\bar{\gamma}_A - \frac{i}{6} \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \left(\bar{\psi}\gamma^d\right)_A \right) \\
+ & \; \bar{\lambda}_A \wedge \left((\delta_0^g \gamma^A) + \frac{i}{6} \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \left(\gamma^d (\delta_0^g \psi)\right)^A + \frac{i}{2} \varepsilon_{abcd} (\delta_0^g e^a) \wedge e^b \wedge e^c \left(\gamma^d \psi\right)^A \right) \\
- & \; \lambda^A \wedge \left((\delta_0^g \bar{\gamma}_A) - \frac{i}{6} \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \left(\gamma^d (\delta_0^g \bar{\psi})\right)_A - \frac{i}{2} \varepsilon_{abcd} (\delta_0^g e^a) \wedge e^b \wedge e^c \left(\gamma^d \bar{\psi}\right)_A \right) . \qquad (77)
\end{aligned}
$$

Заменом (62) у (77) се добија:

$$
\begin{aligned}
\delta_0^g S_{\rm Dirac} \;=\; & \int (\delta_0^g \bar{\lambda}_A + \epsilon_{\mathfrak{g}}{}^\alpha \rhd_{\alpha A}{}^B \bar{\lambda}_B) \wedge \left(\gamma^A + \frac{i}{6} \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \left(\gamma^d \psi\right)^A \right) \\
- & \; (\delta_0^g \lambda^A + \epsilon_{\mathfrak{g}}{}^\alpha \rhd_\alpha{}^A{}_B \lambda^B) \wedge \left(\bar{\gamma}_A - \frac{i}{6} \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \left(\bar{\psi}\gamma^d\right)_A \right) \\
+ & \; \bar{\lambda}_A \wedge \left(\nabla \epsilon_{\mathfrak{l}}{}^A + \frac{i}{2} \varepsilon_{abcd} (\nabla \epsilon_{\mathfrak{n}}{}^a) \wedge e^b \wedge e^c \left(\gamma^d \psi\right)^A \right) \\
- & \; \lambda^A \wedge \left(\nabla \bar{\epsilon}_{\mathfrak{l}A} - \frac{i}{2} \varepsilon_{abcd} (\nabla \epsilon_{\mathfrak{n}}{}^a) \wedge e^b \wedge e^c \left(\bar{\psi}\gamma^d\right)_A \right) . \qquad (78)
\end{aligned}
$$

Ова веза такође нарушава само $\tilde{N}$ и $\tilde{L}$ симетрије, слично као и веза за скаларна поља. Варијација нових множитеља се може изабрати да буде:

$$
\delta_0^g \bar{\lambda}_A = \epsilon_{\mathfrak{g}}{}^\alpha \rhd_\alpha{}^B{}_A \bar{\lambda}_B \,, \qquad \delta_0^g \lambda^A = \epsilon_{\mathfrak{g}}{}^\alpha \rhd_\alpha{}^A{}_B \lambda^B \,. \qquad (79)
$$

3.1.4 Веза за Јанг-Милсова поља

Веза за Јанг-Милсова поља је слична вези за гравитационо поље, али садржи више Лагранжевих множитеља. Применом исте процедуре се добија:

$$
\begin{aligned}
\delta_0^g S_{\text{Yang-Mills}} &= \int \delta_0^g \lambda^\alpha \wedge \left(B_\alpha - 12C^{\alpha\beta} M_{\beta ab} e^a \wedge e^b\right) \\
&+ \lambda^\alpha \wedge \left(\delta_0^g B_\alpha - 12C^{\alpha\beta} \delta_0^g M_{\beta ab} e^a \wedge e^b - 24C^{\alpha\beta} M_{\beta ab} \delta_0^g e^a \wedge e^b\right) \\
&+ \delta_0^g \zeta^{\alpha ab} \left(M_{\alpha ab} \varepsilon_{cdef} e^c \wedge e^d \wedge e^e \wedge e^f - F_\alpha \wedge e_a \wedge e_b\right) \\
&+ \zeta^{\alpha ab} \left((\delta_0^g M_{\alpha ab}) \varepsilon_{cdef} e^c \wedge e^d \wedge e^e \wedge e^f + 4M_{\alpha ab} \varepsilon_{cdef} (\delta_0^g e^c) \wedge e^d \wedge e^e \wedge e^f\right. \\
&- \left.(\delta_0^g F_\alpha) \wedge e_a \wedge e_b - 2F_\alpha \wedge (\delta_0^g e_a) \wedge e_b\right) ,
\end{aligned}
\tag{80}
$$

где је

$$
\delta_0^g F_\alpha = \epsilon_{\mathfrak{g}}{}^\beta F^\gamma \rhd_{\alpha\beta\gamma} \ . \tag{81}
$$

Варијација јачине поља (81) је добијена по дефиницији (7) коришћењем (62). Комбинација једначина (62), (80) и (81) даје:

$$
\begin{aligned}
\delta_0^g S_{\text{Yang-Mills}} &= \int \left(\delta_0^g \lambda^\alpha + \lambda^\gamma \epsilon_{\mathfrak{g}}{}^\beta \rhd_{\gamma\beta}{}^\alpha\right) \wedge \left(B_\alpha - 12C^{\alpha\beta} M_{\beta ab} e^a \wedge e^b\right) \\
&+ \left(\delta_0^g M_{\alpha ab} - \epsilon_{\mathfrak{g}}{}^\beta M^\gamma{}_{ab} \rhd_{\alpha\beta\gamma}\right) \left(\zeta^{\alpha ab} \varepsilon_{cdef} e^c \wedge e^d \wedge e^e \wedge e^f - 12C^\alpha{}_\beta \lambda^\beta \wedge e^a \wedge e^b\right) \\
&+ \left(\delta_0^g \zeta^{\alpha ab} + \rhd_{\gamma\beta}{}^\alpha \zeta^{\gamma ab} \epsilon_{\mathfrak{g}}{}^\beta\right) \left(M_{\alpha ab} \varepsilon_{cdef} e^c \wedge e^d \wedge e^e \wedge e^f - F_\alpha \wedge e_a \wedge e_b\right) \\
&+ \lambda^\alpha \wedge \left(-\nabla \epsilon_{\mathfrak{m}\alpha} - \epsilon_{\mathfrak{l}}{}^A \rhd_{\alpha A}{}^B D_B + \epsilon_{\mathfrak{n}}{}^a \rhd_{\alpha a}{}^b \beta_b\right. \\
&- \left.\epsilon_{\mathfrak{h}}{}^a \wedge e_b \rhd_{\alpha a}{}^b + 24C^{\alpha\beta} M_{\beta ab} (\nabla \epsilon_{\mathfrak{n}}{}^a) \wedge e^b\right) \\
&+ \zeta^{\alpha ab} \left(-4M_{\alpha ab} \varepsilon_{cdef} (\nabla \epsilon_{\mathfrak{n}}{}^c) \wedge e^d \wedge e^e \wedge e^f + 2F_\alpha \wedge (\nabla \epsilon_{\mathfrak{n}a}) \wedge e_b\right) .
\end{aligned}
\tag{82}
$$

Слично као и у случају везе за гравитационо поље, све симетрије $\tilde{H}$, $\tilde{L}$, $\tilde{N}$ и $\tilde{M}$, осим $\tilde{G}$, су нарушене. Варијација нових множитеља се може изабрати да буде:

$$
\delta_0^g \lambda^\alpha = \epsilon_{\mathfrak{g}}{}^\beta \lambda^\gamma \rhd_{\beta\gamma}{}^\alpha , \qquad \delta_0^g M_{\alpha ab} = \epsilon_{\mathfrak{g}}{}^\beta M^\gamma{}_{ab} \rhd_{\alpha\beta\gamma} , \qquad \delta_0^g \zeta^{\alpha ab} = \epsilon_{\mathfrak{g}}{}^\beta \zeta^{\gamma ab} \rhd_{\beta\gamma}{}^\alpha . \tag{83}
$$

3.1.5 Везе за Хигсов и Јукавин потенцијал, спинска и веза са космолошком константом

Варијација везе за Хигсов потенцијал даје

$$
\begin{aligned}
\delta_0^g S_{\text{Higgs}} &= -\frac{1}{3}\chi \int 2(\phi_A \phi^A - v^2) \phi_A (\delta_0^g \phi^A) \varepsilon^{abcd} e_a \wedge e_b \wedge e_c \wedge e_d \\
&+ (\phi_A \phi^A - v^2)^2 \varepsilon^{abcd} (\delta_0^g e_a) \wedge e_b \wedge e_c \wedge e_d .
\end{aligned}
\tag{84}
$$

Одатле следи

$$\delta_0^g S_{\rm Higgs} = \frac{1}{3}\chi \int (\phi_A\phi^A - v^2)^2 \varepsilon^{abcd} (\nabla \epsilon_{\mathfrak{n}a}) \wedge e_b \wedge e_c \wedge e_d \,, \tag{85}$$

након коришћења идентитета $\triangleright_{\alpha AB} = -\triangleright_{\alpha BA}$ из Теореме 1. Ова веза не нарушава $\tilde{G}$ симетрију јер њен параметар нестаје из једначине (85), иако је иницијално био присутан у дефиницији варијација ϕ^A и e_a, па је нарушена само $\tilde{N}$ симетрија.

Варијација везе за Јукавин потенцијал је:

$$\begin{aligned} \delta_0^g S_{\rm Yukawa} &= -\frac{2}{4!}\int Y_{ABC}\delta_0^g(\bar{\psi}^A\psi^B)\phi^C \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d \\ &+ Y_{ABC}\bar{\psi}^A\psi^B \delta_0^g\phi^C \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d + 4Y_{ABC}\bar{\psi}^A\psi^B\phi^C \varepsilon_{abcd} \delta_0^g e^a \wedge e^b \wedge e^c \wedge e^d \,. \end{aligned} \tag{86}$$

Поново, заменом варијације поља (62) у (86) се добија:

$$\delta_0^g S_{\rm Yukawa} = \frac{1}{3}\int Y_{ABC}\bar{\psi}^A\psi^B\phi^C \varepsilon_{abcd} (\nabla \epsilon_{\mathfrak{n}}{}^a) \wedge e^b \wedge e^c \wedge e^d \,. \tag{87}$$

Као и у претходном случају, само је $\tilde{N}$ симетрија нарушена, зато што је Y_{ABC} матрица дефинисана на такав начин да чува $\tilde{G}$ симетрију.

Спинска веза не нарушава $\tilde{G}$ симетрију из истих разлога као и претходне две везе, али нарушава $\tilde{N}$ и $\tilde{H}$ симетрије:

$$\begin{aligned} \delta_0^g S_{\rm spin} &= 2\pi i l_p^2 \varepsilon_{abcd} \int \left(\delta_0^g(\bar{\psi}\gamma_5\gamma^a\psi) e^b \wedge e^c \wedge \beta^d + \bar{\psi}\gamma_5\gamma^a\psi \delta_0^g(e^b \wedge e^c \wedge \beta^d) \right) \,, \\ &= -2\pi i l_p^2 \varepsilon_{abcd} \int \bar{\psi}\gamma_5\gamma^a\psi \left(2(\nabla \epsilon_{\mathfrak{n}}{}^b) \wedge e^c \wedge \beta^d + e^b \wedge e^c \wedge (\nabla \epsilon_{\mathfrak{h}}{}^d) \right) \,. \end{aligned} \tag{88}$$

Коначно, веза за космолошку константу нарушава само $\tilde{N}$ симетрију:

$$\delta_0^g S_{\rm CC} = -\int \frac{1}{24\pi l_p^2} \Lambda \varepsilon_{abcd} \delta_0^g e^a \wedge e^b \wedge e^c \wedge e^d = -\int \frac{1}{24\pi l_p^2} \Lambda \varepsilon_{abcd} \nabla \epsilon_{\mathfrak{n}}{}^a \wedge e^b \wedge e^c \wedge e^d \,. \tag{89}$$

3.1.6 Преглед нарушења симетрије

У циљу сумирања резултата у овом одељку, може се конструисати табела симетрија и веза. Свако поље у табели означено са × одговара постојању нарушења дате симетрије датом везом:

	$S_{\rm grav}$	$S_{\rm scal}$	$S_{\rm Dirac}$	$S_{\rm Yang\text{-}Mills}$	$S_{\rm Higgs}$	$S_{\rm Yukawa}$	$S_{\rm spin}$	$S_{\rm CC}$
$\tilde{G}$								
$\tilde{H}$	×			×			×	
$\tilde{L}$	×	×	×	×				
$\tilde{M}$	×			×				
$\tilde{N}$	×	×	×	×	×	×	×	×

Из ове табеле је могуће закључити неколико особина. Прва, $\tilde{G}$ симетрија је очувана у свим везама, док је $\tilde{N}$ симетрија нарушена свим везама. Друга, гравитациона и Јанг-Милсова веза нарушавају све симетрије осим $\tilde{G}$, и ове везе су једине такве. Коначно, везе за Хигсов и Јукавин потенцијал, као и веза за космолошку константу нарушавају само $\tilde{N}$ симетрију.

Додатно се, поред ових веза, могу испитати и везе које се не појављују у дејству (22) али ће се појавити у следећем поглављу о спонтаном нарушењу симетрије. Прва од ових веза је веза за масу скаларних поља:

$$S_{\text{scalar mass}} = -\frac{m^2}{4!}\varepsilon_{abcd}\int \phi_A\phi^A e^a\wedge e^b\wedge e^c\wedge e^d\,. \tag{90}$$

Варијација ове везе је

$$\delta_0^g S_{\text{scalar mass}} \;=\; -\frac{m^2}{4!}\varepsilon_{abcd}\int \left(2(\delta_0^g\phi_A)\phi^A e^a\wedge e^b\wedge e^c\wedge e^d + 4\phi_A\phi^A(\delta_0^g e^a)\wedge e^b\wedge e^c\wedge e^d\right)\,, \tag{91}$$

односно:

$$\delta_0^g S_{\text{scalar mass}} = \frac{m^2}{3!}\varepsilon_{abcd}\int \phi_A\phi^A(\nabla\epsilon_{\mathfrak{n}}{}^a)\wedge e^b\wedge e^c\wedge e^d\,. \tag{92}$$

Одатле се може закључити да веза (90) нарушава само $\tilde{N}$ симетрију.

Друга нова веза је веза за масу Дираковил поља:

$$S_{\text{Dirac mass}} = -\frac{m}{12}\varepsilon_{abcd}\int \bar{\psi}_A\psi^A e^a\wedge e^b\wedge e^c\wedge e^d\,. \tag{93}$$

Њена варијација је

$$\begin{aligned}\delta_0^g S_{\text{Dirac mass}} \;&=\; -\frac{m}{12}\varepsilon_{abcd}\int \Big((\delta_0^g\bar{\psi}_A)\psi^A e^a\wedge e^b\wedge e^c\wedge e^d + \bar{\psi}_A(\delta_0^g\psi^A)e^a\wedge e^b\wedge e^c\wedge e^d\\ &+\; 4\bar{\psi}_A\psi^A(\delta_0^g e^a)\wedge e^b\wedge e^c\wedge e^d\Big)\,,\end{aligned} \tag{94}$$

што је једнако:

$$\delta_0^g S_{\text{Dirac mass}} = \frac{m}{3}\varepsilon_{abcd}\int \bar{\psi}_A\psi^A(\nabla\epsilon_{\mathfrak{n}}{}^a)\wedge e^b\wedge e^c\wedge e^d\,. \tag{95}$$

Слично као у случају масе скаларних поља, и ова веза нарушава само $\tilde{N}$ симетрију.

Трећа нова веза је Прокина веза. Ова веза се експлицитно појављује у следећем поглављу у једначини (103), као део дискусије дејства за Прокино поље. Облик ове везе је:

$$S_{\text{Proca}} = \int \Theta^{\alpha ab}\wedge\left(\Xi_{\alpha abc}\varepsilon^{cdef}e_d\wedge e_e\wedge e_f + \frac{M}{g}\alpha_\alpha\wedge e_a\wedge e_b\right) + \frac{M}{g}\alpha^\alpha\wedge\Xi_{\alpha abc}e^a\wedge e^b\wedge e^c\,. \tag{96}$$

Варијација везе даје:

$$\begin{aligned}\delta_0^g S_{\text{Proca}} \;&=\; \int\Bigg[\frac{M}{g}\left(\delta_0^g\alpha^\alpha\wedge\Xi_{\alpha abc}e^a + \alpha^\alpha\wedge\delta_0^g\Xi_{\alpha abc}e^a + 3\alpha^\alpha\wedge\Xi_{\alpha abc}\delta_0^g e^a\right)\wedge e^b\wedge e^c\\ &+\; \delta_0^g\Theta^{\alpha ab}\wedge\left(\Xi_{\alpha abc}\varepsilon^{cdef}e_d\wedge e_e\wedge e_f + \frac{M}{g}\alpha_\alpha\wedge e_a\wedge e_b\right)\\ &+\; \Theta^{\alpha ab}\wedge\left(\delta_0^g\Xi_{\alpha abc}\varepsilon^{cdef}e_d\wedge e_d\wedge e_e\wedge e_f + 3\Xi_{\alpha abc}\varepsilon^{cdef}\delta_0^g e_d\wedge e_e\wedge e_f\right)\\ &+\; \frac{M}{g}\Theta^{\alpha ab}\wedge\left(\delta_0^g\alpha_\alpha\wedge e_a\wedge e_b + 2\alpha_\alpha\wedge\delta_0^g e_a\wedge e_b\right)\Bigg]\,.\end{aligned} \tag{97}$$

Заменом варијације конексије α и поља тетрада, као и коришћењем чињенице да је $\triangleright_{\alpha a}{}^{b} = 0$, добија се:

$$
\begin{aligned}
\delta_0^g S_{\text{Proca}} &= \int \Bigg[\delta_0^g \Theta^{\alpha ab} \wedge \left(\Xi_{\alpha abc} \varepsilon^{cdef} e_d \wedge e_e \wedge e_f + \frac{M}{g} \alpha_\alpha \wedge e_a \wedge e_b \right) \\
&+ \delta_0^g \, \Xi_{\alpha abc} \left(\frac{M}{g} \alpha^\alpha \wedge e^a \wedge e^b \wedge e^c + \Theta^{\alpha ab} \wedge \varepsilon^{cdef} e_d \wedge e_e \wedge e_f \right) \\
&+ \frac{M}{g} \nabla \epsilon_{\mathfrak{g}}{}^{\alpha} \wedge \left(\Theta^{\alpha ab} \wedge e_a \wedge e_b - \Xi_{\alpha abc} e^a \wedge e^b \wedge e^c \right) \\
&- 3\, \Xi_{\alpha abc} \left(\frac{M}{g} \alpha^\alpha \wedge \nabla \epsilon_{\mathfrak{n}}{}^{a} \wedge e^b \wedge e^c + \Theta^{\alpha ab} \varepsilon^{cdef} \nabla \epsilon_{\mathfrak{n}d} \wedge e_e \wedge e_f \right) \\
&- 2 \frac{M}{g} \Theta^{\alpha ab} \wedge \alpha_\alpha \wedge \nabla \epsilon_{\mathfrak{n}a} \wedge e_b \Bigg] .
\end{aligned} \tag{98}
$$

Сада је могуће елиминисати прва два реда из једначине избором варијације множитеља:

$$
\delta_0^g \Theta^{\alpha ab} = 0 \,, \qquad \delta_0^g \, \Xi_{\alpha abc} = 0 \,. \tag{99}
$$

Међутим, поред нарушења $\tilde{N}$ симетрије, Прокина веза је једина која нарушава и $\tilde{G}$ симетрију, јер се њен параметар експлицитно појављује у трећем реду.

Коначно, могуће је конструисати сличну табелу и за ове три везе:

	$S_{\text{scalar mass}}$	$S_{\text{Dirac mass}}$	S_{Proca}
$\tilde{G}$			$\times$
$\tilde{H}$			
$\tilde{L}$			
$\tilde{M}$			
$\tilde{N}$	$\times$	$\times$	$\times$

Овиме је завршена анализа експлицитног нарушења симетрије 3BF дејства са везама.

3.2 Спонтано нарушење симетрије и Хигсов механизам

Уобичајене формулације дејства за Стандардни Модел елементарних честица не укључују 3BF дејство са везама, већ су најчешће изражене помоћу Лагранжијана састављених од обичних тензора. Потом се спровођењем низа корака по угледу на Хигсов механизам, Лагранжијан преписује у облик у ком градијентна симетрија није експлицитно изражена. Додатна претпоставка је да вакуумско стање делимично нарушава ову градијентну симетрију на њену подгрупу и одговара делимичном фиксирању калибрације Лагранжијана, што доводи до спонтаног нарушења његове симетрије.

Сада је у контексту 3BF теорије са везама природно поставити питање да ли се Хигсов механизам може спровести над дејством (22), које представља Стандардни Модел изражен у другом облику, и репродуковати исти резултати као у случају стандардне дефиниције теорије.

3.2.1 3BF дејство са везама за Прокино поље

У циљу проучавања теорије електрослабих интеракција, спонтаног нарушења симетрије и Хигсовог механизма у контексту виших градијентних теорија је важно конструисати дејство за Прокино поље као 3BF дејство са везама, јер ће Хигсов механизам да генерише масене чланове у дејству за градијентна поља. Формулација 3BF дејства са везама за Прокино поље и његово проучавање ће касније олакшати анализу Хигсовог механизма.

У процесу дефинисања 3BF дејства са везама за Прокино поље, прво је потребно изабрати конкретну 3-групу. Уобичајен избор за три Лијеве групе које чине структуру 3-групе је:

$$G = SO(3,1) \times SU(N)\,, \qquad H = \mathbb{R}^4\,, \qquad L = \{\mathbb{1}_L\}\,. \tag{100}$$

Овај избор одговара $SU(N)$ Јанг-Милсовом пољу спрегнутим са Ајнштајн-Картановом гравитацијом без скаларних поља и фермиона (пошто је L група тривијална). Избор тривијалне L групе повлачи за собом услов да су и Пајферово подизање и хомоморфизам δ тривијални, као и дејство $\triangleright$ групе G на групу L. Преостаје да се дефинише хомоморфизам ∂ и дејство $\triangleright$ групе G на саму себе и на групу H. Као и у случају 3-групе за Стандардни Модел, хомоморфизам ∂ се бира да буде тривијалан, а дејство $\triangleright$ је дефинисано на следећи начин. Дејство групе G на саму себе је дато једначинама (13) и (14), слично као у случају Стандардног Модела, док је дејство групе G на групу H такође дато једначинама (15).

Такође, да би дејство било дефинисано, потребно је изабрати конкретне симетричне, недегенерисане, инваријантне, билинеарне форме $\langle\_\,,\_\rangle_{\mathfrak{g}}$ и $\langle\_\,,\_\rangle_{\mathfrak{h}}$. Њихов избор је одређен једначинама (17) и (18), редом, док је $\langle\_\,,\_\rangle_{\mathfrak{l}}$ тривијална. Ови избори поједностављују 3BF дејство тако да оно постаје 2BF дејство, специјалан случај (23), дато са:

$$S_{2BF} = \int B_\alpha \wedge F^\alpha + B^{[ab]} \wedge R_{[ab]} + e_a \wedge \nabla \beta^a\,. \tag{101}$$

Овде је први члан BF дејство за $SU(N)$ Јанг-Милсово поље, док преостала два сабирка одговарају гравитационом пољу.

Након дефинисања тополошког дејства теорије, потребно је одредити везе којима се оно деформише тако да даје добре једначине кретања за Прокино и гравитационо поље. Веза за гравитационо поље је поново дата једначином (24), док се за Прокино поље морају додати везе за Јанг-Милсово поље (27), као и додатна веза, звана Прокина веза:

$$S = S_{2BF} + S_{\text{grav}} + S_{\text{Yang-Mills}} + S_{\text{Proca}}\,. \tag{102}$$

Прокина веза има следећи облик:

$$S_{\text{Proca}} = \int \Theta^{\alpha ab} \wedge \left(\Xi_{\alpha abc} \varepsilon^{cdef} e_d \wedge e_e \wedge e_f + \frac{M}{g} \alpha_\alpha \wedge e_a \wedge e_b \right) + \frac{M}{g} \alpha^\alpha \wedge \Xi_{\alpha abc} e^a \wedge e^b \wedge e^c\,, \tag{103}$$

где су 1-форма $\Theta^{\alpha ab}$ и 0-форма $\Xi_{\alpha abc}$ нови Лагранжеви множитељи, M је нови параметар, док је g константа интеракције Јанг-Милсовог поља, која одговара конкретном избору билинеарне форме интеракционих константи у (27) као:

$$C_{\alpha\beta} = \frac{1}{g^2} g_{\alpha\beta}\,. \tag{104}$$

У циљу провере да дејство (102) заиста одговара Прокином пољу, потребно је одредити једначине кретања. Слично као у случају дејства за Стандардни Модел, варијација дејства по Лагранжевим множитељима даје алгебарске једначине које се могу решити по множитељима,

$$
\begin{aligned}
M_{\alpha ab} &= -\frac{1}{48}\varepsilon_{abcd}e^c{}_\mu e^d{}_\nu F_\alpha{}^{\mu\nu}\,, & \lambda_{\alpha\mu\nu} &= -F_{\alpha\mu\nu}\,, & \zeta^{\alpha ab} &= \frac{1}{4g^2}\varepsilon^{abcd}e_{c\mu}e_{d\nu}F^{\alpha\mu\nu}\,,\\
\Theta^{\alpha ab}{}_\mu &= \frac{M}{6g}\varepsilon^{abcd}\alpha^\alpha{}_\nu e_c{}^\nu e_{d\mu}\,, & \lambda_{[ab]\mu\nu} &= R_{[ab]\mu\nu}\,, & \Xi_{\alpha abc} &= \frac{M}{6g}\varepsilon_{abcd}\alpha_{\alpha\mu}e^{d\mu}\,,\\
B_{\alpha\mu\nu} &= -\frac{e}{2g^2}\varepsilon_{\mu\nu\rho\sigma}F_\alpha{}^{\rho\sigma}\,, & \beta^a{}_{\mu\nu} &= 0\,, & B_{[ab]\mu\nu} &= \frac{1}{8\pi l_p^2}\varepsilon_{abcd}e^c{}_\mu e^d{}_\nu\,,
\end{aligned}
\tag{105}
$$

потом, Ајнштајнову једначину (41) за тензор енергије-импулса који је једнак:

$$
T_{\mu\nu} = \frac{1}{g^2}\left(F^\alpha{}_{\mu\rho}F_{\alpha\nu}{}^\rho - \frac{1}{4}g_{\mu\nu}F^\alpha{}_{\rho\sigma}F_\alpha{}^{\rho\sigma}\right) + \frac{M^2}{g^2}\left(\alpha_{\alpha\mu}\alpha^\alpha{}_\nu - \frac{1}{2}g_{\mu\nu}\alpha_\alpha{}^\rho\alpha^\alpha{}_\rho\right)\,, \tag{106}
$$

као и једначину за спинску конексију (33) и једначину за торзију $T^a \equiv \nabla e^a = 0$, и коначно, једначину кретања за поље векторских бозона

$$
\nabla_\mu F^{\alpha\mu}{}_\nu - M^2\alpha^\alpha{}_\nu = 0\,, \tag{107}
$$

где је $F^\alpha{}_{\mu\nu}$ стандардни тензор јачине Јанг-Милсовог поља за $SU(N)$ конексију $\alpha^\alpha{}_\mu$. Ово је управо једначина кретања за Прокино поље масе M.

Додатно се осим решавањем једначина кретања може демонстрирати да дејство (102) одговара дејству за Прокино поље елиминацијом свих помоћних поља. Пошто су помоћна поља на једначинама кретања задата алгебарским једначинама динамичких поља, њихове једначине кретања се могу заменити натраг у дејство чиме се добија теорија другог реда. Конкретно, заменом свих једначина за помоћна поља (105) у (102), се добија управо стандардно дејство за Прокино поље спрегнуто са Ајнштајн-Картановом гравитацијом:

$$
S = \int \frac{1}{16\pi l_p^2}\varepsilon^{abcd}\,R_{ab}\wedge e_c\wedge e_d - \frac{1}{g^2}F_\alpha\wedge\star F^\alpha - \frac{1}{4!}\frac{M^2}{g^2}\alpha^\alpha{}_\mu\alpha_\alpha{}^\mu\,\varepsilon^{abcd}\,e_a\wedge e_b\wedge e_c\wedge e_d\,. \tag{108}
$$

Овде је $\alpha_\alpha{}^\mu \equiv \alpha_{\alpha\nu}g^{\mu\nu}$ и $g^{\mu\nu} = \eta^{ab}e_a{}^\mu e_b{}^\nu$. Такође, $\star F$ означава Хоџов дуал 2-форме F:

$$
\star F^\alpha = \frac{1}{4}F^\alpha{}_{cd}\varepsilon^{abcd}e_a\wedge e_b\,. \tag{109}
$$

Електрослабе интеракције у Стандардном Моделу садрже више Прокиних поља са различитим масама M. Да би овакав случај био обухваћен у теорији, потребно је уопштити дејство (102) тако да садржи више Прокиних поља. Ово уопштење се може извршити избором 3-групе са модификованом групом G у облику:

$$
G = SO(3,1)\times\prod_i U(1)\times\prod_j SU(N_j)\,. \tag{110}
$$

У односу на претходни избор три групе (100), може се уочити да је подгрупа $SU(N)$ у G замењена вишеструким копијама $U(1)$ и $SU(N_j)$ групе, у зависности од тога колико је Прокиних поља садржано у теорији. Структура 3-групе остаје непромењена, у смислу да је дејство ▷

само уопштено са $SU(N)$ случаја на стандардан начин, тако да једначине (13), (14) и (15) остају задовољене за општи избор (110).

Чак и у случају овог општијег избора 3-групе, дејство теорије и даље има исти облик као у једначини (102), али сада чланови S_{2BF} и $S_{\text{Yang-Mills}}$ одговарају новоизабраним групама унутрашње симетрије, а билинеарна форма константи интеракције $C_{\alpha\beta}$ не мора више да има облик (104), већ може да зависи од више константи g_i, од којих по једна одговара сваком чиниоцу у производу (110). Једини захтев који $C_{\alpha\beta}$ мора да задовољи је да буде симетрична, недегенерисана и G-инваријантна, будући да су јој својствене вредности $1/g_i^2$. Коначно, сабирак S_{Proca} постаје компликованији и има облик:

$$S_{\text{Proca}} = \int \Theta^{\alpha ab} \wedge \left(\Xi_{\alpha abc} \varepsilon^{cdef} e_d \wedge e_e \wedge e_f + N_{\alpha\beta}\, \alpha^\beta \wedge e_a \wedge e_b \right) + \alpha^\alpha \wedge \tilde{N}_\alpha{}^\beta \Xi_{\beta abc} e^a \wedge e^b \wedge e^c \,. \tag{111}$$

Ова веза садржи нову билинеарну форму $N_{\alpha\beta}$ и нову матрицу $\tilde{N}_\alpha{}^\beta$, које су константне, произвољне и представљају нове слободне параметре у дејству. У циљу разумевања њихове физичке интерпретације, потребно је одредити једначине кретања. Прво се добијају једначине кретања по множитељима:

$$\begin{array}{rclrclrcl}
M_{\alpha ab} & = & -\dfrac{1}{48}\varepsilon_{abcd} e^c{}_\mu e^d{}_\nu F_\alpha{}^{\mu\nu} \,, & \lambda_{\alpha\mu\nu} & = & -F_{\alpha\mu\nu} \,, & \zeta_\alpha{}^{ab} & = & \dfrac{1}{4} C_{\alpha\beta} \varepsilon^{abcd} e_{c\mu} e_{d\nu} F^{\beta\mu\nu} \,, \\
\Theta^{\alpha ab}{}_\mu & = & \dfrac{1}{6} \tilde{N}_\beta{}^\alpha \varepsilon^{abcd} \alpha^\beta{}_\nu e_c{}^\nu e_{d\mu} \,, & \lambda_{[ab]\mu\nu} & = & R_{[ab]\mu\nu} \,, & \Xi_{\alpha abc} & = & \dfrac{1}{6} N_{\alpha\beta} \varepsilon_{abcd} \alpha^\beta{}_\mu e^{d\mu} \,, \\
B_{\alpha\mu\nu} & = & -\dfrac{e}{2} C_{\alpha\beta} \varepsilon_{\mu\nu\rho\sigma} F^{\beta\rho\sigma} \,, & \beta^a{}_{\mu\nu} & = & 0 \,, & B_{[ab]\mu\nu} & = & \dfrac{1}{8\pi l_p^2} \varepsilon_{abcd} e^c{}_\mu e^d{}_\nu \,,
\end{array} \tag{112}$$

где се појављују нови параметри $C_{\alpha\beta}$, $N_{\alpha\beta}$ и $\tilde{N}_\alpha{}^\beta$. Следећа је једначина кретања за торзију $\nabla e^a = 0$ која је остала непромењена, док су једначина кретања и тензор енергије-импулса за векторска поља у облику:

$$\nabla_\mu F^{\alpha\mu}{}_\nu - M^\alpha{}_\beta \alpha^\beta{}_\nu = 0 \,, \tag{113}$$

$$T_{\mu\nu} = C_{\alpha\beta} \left(F^\alpha{}_{\mu\rho} F^\beta{}_\nu{}^\rho - \frac{1}{4} g_{\mu\nu} F^\alpha{}_{\rho\sigma} F^{\beta\rho\sigma} \right) + C_{\alpha\beta} M^\beta{}_\gamma \left(\alpha^\alpha{}_\mu \alpha^\gamma{}_\nu - \frac{1}{2} g_{\mu\nu} \alpha^{\alpha\rho} \alpha^\gamma{}_\rho \right) . \tag{114}$$

Овде је нова матрица $M^\alpha{}_\beta$ добијена из $C_{\alpha\beta}$, $N_{\alpha\beta}$ и $\tilde{N}_\alpha{}^\beta$ као:

$$M^\alpha{}_\beta = \frac{1}{2} \left(C^{-1} \right)^{\alpha\gamma} \left(\tilde{N}_\gamma{}^\delta N_{\delta\beta} + \tilde{N}_\beta{}^\delta N_{\delta\gamma} \right) . \tag{115}$$

Ова матрица се може интерпретирати као матрица квадрата маса. Ову матрицу је могуће конструисати јер је билинеарна форма $C_{\alpha\beta}$ недегенерисана, па је самим тим и инвертибилна. У циљу интерпретације $M^\alpha{}_\beta$ као матрице квадрата маса, параметри $C_{\alpha\beta}$, $N_{\alpha\beta}$ и $\tilde{N}_\alpha{}^\beta$ морају да се изаберу тако да је (115) ненегативно дефинитна. У том случају се избором базиса у Лијевој алгебри $\mathfrak{g}$ који одговара својственом базису $M^\alpha{}_\beta$, са својственим вредностима означеним са $M^2_{(\alpha)}$, матрица квадрата маса може записати у облику

$$M^\alpha{}_\beta = M^2_{(\alpha)} \delta^\alpha_\beta \,, \tag{116}$$

где заграде око индекса α означавају да се по индексу не сумира. Једначина кретања (113) је у овом базису представљена скупом распрегнутих једначина

$$\nabla_\mu F^{\alpha\mu}{}_\nu - M^2_{(\alpha)} \alpha^\alpha{}_\nu = 0 \,. \tag{117}$$

Овај скуп једначина за више Прокиних поља, потенцијално различитих маса $M_{(\alpha)}$, објашњава због чега се $M^{\alpha}{}_{\beta}$ може интерпретирати као матрица квадрата маса. Такође, треба приметити да су добијена једначина кретања (117), једначина тензора енергије-импулса (114), природна уопштења једначина (107) и (106), редом, добијених за једно Прокино поље. Штавише, као и у случају једног Прокиног поља, алгебарске једначине кретања за помоћна поља (112) се исто могу заменити натраг у дејство (102) са (111) чиме се добија стандардна теорија другог реда за више Прокиних поља спрегнутих са Ајнштајн-Картановом гравитацијом.

Да би дејство за Прокино поље било могуће успешно упоредити са дејством добијеним у наредним одељцима као резултат Хигсовог механизма члан по члан, потребно је још мало модификовати дејство. Конкретно, потребно је модификовати Прокину везу (111) увођењем још два Лагранжева множитеља, 1-форму θ^{α} и 3-форму ρ_{α}, на следећи начин:

$$\begin{aligned} S_{\text{Proca}} &= \int \Theta^{\alpha ab} \wedge \left(\Xi_{\alpha abc} \varepsilon^{cdef} e_d \wedge e_e \wedge e_f + N_{\alpha\beta}\, \alpha^{\beta} \wedge e_a \wedge e_b \right) + \alpha^{\alpha} \wedge \tilde{N}_{\alpha}{}^{\beta} \rho_{\beta} \\ &+ \theta^{\alpha} \wedge \left(\rho_{\alpha} - \Xi_{\alpha abc} e^a \wedge e^b \wedge e^c \right) . \end{aligned} \tag{118}$$

Два додатна Лагранжева множитеља повећавају конфигурациони простор на одговарајући начин, тако да он постаје компатибилан са конфигурационим простором који се појављује у одељку о Хигсовом механизму. Осим тога, множитељи не мењају ни једну другу особину дејства за Прокина поља. Конкретно, једначине кретања (112) и (117), као и тензор енергије-импулса (114) и једначина за торзију остају непромењене. Наравно, повећање конфигурационог простора значи да постоје додатне једначине кретања за два нова множитеља:

$$\theta^{\alpha} = -\tilde{N}_{\beta}{}^{\alpha} \alpha^{\beta}{}_{\mu} \,, \qquad \rho_{\alpha\nu\rho\sigma} = e\, M_{\alpha\beta}\, \varepsilon_{\mu\nu\rho\sigma} \alpha^{\beta\mu} \,. \tag{119}$$

Као и до сада, и ове једначине су алгебарске и могу се заменити натраг у дејство чиме се добија стандардна теорија другог реда за Прокина поља.

Овим је закључена анализа реформулација Прокиног дејства заснованог на вишим градијентним теоријама. Облик сабирака у Прокиној вези (118) тачно одговара типу сабирака који се добијају у дејству за Стандардни Модел након спонтаног нарушења симетрије електрослабе интеракције.

3.2.2 3BF дејство са везама за теорију електрослабих интеракција

Да би Хигсов механизам био спроведен на што једноставнији начин, поступак неће бити спроведен над целим дејством (22) већ само над електрослабим сектором и то без фермиона, који ће бити разматрани накнадно. Другим речима, одговарајући избор 3-групе је:

$$G = SO(3,1) \times SU(2) \times U(1) \,, \qquad H = \mathbb{R}^4 \,, \qquad L = \mathbb{C}^4 \,. \tag{120}$$

Група G садржи Лоренцову подгрупу, подгрупу за слаб изоспин $SU(2)$ и подгрупу за слаб хипернабој $U(1)$. Група H остаје иста као до сада, описујући транслације, док је група L редукована тако да описује само дублет комплексних скаларних поља. Хомоморфизми δ и ∂ остају тривијални, као и Пајферово подизање $\{\_\,,\_\}_{\text{pf}}$. Коначно, дејство групе је следеће. Група на саму себе делује коњугацијом, Лоренцова подгрупа делује на стандардан начин на групу H и тривијално на групу L, јер су све компоненте поља из групе L скаларна поља. Слаб изоспин и хипернабој делују тривијално на групу H, док делују нетривијално на групу

L. Да би ово дејство било експлицитно изражено, корисно је увести матричну репрезентацију генератора T_A групе L, као:

$$T_1 = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}\,, \qquad T_2 = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}\,, \qquad T_3 = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}\,, \qquad T_4 = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}\,. \tag{121}$$

Даље, ако се генератори изоспина означе са τ_i $(i = 1,2,3)$, и генератор хипернабоја са τ_0, добија се:

$$\tau_\alpha \triangleright T_A = \triangleright_{\alpha A}{}^B T_B\,, \tag{122}$$

где индекси α узимају вредности из скупа $0,\dots,3$, што једнобразно пребројава све генераторе (τ_0, τ_i) групе $SU(2)\times U(1)$. Коефицијенти су експлицитно дати као:

$$\begin{aligned}
\triangleright_{0A}{}^B &= \frac{i}{2}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}\,, \qquad \triangleright_{1A}{}^B = \frac{i}{2}\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&-1\\0&0&-1&0\end{pmatrix}\,,\\
\triangleright_{2A}{}^B &= \frac{i}{2}\begin{pmatrix}0&i&0&0\\-i&0&0&0\\0&0&0&i\\0&0&-i&0\end{pmatrix}\,, \qquad \triangleright_{3A}{}^B = \frac{i}{2}\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&1\end{pmatrix}\,.
\end{aligned} \tag{123}$$

Треба приметити да генератори $SU(2)\times U(1)$ задовољавају уобичајене комутационе релације

$$f_{\alpha\beta\gamma} = \begin{cases} -\varepsilon_{\alpha\beta\gamma}\,, & \text{за } \alpha,\beta,\gamma \neq 0\,,\\ 0 & \text{иначе.}\end{cases} \tag{124}$$

Овим је фиксиран избор 3-групе. Даље, билинеарне форме су дефинисане на природан начин и за групе G и H су дефинисане као у (17) и (18), док је за групу L избор помало необичан:

$$g_{AB} = \frac{1}{2}\begin{pmatrix}0&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&0\end{pmatrix}\,. \tag{125}$$

Ово је само привидно необично јер је базис у алгебри $\mathfrak{l}$ у облику

$$\phi \equiv \phi^A T_A \equiv \phi_+ T_1 + \phi_0 T_2 + \phi_+^\dagger T_3 + \phi_0^\dagger T_4 = \begin{pmatrix}\phi_+\\ \phi_0\\ \phi_+^\dagger\\ \phi_0^\dagger\end{pmatrix}\,. \tag{126}$$

Међутим, преласком у нови базис у алгебри $\mathfrak{l}$ датим са

$$\tilde{T}_1 = T_1 + T_3\,, \qquad \tilde{T}_2 = iT_1 - iT_3\,, \qquad \tilde{T}_3 = T_2 + T_4\,, \qquad \tilde{T}_4 = iT_2 - iT_4\,, \tag{127}$$

се исти елемент алгебре може преписати у облику

$$\phi = \phi_1\tilde{T}_1 + \phi_2\tilde{T}_2 + \phi_3\tilde{T}_3 + \phi_4\tilde{T}_4\,, \tag{128}$$

где су $\phi_1, \ldots, \phi_4$ реалне компоненте, и где постоји природна веза између коефицијената:

$$\phi_+ = \phi_1 + i\phi_2\,, \qquad \phi_0 = \phi_3 + i\phi_4\,, \qquad \phi_+^\dagger = \phi_1 - i\phi_2\,, \qquad \phi_0^\dagger = \phi_3 - i\phi_4\,. \tag{129}$$

У базису $\tilde{T}_A$ је билинеарна форма (125) репрезентована дијагоналном јединичном матрицом. Базис $\tilde{T}_A$ је користан јер је својствен за билинеарну форму и у њему компоненте поља постају реалне, док је базис T_A користан јер је својствен за операторе слабог изоспина (његове једне компоненте) и слабог хипернабоја (а испоставиће се и оператора наелектрисања). У овом поглављу ће једначине бити записиване у оба базиса према потреби.

Треба уочити да се (128) може разумети као елемент 4-димензионалне реалне Лијеве алгебре $\mathfrak{l} = \mathbb{R}^4$, или као 2-димензионалне комплексне Лијеве алгебре $\mathfrak{l} = \mathbb{C}^2$ (која је имплицитно коришћена у највећем броју стандардних уџбеника о Хигсовом механизму). Са друге стране, (126) је елемент 4-димензионалне комплексне Лијеве алгебре $\mathfrak{l} = \mathbb{C}^4$, која је комплексификација $\mathbb{R}^4$. Како би прелазак из T_A базиса у $\tilde{T}_A$ базис и назад био што једноставнији, корисно је радити у комплексификованој $\mathfrak{l} = \mathbb{C}^4$ алгебри, због чега ово и јесте избор групе $L = \mathbb{C}^4$ у електрослабој 3-групи (120).

Након избора 3-групе и билинеарних форми, дејство за електрослабу теорију се може записати као:

$$S = S_{3BF} + S_{\text{grav}} + S_{\text{scal}} + S_{\text{Yang-Mills}} + S_{\text{Higgs}} + S_{\text{CC}}\,. \tag{130}$$

Сличног је облика као (22), где су везе за фермионе изостављене. Билинеарна форма интеракционих константи је у $S_{\text{Yang-Mills}}$ дата са

$$C_{\alpha\beta} = \begin{pmatrix} \frac{1}{g_0^2} & 0 & 0 & 0 \\ 0 & \frac{1}{g_1^2} & 0 & 0 \\ 0 & 0 & \frac{1}{g_1^2} & 0 \\ 0 & 0 & 0 & \frac{1}{g_1^2} \end{pmatrix}\,, \tag{131}$$

и има структуру $SU(2) \times U(1)$ групе.

3.2.3 Преглед Хигсовог механизма

Постоје три главна корака у спровођењу Хигсовог механизма:

- дискусија стабилног вакуума,
- увођење смене променљивих,
- фиксирање калибрације скаларних поља.

У циљу разумевања детаља Хигсовог механизма спроведеног над дејством (130), корисно је поновити ове главне кораке коришћењем нових променљивих и нотације.

Анализа стабилног вакуума је идентична анализи у случају уобичајеног Хигсовог механизма. Веза S_{Higgs} уводи потенцијал за скаларна поља у облику

$$V(\phi) = 2\chi \left(\phi^A \phi_A - v^2\right)^2\,, \tag{132}$$

и може се уочити да стабилан вакуум није јединствен, већ представља било коју тачку на 3-сфери $\phi^A\phi_A = v^2$ у конфигурационом простору. Да би дејство било записано у облику у ком сва поља имају вредност једнаку нули у тачки у конфигурационом простору која одговара изабраном вакууму, потребно је увести смену промењљивих из $(\phi_1, \phi_2, \phi_3, \phi_4)$ у $(\phi_1, \phi_2, h, \phi_4)$, где је $h(x)$ ново скаларно поље добијено транслацијом ϕ_3 за v:

$$\phi_3(x) = v + h(x)\,. \tag{133}$$

Ово одговара избору тачке $(0, 0, v, 0)$ на 3-сфери за вакуум, према конвенцији. Наравно, ова конвенција је потпуно произвољна и ништа у остатку анализе не зависи од ње. Смена променљивих је дата у базису $\tilde{T}_A$, док у оригиналном базису T_A она гласи:

$$\phi^A = \begin{pmatrix} \phi_+ \\ \phi_0 \\ \phi_+^\dagger \\ \phi_0^\dagger \end{pmatrix} = \begin{pmatrix} \phi_1 + i\phi_2 \\ v + h + i\phi_4 \\ \phi_1 - i\phi_2 \\ v + h - i\phi_4 \end{pmatrix}\,. \tag{134}$$

Коначно, на основу ове релације се може уочити да су компоненте ϕ_1, ϕ_2 и ϕ_4 суштински еквивалентне (у линеарном развоју) параметрима $\mathfrak{g}$-градијентне трансформације

$$\phi \to \phi' = e^{\xi^\alpha \tau_\alpha} \triangleright \phi\,. \tag{135}$$

Конкретно, коришћењем дејства (122) генератора алгебре $\mathfrak{g}$ на генераторе алгебре $\mathfrak{l}$, може се поћи од тренутног стања и избора параметара градијентне трансформације,

$$\phi^A = \begin{pmatrix} 0 \\ v + h \\ 0 \\ v + h \end{pmatrix}, \qquad \xi^\alpha(\phi^A) = \frac{1}{v} \begin{pmatrix} \phi_4 \\ 2\phi_2 \\ 2\phi_1 \\ -\phi_4 \end{pmatrix} + \mathcal{O}(\phi^2)\,, \tag{136}$$

а потом, применом трансформације (135) на стање изнад, добити:

$$\phi'^A = e^{\xi^\alpha(\phi)\tau_\alpha} \begin{pmatrix} 0 \\ v + h \\ 0 \\ v + h \end{pmatrix} = \begin{pmatrix} \phi_1 + i\phi_2 \\ v + h + i\phi_4 \\ \phi_1 - i\phi_2 \\ v + h - i\phi_4 \end{pmatrix}\,. \tag{137}$$

Због тога је очигледно да се поља ϕ_1, ϕ_2 и ϕ_4 могу разумети као степени слободе фиксирања калибрације дати релацијом (136), па је јасно да је само поље h физичко, обзиром да се не може уклонити $\mathfrak{g}$-градијентном трансформацијом.

Даље, може се уочити да и након уклањања поља ϕ_1, ϕ_2 и ϕ_4 коришћењем градијентне трансформације, стање ϕ^A у (136) и даље остаје очувано при дејству $U(1)$ подгрупе G. Ако се генератор ове групе означи са Q, једноставно је уочити из (123) да је услов стабилизације вакуума $Q \triangleright \phi = 0$ задовољен за

$$Q = \tau_0 + \tau_3\,. \tag{138}$$

Ова једначина је позната под називом Гел-Ман–Нишиџимина формула (за електрослабе интеракције). Феноменолошки, Q одговара електромагнетном набоју q, конкретно q је својствена вредност оператора $-iQ$, и одговарајућа градијентна група стабилизатора $U(1)$ је градијентна група симетрије електродинамике. Из једначине стабилизатора је очигледно да Хигсово поље није наелектрисано, јер је својствено оператору $-iQ$ за својствену вредност $q = 0$.

Треба још једном нагласити да претходни резултати не зависе од избора вакуума $(0, 0, v, 0)$ на 3-сфери. Ако би уместо ове тачке била одабрана нека друга, разликовало би се једино решење једначине стабилизатора $Q \triangleright \phi = 0$, али би се то решење увек могло записати у општем облику као:

$$Q = \tau_0 + \vec{\alpha} \cdot \vec{\tau} , \qquad \vec{\alpha} \in \mathbb{R}^3 , \qquad \|\vec{\alpha}\|^2 = 1 . \tag{139}$$

Овде је $\vec{\tau}$ тројка (τ_1, τ_2, τ_3). Међутим, набој Хигсовог поља и даље остаје једнак нули, јер је он последица једначине стабилизатора вакуума.

3.2.4 Трансформација дејства

Следећи корак је замена уведених смена у дејство и фиксирање калибрације применом градијентне трансформације која уклања компоненте поља ϕ_1, ϕ_2 и ϕ_4,

$$\phi^A \to \left(e^{-\xi} \triangleright \phi\right)^A = \begin{pmatrix} 0 \\ v+h \\ 0 \\ v+h \end{pmatrix} , \tag{140}$$

где је $\xi \equiv \xi^\alpha \tau_\alpha$, а параметри ξ^α су дати у (136). У циљу испитивања шта се догађа са дејством (130), потребно је утврдити дејство градијентне трансформације на сва поља у дејству. Трансформације 3-конексија $(\alpha, \omega, \beta, \tilde{\gamma})$ су дате као:

$$\alpha' = e^{-\xi} \left(\alpha + \mathrm{d}\right) e^{\xi} , \qquad \omega' = \omega , \qquad \beta' = \beta , \qquad \tilde{\gamma}' = e^{-\xi} \triangleright \tilde{\gamma} . \tag{141}$$

Одговарајуће кривине се трансформишу према:

$$F' = e^{-\xi} F e^{\xi} , \qquad R' = R , \qquad \mathcal{G}' = \mathcal{G} , \qquad \mathcal{H}' = e^{-\xi} \triangleright \mathcal{H} . \tag{142}$$

Трансформације Лагранжевих множитеља који се појављују у тополошком делу дејства B_α, $B_{[ab]}$, e_a и ϕ^A, су дате са:

$$B'_\alpha = \left(e^{-\xi} B e^{\xi}\right)_\alpha , \qquad B'_{[ab]} = B_{[ab]} , \qquad e'_a = e_a , \tag{143}$$

док је трансформација ϕ^A већ одређена у (140). Даље, Лагранжеви множитељи из алгебре $\mathfrak{g}$ који се појављују у везама λ_α, $\lambda_{[ab]}$, $M_{\alpha ab}$ и $\zeta_{\alpha ab}$, се трансформишу као:

$$\lambda'_\alpha = \left(e^{-\xi} \lambda e^{\xi}\right)_\alpha , \qquad \lambda'_{[ab]} = \lambda_{[ab]} , \qquad M'_{\alpha ab} = \left(e^{-\xi} M e^{\xi}\right)_{\alpha ab} , \qquad \zeta'_{\alpha ab} = \left(e^{-\xi} \zeta e^{\xi}\right)_{\alpha ab} . \tag{144}$$

Лагранжеви множитељи из алгебре $\mathfrak{l}$ који се појављују у везама $\tilde{\lambda}_A$, Λ_{abA} и H_{abcA}, се трансформишу:

$$\tilde{\lambda}'_A = \left(e^{-\xi} \triangleright \tilde{\lambda}\right)_A , \qquad \Lambda'_{abA} = \left(e^{-\xi} \triangleright \Lambda\right)_{abA} , \qquad H'_{abcA} = \left(e^{-\xi} \triangleright H\right)_{abcA} . \tag{145}$$

Коначно, везе садрже коваријантни извод $\nabla\phi$, који се трансформише као

$$(\nabla\phi)' = e^{-\xi} \triangleright (\nabla\phi) , \tag{146}$$

што је и очекивано.

Додатно, дејство садржи и билинеарну форму интеракционих константи, $C_{\alpha\beta}$, дату са (131). Може се уочити да је она већ део по део пропорционална већ уведеној билинеарној форми $\langle \_ \, , \_ \rangle_{\mathfrak{g}}$, на следећи начин:

$$C_{\alpha\beta} = \mathcal{C}(\tau_\alpha, \tau_\beta) \equiv \frac{\delta^j_\alpha \delta^k_\beta}{g_1^2} \langle \tau_j, \tau_k \rangle_{\mathfrak{g}} + \frac{\delta^0_\alpha \delta^0_\beta}{g_0^2} \langle \tau_0, \tau_0 \rangle_{\mathfrak{g}} . \tag{147}$$

Два сабирка у суми одговарају билинеарним формама $\langle \_ \, , \_ \rangle_{\mathfrak{su}(2)}$ и $\langle \_ \, , \_ \rangle_{\mathfrak{u}(1)}$, редом. Будући да се градијентна трансформација може представити у облику $e^{-\xi^i \tau_i} \times e^{-\xi^0 \tau_0}$, захваљујући директном производу у структури групе $SU(2) \times U(1)$, сваки чинилац у градијентној трансформацији оставља одговарајућу билинеарну форму инваријантном,

$$\langle e^{-\xi^i \tau_i} \triangleright \tau_j , e^{-\xi^i \tau_i} \triangleright \tau_k \rangle_{\mathfrak{su}(2)} = \langle \tau_j , \tau_k \rangle_{\mathfrak{su}(2)} \, , \qquad \langle e^{-\xi^0 \tau_0} \triangleright \tau_0 , e^{-\xi^0 \tau_0} \triangleright \tau_0 \rangle_{\mathfrak{u}(1)} = \langle \tau_0 , \tau_0 \rangle_{\mathfrak{u}(1)} \, , \tag{148}$$

као последицу постулиране G-инваријантности билинеарне форме $\langle \_ \, , \_ \rangle_{\mathfrak{g}}$. Ово имплицира да билинеарна форма константи интеракције такође мора да буде инваријантна на градијентне трансформације:

$$C'_{\alpha\beta} = C_{\alpha\beta} \, . \tag{149}$$

Сада је могуће испитати трансформацију целог дејства у односу на (140). Конкретно, дејство (130) је функционал свих горе поменутих поља,

$$\alpha^\alpha, \; \omega^{[ab]}, \; \beta^a, \; \tilde{\gamma}^A, \; B_\alpha, \; B_{[ab]}, \; e_a, \; \lambda_\alpha, \; \lambda_{[ab]}, \; M_{\alpha ab}, \; \zeta_{\alpha ab}, \; \tilde{\lambda}_A, \; \Lambda_{abA}, \; H_{abcA}, \; \phi^A, \tag{150}$$

или другим речима, горе наведена листа поља дефинише цео кинематички конфигурациони простор дејства. Међутим, нису сви сабирци у дејству функције поља ϕ^A. Због тога трансформација (140) делује тривијално на сабирке који су независни од ϕ^A, док се сабирци зависни од ϕ^A трансформишу нетривијално, редукујући конфигурациони простор на мањи, дефинисан над пољима

$$\alpha^\alpha, \; \omega^{[ab]}, \; \beta^a, \; \tilde{\gamma}^A, \; B_\alpha, \; B_{[ab]}, \; e_a, \; \lambda_\alpha, \; \lambda_{[ab]}, \; M_{\alpha ab}, \; \zeta_{\alpha ab}, \; \tilde{\lambda}_A, \; \Lambda_{abA}, \; H_{abcA}, \; h, \tag{151}$$

која се од почетног скупа разликују за смену $(\phi^1, \phi^2, \phi^3, \phi^4) \to (0, 0, v+h, 0)$. Задатак се сада своди на одређивање облика дејства $\tilde{S}$ које је дефинисано на редукованом конфигурационом простору, шематски дефинисано трансформацијом:

$$S[\dots, \phi^A] \quad \xrightarrow{\;e^{-\xi}\;} \quad \tilde{S}[\dots, h] \equiv S[\dots, \phi^A] \Big|_{\substack{\phi^1=\phi^2=\phi^4=0 \\ \phi^3 = v+h}} \, . \tag{152}$$

Може се одмах уочити да се сабирци S_{grav}, $S_{\text{Yang-Mills}}$ и S_{CC} трансформишу тривијално јер не зависе од поља ϕ:

$$S_{\text{grav}} \xrightarrow{e^{-\xi}} \tilde{S}_{\text{grav}} = S_{\text{grav}} \, , \quad S_{\text{Yang-Mills}} \xrightarrow{e^{-\xi}} \tilde{S}_{\text{Yang-Mills}} = S_{\text{Yang-Mills}} \, , \quad S_{\text{CC}} \xrightarrow{e^{-\xi}} \tilde{S}_{\text{CC}} = S_{\text{CC}} \, . \tag{153}$$

Штавише, $2BF$ парче од S_{3BF} сабирка се такође трансформише тривијално из истог разлога.

Са друге стране, трећи сабирак у S_{3BF} делу, као и S_{scal} и S_{Higgs} захтевају више пажње. За почетак, S_{Higgs} сабирак је облика

$$S_{\text{Higgs}} = - \int \frac{1}{4!} V(\phi) \, \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d \, , \tag{154}$$

па деловањем (152) Хигсов потенцијал $V(\phi)$ (погледати (132)) постаје

$$V(\phi) \quad \xrightarrow{e^{-\xi}} \quad V(h) \equiv 8v^2 \chi h^2 + 8v\chi h^3 + 2\chi h^4 \, . \tag{155}$$

Због тога је очигледно да важи

$$S_{\text{Higgs}} \quad \xrightarrow{e^{-\xi}} \quad \tilde{S}_{\text{Higgs}} = - \int \frac{1}{4!} V(h) \, \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d \, . \tag{156}$$

Из облика квадратичног члана у резултујућем потенцијалу (155) и општег облика масеног члана за једно скаларно поље (90), може се закључити да је маса Хигсовог поља:

$$m = 2v\sqrt{2\chi}\,. \tag{157}$$

Облик потенцијала се може изабрати и на другачији начин у односу на (132), на пример

$$V_{\rm alt}(\phi) = 2\chi\left(\phi^A\phi_A\right)^2 - 4\chi v^2\phi^A\phi_A\,. \tag{158}$$

Овај потенцијал се разликује од (132) за константан сабирак $2\chi v^4$, који се онда сабере са $S_{\rm CC}$ и даје другачију вредност космолошке константе. Међутим, потенцијал (132) не генерише додатни сабирак, па CC сабирак у дејству остаје исти након спонтаног нарушења симетрије.

Следећи сабирак је S_{3BF} члан. Само последњи сабирак у S_{3BF} зависи од поља ϕ, док остатак не зависи и може се означити са S_{2BF}. Затим, коришћењм одговарајуће промене базиса $T_A \to \tilde{T}_A$ у Лијевој алгебри $\mathfrak{l}$ (погледати (127)), са додатном нотацијом за индексе $A \to (\bar{A}, H)$ где је $\bar{A} \in \{1,2,4\}$ и $H \equiv 3$, добија се

$$\begin{aligned} S_{3BF} = S_{2BF} + \int \phi^A\nabla\tilde{\gamma}_A \quad \xrightarrow{e^{-\xi}} \quad & S_{2BF} + \int (v+h)\left((\nabla\tilde{\gamma})_0 + (\nabla\tilde{\gamma})_{0^\dagger}\right) \\ &= S_{2BF} + \int (v+h)\left({\rm d}\tilde{\gamma}_H + \rhd_\alpha{}^{\bar{A}}{}_H\,\alpha^\alpha\wedge\tilde{\gamma}_{\bar{A}}\right) \\ &= S_{2BF} + \int h{\rm d}\tilde{\gamma}_H + v{\rm d}\tilde{\gamma}_H + (v+h)\rhd_\alpha{}^{\bar{A}}{}_H\,\alpha^\alpha\wedge\tilde{\gamma}_{\bar{A}} \\ &= \tilde{S}_{3BF} + \int v{\rm d}\tilde{\gamma}_H + (v+h)\rhd_\alpha{}^{\bar{A}}{}_H\,\alpha^\alpha\wedge\tilde{\gamma}_{\bar{A}}\,, \end{aligned} \tag{159}$$

где је ново дејство $\tilde{S}_{3BF}$ дефинисано као функционал над пољима из редукованог конфигурационог простора (151) као

$$\tilde{S}_{3BF} = S_{2BF} + \int h{\rm d}\tilde{\gamma}_H\,. \tag{160}$$

У наредном одељку ће бити дискутована његова одговарајућа 3-група. Коначно, може се закључити да важи

$$S_{3BF} \quad \xrightarrow{e^{-\xi}} \quad \tilde{S}_{3BF} + \int v{\rm d}\tilde{\gamma}_H + (v+h)\rhd_\alpha{}^{\bar{A}}{}_H\,\alpha^\alpha\wedge\tilde{\gamma}_{\bar{A}}\,, \tag{161}$$

где ће додатни чланови бити груписани заједно са другим додатним члановима из преосталих делова дејства и дискутовани заједно.

У једначинама (159), (160) и (161) је коришћен базис (127) у Лијевој алгебри $\mathfrak{l}$, тако да се може увести $\tilde{\gamma}_H \equiv \tilde{\gamma}_0 + \tilde{\gamma}_{0^\dagger}$. Дејство $\rhd$ је било репрезентовано матрицама (123) у оригиналном базису T_A, који је сада подељен на следеће компоненте:

$$\rhd_{\alpha H}{}^{\bar{A}}\,, \qquad \rhd_{\alpha\bar{A}}{}^{H}\,, \qquad \rhd_{\alpha\bar{A}}{}^{\bar{B}}\,, \qquad \rhd_{\alpha H}{}^{H}\,. \tag{162}$$

Пошто је билинеарна форма g_{AB} у овом базису дијагонална, и заправо једнака $g_{AB} = \delta_{AB}$, према Теореми 1 са почетка ове главе, дијагонални елементи свих ових компонената морају да буду једнаки нули. Конкретно, $\rhd_{\alpha H}{}^H = 0$, што даље имплицира да је $\nabla\tilde{\gamma}_H \equiv {\rm d}\tilde{\gamma}_H$ и оправдава идентификацију (160). Штавише, комоненте $\rhd_{\alpha\bar{A}}{}^{\bar{B}}$ не фигуришу у једначинама (159), (160) и

(161) и не појављују се нигде. То значи да су једине релевантне компоненте оне које се могу изразити у матричном облику као:

$$\rhd_{\alpha H}{}^{\bar{A}} = \frac{1}{2}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad \rhd_{\alpha \bar{A}}{}^{H} = \frac{1}{2}\begin{pmatrix} 0 & 0 & -1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{163}$$

Треба уочити да је овде α индекс који пребројава врсте, а $\bar{A}$ индекс који пребројава колоне, док је индекс $H \equiv 3$ константан.

Коначно, последњи сабирак у дејству је S_{scal}, који има почетни облик

$$S_{\text{scal}} = \int \tilde{\lambda}^A \wedge \left(\tilde{\gamma}_A - H_{abcA} e^a \wedge e^b \wedge e^c\right) + \Lambda^{abA} \wedge H_{abcA} \varepsilon^{cdef} e_d \wedge e_e \wedge e_f - \Lambda^{abA} \wedge (\nabla\phi)_A \wedge e_a \wedge e_b\,, \tag{164}$$

и слично као S_{3BF}, зависи од ϕ^A само у последњем сабирку, док је остатак независан од ϕ^A. Раздвајањем индекса A на $(\bar{A}, H)$, сабирак се трансформише у:

$$\begin{aligned} S_{\text{scal}} \xrightarrow{e^{-\xi}} & \int \tilde{\lambda}^H \wedge \left(\tilde{\gamma}_H - H_{abcH} e^a \wedge e^b \wedge e^c\right) + \Lambda^{abH} \wedge H_{abcH} \varepsilon^{cdef} e_d \wedge e_e \wedge e_f \\ & -\Lambda^{abH} \wedge \mathrm{d}h \wedge e_a \wedge e_b + \tilde{\lambda}^{\bar{A}} \wedge \left(\tilde{\gamma}_{\bar{A}} - H_{abc\bar{A}} e^a \wedge e^b \wedge e^c\right) \\ & +\Lambda^{ab\bar{A}} \wedge H_{abc\bar{A}} \varepsilon^{cdef} e_d \wedge e_e \wedge e_f - \Lambda^{ab\bar{A}} \wedge \alpha^\alpha \rhd_\alpha{}^H{}_{\bar{A}} (v+h) \wedge e_a \wedge e_b\,. \end{aligned} \tag{165}$$

Треба уочити да први ред са десне стране заједно са првим сабирком у другом реду одговара тачно скаларној вези за једно реално скаларно поље, као функционал на редукованом конфигурационом простору (151). Ова веза се може означити са $\tilde{S}_{\text{scal}}$, па за цео израз важи:

$$\begin{aligned} S_{\text{scal}} \xrightarrow{e^{-\xi}} \tilde{S}_{\text{scal}} + \int & \tilde{\lambda}^{\bar{A}} \wedge \left(\tilde{\gamma}_{\bar{A}} - H_{abc\bar{A}} e^a \wedge e^b \wedge e^c\right) \\ & + \Lambda^{ab\bar{A}} \wedge H_{abc\bar{A}} \varepsilon^{cdef} e_d \wedge e_e \wedge e_f \\ & - \Lambda^{ab\bar{A}} \wedge \alpha^\alpha \rhd_\alpha{}^H{}_{\bar{A}} (v+h) \wedge e_a \wedge e_b\,, \end{aligned} \tag{166}$$

где ће поново додатни сабирци бити груписани са осталим додатним сабирцима из дејства.

Након разматрања свих делова дејства (130) појединачно, при деловању трансформације (152), сви делови се могу поново сабрати и упоредити са дејством за масено скаларно поље и Прокина поља. Међутим, да би поређење било што једноставније, корисно је увести још нове нотације. Конкретно, може се увести билинеарна форма $\kappa^{\alpha\beta}$ која задовољава следећи идентитет:

$$\kappa^{\alpha\beta} \rhd_{\alpha H}{}^{\bar{A}} \rhd_{\beta \bar{B}}{}^{H} = -\frac{1}{4} \delta^{\bar{A}}_{\bar{B}}\,. \tag{167}$$

Избор ове билинеарне форме није јединствен. Конкретно, будући да су матрице (163) ранга 3, постоји пројектор $P_\alpha{}^\beta$ који задовољава

$$P_\alpha{}^\beta P_\beta{}^\gamma = P_\alpha{}^\gamma\,, \qquad P_\alpha{}^\alpha = 3\,, \qquad P_{\alpha\beta} = P_{\beta\alpha}\,, \qquad P_\alpha{}^\beta \rhd_{\beta H}{}^{\bar{A}} = \rhd_{\alpha H}{}^{\bar{A}}\,. \tag{168}$$

Треба уочити да пројектор (168) такође задовољава идентитет $P_\alpha{}^\beta \rhd_{\beta \bar{A}}{}^H = \rhd_{\alpha \bar{A}}{}^H$, јер $\rhd_{\alpha \bar{A}}{}^H$ садржи исте компоненте као $\rhd_{\alpha H}{}^{\bar{A}}$ до на знак минус, погледати (163). Због тога је билинеарна форма $\kappa^{\alpha\beta}$ дефинисана до на сабирак у облику

$$\kappa^{\alpha\beta} \to \kappa^{\alpha\beta} + \left[\delta^{(\alpha}_\gamma - P_\gamma{}^{(\alpha}\right] A^{\beta)\gamma}\,, \tag{169}$$

где је $A^{\alpha\beta}$ произвољна матрица, док заграде око индекса означавају симетризацију. Може се уочити да је израз у загради ортогонални пројектор који пресликава у језгро матрице (163). Ова произвољност гарантује да се билинеарна форма $\kappa^{\alpha\beta}$ може изабрати тако да буде инвертибилна. Пројектор $P_\alpha{}^\beta$ се може директно одредити из дефиниције (168) и матрица (163), док се један конкретан избор билинеарне форме $\kappa^{\alpha\beta}$ може добити из (167), тако да могу да буду записани у матричној форми као:

$$P_\alpha{}^\beta = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}, \qquad \kappa^{\alpha\beta} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{170}$$

Додатно, осим пројектора и $\kappa^{\alpha\beta}$, се могу увести и додатне величине:

$$\theta^\alpha \equiv -2\kappa^{\alpha\beta}\triangleright_\beta{}^H{}_{\bar{A}}\tilde{\lambda}^{\bar{A}}, \qquad \Theta^{\alpha ab} \equiv -2\kappa^{\alpha\beta}\triangleright_\beta{}^H{}_{\bar{A}}\Lambda^{ab\bar{A}}, \qquad \rho_\alpha \equiv 2\triangleright_\alpha{}^{\bar{A}}{}_H\tilde{\gamma}_{\bar{A}}, \qquad \Xi_{\alpha abc} \equiv 2\triangleright_\alpha{}^{\bar{A}}{}_H H_{abc\bar{A}}. \tag{171}$$

Ове четири величине задовољавају четири фундаментална идентитета,

$$\theta^\alpha\wedge\rho_\alpha = \tilde{\lambda}^{\bar{A}}\wedge\tilde{\gamma}_{\bar{A}}, \qquad \theta^\alpha\Xi_{\alpha abc} = \tilde{\lambda}^{\bar{A}}H_{abc\bar{A}}, \qquad \Theta^{\alpha ab}\wedge\rho_\alpha = \Lambda^{ab\bar{A}}\wedge\tilde{\gamma}_{\bar{A}}, \qquad \Theta^{\alpha ab}\Xi_{\alpha cde} = \Lambda^{ab\bar{A}}H_{cde\bar{A}}, \tag{172}$$

који су директна последица идентитета (167). Сврха увођења ових величина је у томе да се помоћу њих могу елиминисати $\bar{A}$ индекси из једначина. Треба уочити да је и у дефиницијама (171) и у идентитетима (172) сумирано по индексима $\bar{A}$ са десне стране, док се они уопште не појављују са леве стране.

Важно је нагласити да произвољност $\kappa^{\alpha\beta}$ у (169) уводи промене у дејству. Ово је у вези са чињеницом да смена променљивих (171) уводи додатне променљиве које не постоје у оригиналном дејству. Захтев да ове додатне променљиве не постоје, односно да на левој страни идентитета (172) постоји исти број компонената као на десној страни, редукује произвољност (169) билинеарне форме $\kappa^{\alpha\beta}$ која сада има облик:

$$\kappa^{\alpha\beta} \to \kappa^{\alpha\beta} + \Big[\delta^\alpha_\gamma - P_\gamma{}^\alpha\Big]A^{\gamma\delta}\Big[\delta^\beta_\delta - P_\delta{}^\beta\Big]. \tag{173}$$

Важно је нагласити и да је и након овог захтева могуће изабрати инвертибилну билинеарну форму $\kappa^{\alpha\beta}$, док је само дејство инваријантно на промену билинеарне форме (173), што значи да теорија не зависи од тог избора. Видети Додатак A.2 за детаљнију анализу и доказ.

Након увођења нових величина и нотације се може наставити са анализом дејства. За дату трансформацију (152) дејства (130), могу се применити дефиниције (171) и идентитети (172) како би се елиминисали сви индекси $\bar{A}, \bar{B}$, а трансформисано дејство $\tilde{S}$ постало функционал над редукованим конфигурационим простором (151). Састављањем резултата (153), (156), (161) и (166), добија се пун облик дејства:

$$\begin{aligned}\tilde{S} \;=\;& S_{\text{grav}} + S_{\text{Yang-Mills}} + S_{\text{CC}} + \tilde{S}_{\text{Higgs}} + \tilde{S}_{3BF} + \tilde{S}_{\text{scal}} \\ &+ \int \Theta^{\alpha ab}\wedge\Big(\Xi_{\alpha abc}\,\varepsilon^{cdef}e_d\wedge e_e\wedge e_f + \frac{v}{2}\kappa^{-1}_{\alpha\beta}P_\gamma{}^\beta\alpha^\gamma\wedge e_a\wedge e_b\Big) \\ &+ \int \frac{v}{2}\alpha^\alpha P_\alpha{}^\beta\wedge\rho_\beta + \theta^\alpha\wedge\big(\rho_\alpha - \Xi_{\alpha abc}e^a\wedge e^b\wedge e^c\big) \\ &+ \frac{1}{2}\int h\,\alpha^\alpha P_\alpha{}^\beta\wedge\big(\rho_\beta - \kappa^{-1}_{\beta\gamma}\Theta^{\gamma ab}\wedge e_a\wedge e_b\big) + v\int \mathrm{d}\tilde{\gamma}_H\,.\end{aligned} \tag{174}$$

Овај облик дејства је сада упоредив са дејством за Прокино поље (102). Специјално, други и трећи ред у (174) се могу упоредити са Прокином везом у (118). Упоређивање слободних параметара члан по члан даје идентификацију слободних параметара у Прокином дејству на следећи начин:

$$N_{\alpha\beta} = \frac{v}{2}\kappa^{-1}_{\alpha\gamma}P_{\beta}{}^{\gamma}\,, \qquad \tilde{N}_{\alpha}{}^{\beta} = \frac{v}{2}P_{\alpha}{}^{\beta}\,. \tag{175}$$

Коришћењем ових израза, може се конструисати матрица квадрата маса (115) и добити

$$M^{\alpha}{}_{\beta} = \frac{v^2}{4}\left(C^{-1}\right)^{\alpha\gamma}P_{\gamma}{}^{\delta}\kappa^{-1}_{\delta\epsilon}P_{\beta}{}^{\epsilon}\,, \tag{176}$$

где је билинеарна форма интеракционих константи $C_{\alpha\beta}$ дефинисана у (131), док су пројектор и билинеарна форма $\kappa^{\alpha\beta}$ дати у (170). На основу тога се за матрицу квадрата маса добија:

$$M^{\alpha}{}_{\beta} = \frac{v^2}{4}\begin{pmatrix} g_0^2 & 0 & 0 & -g_0^2 \\ 0 & g_1^2 & 0 & 0 \\ 0 & 0 & g_1^2 & 0 \\ -g_1^2 & 0 & 0 & g_1^2 \end{pmatrix}\,. \tag{177}$$

Физички релевантан базис у Лијевој алгебри $\mathfrak{g}$ је онај у којем је матрица квадрата маса дијагонална, а одговарајуће својствене вредности могуће интерпретирати као квадрате маса векторских бозона у том базису. Будући да су прва и последња колона у (177) пропорционалне, детерминанта матрице је једнака нули, па је бар једна њена својствена вредност једнака нули. Штавише, матрица је већ у блок-дијагоналној форми, са два једнодимензионална блока g_1^2, одакле се може закључити да су те две својствене вредности једнаке $v^2g_1^2/4$. Коначно, из трага матрице се може добити и четврта својствена вредност, па је цео масени спектар дат са:

$$M_1^2 = 0\,, \qquad M_2^2 = \frac{v^2}{4}g_1^2\,, \qquad M_3^2 = \frac{v^2}{4}g_1^2\,, \qquad M_4^2 = \frac{v^2}{4}(g_0^2+g_1^2)\,. \tag{178}$$

Чињеница да су својствене вредности M_2^2 и M_3^2 једнаке имплицира да својствени базис није једнозначно одређен, и да је потребно додати још специјалније захтеве. Природни избор ових захтева је да базис, осим за матрицу квадрата маса буде својствен и за стабилизатор Q, уведен у (138), будући да он представља електрични набој, чија би вредност требала да буде добро дефинисана за сва описана физичка стања у преферираном базису. Стабилизатор се у ту сврху може изразити у матричном облику $Q_{\alpha}{}^{\beta}$, дефинисан дејством Q на базисни вектор τ_{α}:

$$Q \triangleright \tau_{\alpha} = Q_{\alpha}{}^{\beta}\tau_{\beta}\,. \tag{179}$$

Коришћењем (138) се једноставно добија да су компоненте матрице $Q_{\alpha}{}^{\beta}$

$$Q_{\alpha}{}^{\beta} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\,. \tag{180}$$

Ова матрица има својствене вредности $(0, i, -i, 0)$, што значи да оператор наелектрисања $-iQ$ има својствене вредности:

$$q_1 = 0\,, \qquad q_2 = +1\,, \qquad q_3 = -1\,, \qquad q_4 = 0\,. \tag{181}$$

Стабилизатор такође има две исте својствене вредности, па ни његов својствени базис није једнозначно одређен. У сваком случају, матрица квадрата маса и стабилизатор међусобно комутирају, па имају заједнички својствени базис и овај својствени базис је једнозначно одређен. Нови својствени базис се може изразити преко старог,

$$\tau_A = \tau_0 + \tau_3 \,, \qquad \tau_+ = \frac{\tau_1 + i\tau_2}{\sqrt{2}} \,, \qquad \tau_- = \frac{\tau_1 - i\tau_2}{\sqrt{2}} \,, \qquad \tau_Z = -\frac{g_0^2}{g_0^2+g_1^2}\tau_0 + \frac{g_1^2}{g_0^2+g_1^2}\tau_3 \,, \tag{182}$$

и могу се изразити компоненте 1-форме конексије $\alpha = \alpha^\alpha{}_\mu \mathrm{d}x^\mu \otimes \tau_\alpha$ у новом базису као:

$$A_\mu = \frac{g_1^2}{g_0^2+g_1^2}\alpha^0{}_\mu + \frac{g_0^2}{g_0^2+g_1^2}\alpha^3{}_\mu \,, \quad W^+_\mu = \frac{\alpha^1{}_\mu - i\alpha^2{}_\mu}{\sqrt{2}} \,, \quad W^-_\mu = \frac{\alpha^1{}_\mu + i\alpha^2{}_\mu}{\sqrt{2}} \,, \quad Z_\mu = -\alpha^0{}_\mu + \alpha^3{}_\mu \,. \tag{183}$$

Овде је такође уведена традиционална нотација за векторске бозоне. Електромагнетни набоји добијена четири бозона су већ уграђени у нотацију, док се њихове масе могу прочитати из (178):

$$M_A = 0 \,, \qquad M_{W^\pm} = \frac{v}{2}g_1 \,, \qquad M_Z = \frac{v}{2}\sqrt{g_0^2+g_1^2} \,. \tag{184}$$

У новом базису су матрица квадрата маса и матрица стабилизатора дијагонални, док су билинеарна форма $g_{\alpha\beta}$ и билинеарна форма интеракционих константи $C_{\alpha\beta}$:

$$g_{\alpha\beta} = \begin{pmatrix} 2 & 0 & 0 & \frac{g_1^2-g_0^2}{g_1^2+g_0^2} \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{g_1^2-g_0^2}{g_1^2+g_0^2} & 0 & 0 & \frac{g_1^4+g_0^4}{(g_1^2+g_0^2)^2} \end{pmatrix} \,, \qquad C_{\alpha\beta} = \begin{pmatrix} \frac{g_0^2+g_1^2}{g_0^2 g_1^2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{g_1^2} & 0 \\ 0 & \frac{1}{g_1^2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{g_0^2+g_1^2} \end{pmatrix} \equiv \begin{pmatrix} \frac{1}{g_A^2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{g_W^2} & 0 \\ 0 & \frac{1}{g_W^2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{g_Z^2} \end{pmatrix} . \tag{185}$$

Упоређивањем са дејством у (174), очигледно је да други и трећи ред представљају Прокину везу, па је коначно дејство у облику:

$$\begin{aligned} \tilde{S} \;=\; & S_{\text{grav}} + S_{\text{Yang-Mills}} + S_{\text{CC}} + \tilde{S}_{\text{Higgs}} + \tilde{S}_{3BF} + \tilde{S}_{\text{scal}} + S_{\text{Proca}} \\ & + \frac{1}{2}\int h\,\alpha^\alpha P_\alpha{}^\beta \wedge \left(\rho_\beta - \kappa^{-1}_{\beta\gamma}\Theta^{\gamma ab} \wedge e_a \wedge e_b\right) + v\int \mathrm{d}\tilde{\gamma}_H \,. \end{aligned} \tag{186}$$

Први ред у дејству садржи сабирак који описује једно реално скаларно поље h (Хигсово поље) масе $m = 2v\sqrt{2\chi}$, и четири векторска бозона са масама датим у (184), спрегнутих са гравитацијом и међусобно. Први сабирак у другом реду описује интеракцију између Хигсовог поља и векторских бозона, тако да су све интеракције еквивалентне са интеракцијама у стандардној електрослабој теорији. Други сабирак у другом реду је гранични члан и као такав не утиче на једначине кретања у теорији.

3.2.5 Спонтано нарушење 3-групе симетрије и масени спектар фермиона

Да би анализа претходних резултата била комплетна, потребно је дати одговор на још два питања. Прво питање је шта се догађа са структуром 3-групе у процесу спонтаног нарушења симетрије. Конкретно, полазни 2-укрштени модул који одговара теорији електрослабих интеракција, је заснован на избору група у (120):

$$G = SO(3,1) \times SU(2) \times U(1) \,, \qquad H = \mathbb{R}^4 \,, \qquad L = \mathbb{C}^4 \,. \tag{187}$$

Међутим, након анализе и спровођења Хигсовог механизма, резултујућем дејству (186) више не одговара овај 2-укрштени модул. Уместо њега, једноставно је уочити, одговара 2-укрштени модул заснован на следећем избору група:

$$G = SO(3,1) \times SU(2) \times U(1)\,, \qquad H = \mathbb{R}^4\,, \qquad L = \mathbb{R}\,. \tag{188}$$

Овде постоје два детаља која треба нагласити. Први, да је група L редукована тако да уместо да описује четири комплексна скаларна поља, сада описује једно реално скаларно поље. Ово је директна последица спонтаног нарушења симетрије, конкретно, градијентне трансформације (140) која је примењена да би се уклонила поља ϕ_1, ϕ_2 и ϕ_4. Трансформација је довела до редукције конфигурационог простора остављајући само једно реално скаларно поље h у дејству. Нови облик, редукованог конфигурационог простора је конзистентан са новим избором (188) 2-укрштеног модула, који одговара тополошком $3BF$ дејству (160), добијеном фиксирањем калибрације почетног $3BF$ дејства заснованом на 2-укрштеном модулу (187).

Други детаљ се тиче групе G. Формално, група G у финалном 2-укрштеном модулу је остала иста као што је била у иницијалном 2-укрштеном модулу. Међутим, као што је показано у поглављу о експлицитном нарушењу симетрије, Прокина веза нарушава G групу симетрије, и она је једина веза која то ради. Због тога, иако тополошки 3BF део почетног и крајњег дејства има исти BF сабирак и исте 1-форме конексије које следе из групе G, присуство Прокине везе нарушава групу симетрије G на њену подгрупу $SO(3,1) \times U(1)$, док почетно дејство није садржало Прокину везу и група G није била нарушена. Коначан резултат је такав да крајње дејство има нарушену G симетрију иако је засновано на 2-укрштеном модулу (188) који садржи целу групу G, што је последица присуства Прокине везе.

Друго питање на које треба одговорити је шта се дешава са целим Стандардним Моделом (22) током процеса спонтаног нарушења симетрије. Конкретно у овом поглављу, у његовим претходним деловима, да би анализа била што једноставнија, проучавано је спонтано нарушење симетрије и Хигсов механизам само у специјалном случају теорије електрослабих интеракција. У сваком случају, једноставно је додати преостале три везе S_{Dirac}, S_{Yukawa}, и S_{spin}, као и одговарајући $\langle D \wedge \mathcal{H} \rangle_{\mathfrak{l}}$ сабирак за фермионе и добити дејство за цео Стандардни Модел. Резултат спровођења процедуре спонтаног нарушења симетрије и Хигсовог механизма на оваквом дејству је исти као и у случају електрослабе теорије, до на одговарајуће сабирке који одговарају фермионима и додатној $SU(3)$ градијентној симетрији (која остаје ненарушена јер не учествује у Хигсовом механизму). Сабирци $\langle D \wedge \mathcal{H} \rangle_{\mathfrak{l}}$ за фермионе су једнаки

$$\langle D_f \wedge \mathcal{H}_f \rangle_{\mathfrak{l}} \equiv \bar{\psi}_A (\overset{\rightarrow}{\nabla} \gamma)^A - (\bar{\gamma} \overset{\leftarrow}{\nabla})_A \psi^A\,, \tag{189}$$

и трансформишу се тривијално у односу на (152) као и везе S_{Dirac} и S_{spin}:

$$\begin{aligned} &\langle D_f \wedge \mathcal{H}_f \rangle_{\mathfrak{l}} \ \xrightarrow{e^{-\xi}} \ \langle \tilde{D}_f \wedge \tilde{\mathcal{H}}_f \rangle_{\mathfrak{l}} = \langle D_f \wedge \mathcal{H}_f \rangle_{\mathfrak{l}}\,, \\ S_{\text{Dirac}} \ \xrightarrow{e^{-\xi}} \ \tilde{S}_{\text{Dirac}} = S_{\text{Dirac}}\,, \quad & S_{\text{spin}} \ \xrightarrow{e^{-\xi}} \ \tilde{S}_{\text{spin}} = S_{\text{spin}}\,. \end{aligned} \tag{190}$$

Једина веза која се не трансформише тривијално је веза за Јукавин потенцијал S_{Yukawa}, и она се дели на два сабирка:

$$\begin{aligned} S_{\text{Yukawa}} &= -\int \frac{2}{4!} Y_{ABC} \bar{\psi}^A \psi^B \phi^C \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d \\ \xrightarrow{e^{-\xi}} \quad & -\frac{1}{12} \int v Y_{ABH}\, \bar{\psi}^A \psi^B\, \varepsilon_{abcd}\, e^a \wedge e^b \wedge e^c \wedge e^d - \frac{1}{12} \int Y_{ABH}\, \bar{\psi}^A \psi^B h\, \varepsilon_{abcd}\, e^a \wedge e^b \wedge e^c \wedge e^d\,, \end{aligned} \tag{191}$$

где први сабирак у другом реду одговара вези за масу Диракових поља (93), док је други сабирак нова веза за Јукавин потенцијал $\tilde{S}_{\text{Yukawa}}$, и описује интеракцију између фермиона и Хигсовог поља h. Упоређивањем првог сабирка са везом за масу Диракових поља, закључује се да константе Јукавине интеракције Y_{ABH} јесу пропорционалне масеној матрици фермиона

$$M_{AB} = vY_{ABH} \,, \tag{192}$$

која се састоји од правих маса фермиона и одговарајућих углова мешања. Коначни облик трансформисане везе за Јукавиним потенцијалом постаје:

$$S_{\text{Yukawa}} \quad \xrightarrow{e^{-\xi}} \quad \tilde{S}_{\text{Yukawa}} - \frac{1}{12}\int M_{AB}\bar{\psi}^A\psi^B\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\wedge e^d = \tilde{S}_{\text{Yukawa}} + S_{\text{Dirac mass}} \,. \tag{193}$$

Сада је показано да се Хигсов механизам описан на примеру теорије електрослабих интеракција може уопштити на цео Стандардни Модел (22).

Овиме је закључена анализа Хигсовог механизма и спонтаног нарушења симетрије 3BF дејства са везама. Укратко, резултати Хигсовог механизма су исти као и у стандардној тензорској формулацији Лагранжијана за Стандардни Модел, али су процедура и технички детаљи потпуно другачији, будући да је 3BF дејство функционал на потпуно различитом конфигурационом простору у односу на стандардно дејство за Стандардни Модел.

4 Веза између формулација квантне Ајнштајн-Картанове и 3BF теорије са везама

У претходној глави је показана еквиваленција између класичне 3BF теорије са везама и Стандардног Модела спрегнутог са Ајнштајн-Картановом гравитацијом. Иако су ове две теорије класично еквивалентне, њиховом квантизацијом се потенцијално могу добити различите квантне теорије. У овој глави ће бити успостављена веза између очекиваних вредности опсервабли у две квантне теорије, демонстрирајући да теорије нису потпуно еквивалентне на квантном нивоу, али да постоји добро дефинисана веза између њих. Након тога, у последњем поглављу ће бити разматрани примери на којима се може илустровати разлика у квантним теоријама.

4.1 Квантне опсервабле

Процес конструкције квантне теорије захтева уопштавање математичких резултата везаних за вишеструке интеграле по реалним и Грасмановим бројевима на одговарајуће функционалне интеграле по бозонским и фермионским пољима. Због тога је у првом одељку дат преглед интеграла који ће бити уопштени до функционалног нивоа. У другом одељку ће ови интеграли бити примењени на дефиницију очекиване вредности произвољне опсервабле у квантној 3BF теорији, корак по корак, тако да резултат добијен на крају процедуре може да се интерпретира као очекивана вредност везом придружене опсервабле у квантној ECC теорији. На овај начин ће бити конструисана потпуна непертурбативна веза између две квантне теорије.

4.1.1 Математички резултати

У овом одељку ће бити изложена укупно четири уопштења особина Диракове делте, у неколико специјалних случајева, као и идентитет повезан са Стоксовом теоремом. Идентитети се могу поделити у две групе, бозонску и фермионску. Два бозонска идентитета се могу добити уопштавањем следећих особина Диракове делте на бозонска поља. Први идентитет је

$$\int_{\mathbb{R}} dy\, e^{iyF} = 2\pi\, \delta(F)\,, \qquad F \in \mathbb{R}\,, \tag{194}$$

на основу ког се може добити следећи вишеструки интеграл:

$$\int_{\mathbb{R}} dy \int_{\mathbb{R}^n} dx_k\, H(x_k)\, e^{i(yF(x_k)+G(x_k))} = 2\pi \int_{\mathbb{R}^n} dx_k\, H(x_k)\delta\left(F(x_k)\right)\, e^{iG(x_k)}\,. \tag{195}$$

Потом се, уопштењем овог резултата на ниво функционалног интеграла, добија:

$$\int D\varphi D\phi_k\, H(\phi_k)\, e^{i\int(\varphi\wedge F(\phi_k)+G(\phi_k))} = \mathcal{N} \int D\phi_k\, H(\phi_k)\delta\left(F(\phi_k)\right)\, e^{i\int G(\phi_k)}\,. \tag{196}$$

Други потребан идентитет је:

$$\int_{\mathbb{R}} dy\, \delta(yF-G)\, H(y) = \frac{H(G/F)}{|F|}\,, \qquad F,G \in \mathbb{R}\,, \tag{197}$$

који се може једноставно доказати сменом промењљивих. Одговарајући вишеструки интеграл је

$$\int_{\mathbb{R}} dy \int_{\mathbb{R}^n} dx_k \, \delta\left(y^{aB} F_B{}^A(x_k) - G^{aA}(x_k)\right) H(y, x_k) = \int_{\mathbb{R}^n} dx_k \frac{H\left(G^{aB}(x_k) F^{-1}{}_B{}^A(x_k), x_k\right)}{|F(x_k)|^{|a|}} \,, \tag{198}$$

док је уопштење на ниво функционалног интеграла дато са:

$$\int D\varphi D\phi_k \, \delta\left(\varphi^{aB} F_B{}^A(\phi_k) - G^{aA}(\phi_k)\right) H(\varphi, \phi_k) = \int D\phi_k \frac{1}{|F(\phi_k)|^{|a|}} H\left(G^{aB}(\phi_k) F^{-1}{}_B{}^A(\phi_k), \phi_k\right) \,, \tag{199}$$

где су $F_B{}^A(\phi_k)$ произвољна инвертибилна матрична функција, $G^{aA}(\phi_k)$ и $H(\varphi, \phi_k)$ произвољне функције, и $|a|$ означава број могућих вредности индекса a.

Даље, идентитет у вези са Стоксовом теоремом

$$\begin{aligned} &\int D\varphi D\phi_k \, H(\varphi, \phi_k) \, e^{i \int (\nabla\varphi)\wedge E(\phi_k) + \varphi \wedge F(\phi_k) + G(\phi_k)} \\ &\qquad = \int D\varphi_\partial D\phi_{k\partial} \, e^{i \oint \varphi_\partial \wedge E(\phi_{k\partial})} \int D\varphi D\phi_k \, H(\varphi, \phi_k) \, e^{i \int (-1)^{p-1} \varphi \wedge \nabla E(\phi_k) + \varphi \wedge F(\phi_k) + G(\phi_k)} \,, \end{aligned} \tag{200}$$

који подједнако важи и за бозонска и за фермионска поља. Овде је претпостављено да је φ p-форма, док су φ_∂ и $\phi_{k\partial}$ вредности поља на граници многострукости. Треба нагласити да је главна сврха овог идентитета пребацивање дејства коваријантног извода ∇ са φ на $E(\phi_k)$.

Коначно, два фермионска идентитета која се могу подједнако уопштити са Грасманових бројева на Грасманова поља су дата са

$$\begin{aligned} &\int_{\mathbb{G}^n} d\theta_1 d\theta_2 \dots d\theta_n \; e^{i\theta_1(\theta_2 - \theta_3 - \dots - \theta_k)} \; F(\theta_2, \dots, \theta_n) \\ &\qquad = (-1)^{n-1} i \int_{\mathbb{G}^{n-1}} d\theta_2 \dots d\theta_n \; \delta(\theta_2 - \theta_3 - \dots - \theta_k) \; F(\theta_2, \dots, \theta_n) \,, \end{aligned} \tag{201}$$

и

$$\begin{aligned} &\int_{\mathbb{R}^m} d^m y \int_{\mathbb{R}^m} d^m x \int_{\mathbb{G}^k} d\theta_1 d\theta_2 \dots d\theta_k \int_{\mathbb{G}^{n-k}} d\theta_{k+1} \dots d\theta_n \; e^{i y_a \left(x^a - M^{aij}\theta_i\theta_j\right)} \; F(x, \theta_1, \dots, \theta_n) \\ &\quad = (2\pi)^m \int_{\mathbb{R}^m} d^m x \int_{\mathbb{G}^k} d\theta_1 d\theta_2 \dots d\theta_k \int_{\mathbb{G}^{n-k}} d\theta_{k+1} \dots d\theta_n \; \prod_{a=1}^m \delta\left(x^a - M^{aij}\theta_i\theta_j\right) \; F(x, \theta_1, \dots, \theta_n) \,. \end{aligned} \tag{202}$$

За доказ погледати Додатак A.3. Одговарајући функционални идентитети за Грасманова поља изгледају исто као два идентитета изнад, до на замену мере интеграције $d\theta \to D\theta$, нормализационог фактора $(2\pi)^m \to \mathcal{N}$, и нотације $\prod \delta(x^a) \to \delta(\phi)$. Треба приметити да последњи идентитет садржи Дираков делту од комбинације реалних и Грасманових бројева, па је редослед интеграције важан. Конкретно, прво се мора спровести интеграција по реалним бројевима x^a, па тек онда по Грасмановим бројевима. Такође, Диракова делта од Грасманових бројева је непарна (кососхермитска) као последица чињенице да Грасманови бројеви антикомутирају, па се делта увек појављује у пару са имагинарном јединицом $i \equiv \sqrt{-1}$ у једначини (201).

4.1.2 Очекиване вредности опсервабли

Веза између две квантне теорије се може добити упоређивањем очекиваних вредности опсервабли између две теорије. Да би то било успешно спроведено, потребно је дефинисати ове очекиване вредности у обе теорије:

$$\langle F\rangle_{3BF} = \frac{1}{Z_{3BF}} \int D\phi_i \; F(\phi_k)\; e^{iS_{3BF}[\phi_i]}\,, \qquad \langle F\rangle_{ECC} = \frac{1}{Z_{ECC}} \int D\phi_i \; F(\phi_k)\; e^{iS_{ECC}[\phi_i]}\,, \tag{203}$$

где су суме по стањима дате као:

$$Z_{3BF} = \int D\phi_i \; e^{iS_{3BF}[\phi_i]}\,, \qquad Z_{ECC} = \int D\phi_i \; e^{iS_{ECC}[\phi_i]}\,. \tag{204}$$

Потребно је одмах на почетку дати пар коментара. Прво, за потребе анализе није неопходно задати прецизну дефиницију интеграла по трајекторијама, што значи да није потребно у потпуности дефинисати квантну 3BF и ECC теорију експлицитно. Једини услов који је потребно претпоставити је да су мере у оба интеграла дефинисане на исти начин за обе теорије и да су дефинисане на такав начин да функционални идентитети (196), (199), (200), (201) и (202) и даље важе. На овај начин је могуће дискутовати особине и упоређивати очекиване вредности опсервабли у обе квантне теорије у непертурбативном режиму, упркос чињеници да сами детаљи квантизације теорија нису до краја дефинисани.

Друго, поља ϕ_i у 3BF теорији и ECC теорији припадају њиговим појединачним конфигурационим просторима (47)-(48) и (56), редом. Очигледно је да се опсервабле $F(\phi_k)$ у (203) могу упоређивати само ако припадају заједничком конфигурационом подпростору обе теорије, односно, ако опсервабла F зависи само од поља ϕ_k која припадају редукованом конфигурационом простору (56) који је дефинисан за ECC теорију.

У даљим корацима ће бити примењивани идентитети (196), (199), (200), (201) и (202), корак по корак, у циљу свођења дефиниције опсервабле F у квантној 3BF теорији, на дефиницију очекиване вредности одговарајуће опсервабле прво у квантној EC теорији, а потом и у квантној ECC теорији.

За почетак, потребно је експлицитно записати вредност производа очекиване вредности опсервабле F и суме по стањима у квантној 3BF теорији. Према дефиницији (203) је тај производ једнак:

$$
\begin{aligned}
Z_{3BF}\langle F\rangle_{3BF} \;=\; & \mathcal{N}\int D\alpha D\omega D\beta De D\tilde{\gamma} D\gamma D\bar{\gamma} D\phi D\psi D\bar{\psi} DB D\tilde{\lambda} D\lambda D\bar{\lambda} D\Lambda D\zeta DH DM \\
& \exp\left(i\int B_{\alpha}\wedge F^{\alpha} + B_{[ab]}\wedge R^{[ab]} + e_a\wedge(\nabla\beta)^a + \phi^A\left(\nabla\tilde{\gamma}\right)_A + \bar{\psi}_A\left(\nabla\gamma\right)^A \right. \\
& -\left(\bar{\gamma}\overleftarrow{\nabla}\right)_A\psi^A \\
& +\lambda^{\alpha}\wedge\left(B_{\alpha} - 12C_{\alpha\beta}M^{\beta}{}_{ab}e^a\wedge e^b\right) - \lambda_{[ab]}\wedge\left(B^{[ab]} - \frac{1}{8\pi l_p^2}\varepsilon^{[ab]cd}e_c\wedge e_d\right) \\
& +\tilde{\lambda}^A\wedge\left(\tilde{\gamma}_A - H_{abcA}e^a\wedge e^b\wedge e^c\right) + \bar{\lambda}_A\wedge\left(\gamma^A + \frac{i}{6}\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\left(\gamma^d\psi\right)^A\right) \\
& -\lambda^A\wedge\left(\bar{\gamma}_A - \frac{i}{6}\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\left(\bar{\psi}\gamma^d\right)_A\right) + 2\pi i l_p^2\bar{\psi}_A\gamma_5\gamma^a\psi^A\varepsilon_{abcd}e^b\wedge e^c\wedge\beta^d \\
& +\zeta_{\alpha}{}^{ab}\left(M^{\alpha}{}_{ab}\varepsilon_{cdef}e^c\wedge e^d\wedge e^e\wedge e^f - F^{\alpha}\wedge e_a\wedge e_b\right) \\
& +\Lambda^{abA}\wedge\left(H_{abcA}\varepsilon^{cdef}e_d\wedge e_e\wedge e_f - (\nabla\phi)_A\wedge e_a\wedge e_b\right) \\
& \left. -\frac{1}{12}\left(\chi\left(\phi^A\phi_A - v^2\right)^2 + Y_{ABC}\bar{\psi}^A\psi^B\phi^C + \frac{\Lambda}{8\pi l_p^2}\right)\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\wedge e^d\right) \\
& F\left(\phi_k\right)\,. \qquad (205)
\end{aligned}
$$

Применом функционалног идентитета за Стоксову теорему (200) над сабирцима у другом и сабирком у трећем реду, уклањају се изводи са поља β_a, $\bar{\gamma}_A$, γ^A и $\tilde{\gamma}^A$, како би у следећем кораку могла да се изврши функционална интеграција по њима. Добијен резултат је:

$$
\begin{aligned}
Z_{3BF}\langle F\rangle_{3BF} \;=\; & \mathcal{N}\int D\beta_\partial De_\partial D\tilde{\gamma}_\partial D\gamma_\partial D\bar{\gamma}_\partial D\phi_\partial D\psi_\partial D\bar{\psi}_\partial\, e^{i\oint \phi^A_\partial\tilde{\gamma}_{A\partial}+\bar{\psi}_{A\partial}\gamma^A_\partial+\bar{\gamma}_{A\partial}\psi^A_\partial-e_{a\partial}\wedge\beta^a_\partial} \\
& \int D\alpha D\omega D\beta De D\tilde{\gamma}D\gamma D\bar{\gamma}D\phi D\psi D\bar{\psi}DBD\tilde{\lambda}D\lambda D\bar{\lambda}D\Lambda D\zeta DHDM \\
& \exp\left(i\int B_\alpha\wedge F^\alpha + B_{[ab]}\wedge R^{[ab]} + (\nabla e)_a\wedge\beta^a - (\nabla\phi)^A\wedge\tilde{\gamma}_A\right. \\
& -\left(\bar{\psi}\overleftarrow{\nabla}\right)_A\wedge\gamma^A + \bar{\gamma}_A\wedge(\nabla\psi)^A \\
& +\lambda^\alpha\wedge\left(B_\alpha - 12C_{\alpha\beta}M^\beta{}_{ab}e^a\wedge e^b\right) - \lambda_{[ab]}\wedge\left(B^{[ab]}-\frac{1}{8\pi l_p^2}\varepsilon^{[ab]cd}e_c\wedge e_d\right) \\
& +\tilde{\lambda}^A\wedge\left(\tilde{\gamma}_A - H_{abcA}e^a\wedge e^b\wedge e^c\right) + \bar{\lambda}_A\wedge\left(\gamma^A+\frac{i}{6}\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\left(\gamma^d\psi\right)^A\right) \\
& -\lambda^A\wedge\left(\bar{\gamma}_A-\frac{i}{6}\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\left(\bar{\psi}\gamma^d\right)_A\right) + 2\pi i l_p^2\bar{\psi}_A\gamma_5\gamma^a\psi^A\varepsilon_{abcd}e^b\wedge e^c\wedge\beta^d \\
& +\zeta_\alpha{}^{ab}\left(M^\alpha{}_{ab}\varepsilon_{cdef}e^c\wedge e^d\wedge e^e\wedge e^f - F^\alpha\wedge e_a\wedge e_b\right) \\
& +\Lambda^{abA}\wedge\left(H_{abcA}\varepsilon^{cdef}e_d\wedge e_e\wedge e_f - (\nabla\phi)_A\wedge e_a\wedge e_b\right) \\
& \left.-\frac{1}{12}\left(\chi\left(\phi^A\phi_A - v^2\right)^2 + Y_{ABC}\bar{\psi}^A\psi^B\phi^C + \frac{\Lambda}{8\pi l_p^2}\right)\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\wedge e^d\right) \\
& F(\phi_k)\,.
\end{aligned} \tag{206}
$$

Пажљивом провером горњег израза се може уочити да у колико опсервабла $F(\phi_k)$ не зависи од поља на граници, интеграција по граници не даје допринос. Такође, 3BF теорија формулисана на овакав начин намеће ограничења да су поља материје као и тетрада на граници многострукости једнака нули. Ова ограничења се могу уклонити додавањем одговарајућих граничних чланова у класично 3BF дејство. Један од таквих граничних чланова је и Гибонс-Хокинг-Јорк члан [53, 54, 55].

У следећем кораку се спроводи интеграција по пољима β_a, $B_{[ab]}$, B_α, $\tilde{\gamma}^A$, $\bar{\gamma}_A$, γ_A, $\zeta_\alpha{}^{ab}$ и Λ^{abA}

применом функционалних идентитета (196) и (201), који дају:

$$
\begin{aligned}
Z_{3BF}\langle F\rangle_{3BF} \;=\; & \mathcal{N}\int De_\partial D\phi_\partial D\psi_\partial D\bar{\psi}_\partial\, \delta(\phi_\partial)\delta(\psi_\partial)\delta(\bar{\psi}_\partial)\delta(e_\partial) \\
& \int D\alpha D\omega De D\phi D\psi D\bar{\psi} D\tilde{\lambda} D\lambda D\bar{\lambda} DH DM \\
& \delta\left(F^\alpha+\lambda^\alpha\right)\delta\left(R_{[ab]}-\lambda_{[ab]}\right)\delta\left(\tilde{\lambda}_A-(\nabla\phi)_A\right)\delta\left(\bar{\lambda}_A-\left(\bar{\psi}\overleftarrow{\nabla}\right)_A\right)\delta\left(\lambda^A-(\nabla\psi)^A\right) \\
& \exp\left(i\int -12\lambda^\alpha\wedge C_{\alpha\beta}M^\beta{}_{ab}e^a\wedge e^b+\lambda_{[ab]}\wedge\frac{1}{8\pi l_p^2}\varepsilon^{[ab]cd}e_c\wedge e_d\right. \\
& -\tilde{\lambda}^A\wedge H_{abcA}e^a\wedge e^b\wedge e^c \\
& +\bar{\lambda}_A\wedge\frac{i}{6}\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\left(\gamma^d\psi\right)^A+\lambda^A\wedge\frac{i}{6}\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\left(\bar{\psi}\gamma^d\right)_A \\
& \left.-\frac{1}{12}\left(\chi\left(\phi^A\phi_A-v^2\right)^2+Y_{ABC}\bar{\psi}^A\psi^B\phi^C+\frac{\Lambda}{8\pi l_p^2}\right)\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\wedge e^d\right) \\
& \delta\left(M^\alpha{}_{ab}\varepsilon_{cdef}e^c\wedge e^d\wedge e^e\wedge e^f-F^\alpha\wedge e_a\wedge e_b\right) \\
& \delta\left(H_{abcA}\varepsilon^{cdef}e_d\wedge e_e\wedge e_f-(\nabla\phi)_A\wedge e_a\wedge e_b\right) \\
& \delta\left((\nabla e)_a-2\pi i l_p^2\bar{\psi}_A\gamma_5\gamma^d\psi^A\varepsilon_{abcd}e^b\wedge e^c\right) \\
& F(\phi_k)\,.
\end{aligned}
\tag{207}
$$

Потом, даљом интеграцијом по пољима λ^α, $\lambda_{[ab]}$, $\tilde{\lambda}_A$, $\bar{\lambda}_A$ и λ^A се уклањају Диракове делте из трећег реда и добија:

$$
\begin{aligned}
Z_{3BF}\langle F\rangle_{3BF} \;=\; & \mathcal{N}\int De_\partial D\phi_\partial D\psi_\partial D\bar{\psi}_\partial\, \delta(\phi_\partial)\delta(\psi_\partial)\delta(\bar{\psi}_\partial)\delta(e_\partial) \\
& \int D\alpha D\omega De D\phi D\psi D\bar{\psi} DH DM \\
& \exp\left(i\int 12F^\alpha\wedge C_{\alpha\beta}M^\beta{}_{ab}e^a\wedge e^b+R_{ab}\wedge\frac{1}{16\pi l_p^2}\varepsilon^{abcd}e_c\wedge e_d\right. \\
& -(\nabla\phi)^A\wedge H_{abcA}e^a\wedge e^b\wedge e^c \\
& -\frac{i}{6}\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\wedge\left(\left(\bar{\psi}\overleftarrow{\nabla}\right)_A\gamma^d\psi^A-\bar{\psi}_A\gamma^d(\nabla\psi)^A\right) \\
& \left.-\frac{1}{12}\left(\chi\left(\phi^A\phi_A-v^2\right)^2+Y_{ABC}\bar{\psi}^A\psi^B\phi^C+\frac{\Lambda}{8\pi l_p^2}\right)\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\wedge e^d\right) \\
& \delta\left(M^\alpha{}_{ab}\varepsilon_{cdef}e^c\wedge e^d\wedge e^e\wedge e^f-F^\alpha\wedge e_a\wedge e_b\right) \\
& \delta\left(H_{abcA}\varepsilon^{cdef}e_d\wedge e_e\wedge e_f-(\nabla\phi)_A\wedge e_a\wedge e_b\right) \\
& \delta\left(2\pi i l_p^2\bar{\psi}_A\gamma_5\gamma^d\psi^A\varepsilon_{abcd}e^b\wedge e^c-(\nabla e)_a\right) \\
& F(\phi_k)\,.
\end{aligned}
\tag{208}
$$

Множитељи $M^\alpha{}_{ab}$ и H_{abcA}, по којима још увек није интеграљено су у вези са Хоџовим дуалима јачине поља F^α и $(\nabla\phi)_A$, редом. Применом функционалног идентитета (199) се ови множитељи могу одинтегралити у корист Хоџових дуала (50), што даје:

$$
\begin{aligned}
Z_{3BF}\langle F\rangle_{3BF} \; = \; & \mathcal{N}\int De_\partial D\phi_\partial D\psi_\partial D\bar{\psi}_\partial \,\delta(\phi_\partial)\delta(\psi_\partial)\delta(\bar{\psi}_\partial)\delta(e_\partial) \\
& \int D\alpha D\omega De D\phi D\psi D\bar{\psi}\frac{1}{|e|^{(|\alpha|+|A|(D-1))|[ab]|}} \\
& \exp\left(i\int -F^\alpha\wedge C_{\alpha\beta}\star F^\beta + R_{ab}\wedge\frac{1}{16\pi l_p^2}\varepsilon^{abcd}e_c\wedge e_d - (\nabla\phi)^A\wedge(\star\nabla\phi)_A\right. \\
& -\frac{i}{6}\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\wedge\left(\left(\bar{\psi}\overleftarrow{\nabla}\right)_A\gamma^d\psi^A - \bar{\psi}_A\gamma^d(\nabla\psi)^A\right) \\
& \left.-\frac{1}{12}\left(\chi\left(\phi^A\phi_A - v^2\right)^2 + Y_{ABC}\bar{\psi}^A\psi^B\phi^C + \frac{\Lambda}{8\pi l_p^2}\right)\varepsilon_{abcd}e^a\wedge e^b\wedge e^c\wedge e^d\right) \\
& \delta\left(2\pi i l_p^2\bar{\psi}_A\gamma_5\gamma^d\psi^A\varepsilon_{abcd}e^b\wedge e^c - (\nabla e)_a\right) \\
& F(\phi_k)\,. \qquad (209)
\end{aligned}
$$

Величина D (која се појављује у експоненту детерминанте тетраде) је димензија просторвремена, $D=4$, па је вредност експонента детерминанте тетраде $N=(|\alpha|+|A|(D-1))|[ab]|=144$, узимајући у обзир да је $|\alpha|=12$, $|A|=4$ и $|[ab]|=6$. Може се препознати да је добијени аргумент у експоненту дејство (49) Ајнштајн-Картанове EC теорије, па се резултат може записати као:

$$
\begin{aligned}
Z_{3BF}\langle F\rangle_{3BF} \; = \; & \mathcal{N}\int De_\partial D\phi_\partial D\psi_\partial D\bar{\psi}_\partial \,\delta(\phi_\partial)\delta(\psi_\partial)\delta(\bar{\psi}_\partial)\delta(e_\partial) \\
& \int D\alpha D\omega De D\phi D\psi D\bar{\psi}\;\frac{1}{|e|^N}\;\delta\left(2\pi i l_p^2\bar{\psi}_A\gamma_5\gamma^d\psi^A\varepsilon_{abcd}e^b\wedge e^c - (\nabla e)_a\right) \\
& F(\phi_k)\; e^{iS_{EC}[\phi_k]}\,. \qquad (210)
\end{aligned}
$$

На овом месту је потенцијално могуће успоставити везу између квантне 3BF теорије и квантне EC теорије. На жалост, ово није могуће јер Диракова делта под интегралом намеће додатну јаку везу између торзије и спинског тензора и мора бити одинтеграљена пре покушаја успостављања било какве везе између квантних теорија. Ово је могуће урадити интеграцијом по 1-форми спинске конексије $\omega^{[ab]}$, која је присутна у Дираковој делти као део коваријантног извода ∇ који делује на 1-форму тетраде. Да би интеграција била изведена, потребно је прво трансформисати израз унутар Диракове делте, јер оригинални израз зависи од антисиметричног дела спинске конексије у односу на други индекс и просторвременски индекс. Ова зависност се може уклонити преласком у локално инерцијални координатни систем, где се овај антисиметрични део може одредити увођењем смене променљивих $\omega_{abc}=\omega_{ab\mu}e_c{}^\mu$. Ова смена променљивих мења меру интеграла:

$$
D\omega_{ab\mu} = D\omega_{abc}\left|\frac{\delta\left(\omega_{abc}e^c{}_\mu\right)}{\delta\omega_{efg}}\right| = D\omega_{abc}\left|\delta^{[ef]}_{[ab]}e^g{}_\mu\right| = D\omega_{abc}|e|^{|[ab]|}\,. \qquad (211)
$$

Сада је могуће увести величину A_{abc} која је антисиметрична у односу на други и трећи индекс:

$$
A_{abc} = \frac{1}{2}\left(\omega_{abc} - \omega_{acb}\right)\,. \qquad (212)
$$

Једноставно је показати да компоненте нове променљиве садрже све компоненте спинске конексије (и ниједну више). Да би ово било показано, довољно је одредити вредност следеће линеарне комбинације коришћењем особине антисиметричности спинске конексије по прва два индекса:

$$A_{abc} - A_{bac} - A_{cab} = \frac{1}{2}\left(\omega_{abc} - \omega_{acb} - \omega_{bac} + \omega_{bca} - \omega_{cab} + \omega_{cba}\right) = \omega_{abc}\,. \tag{213}$$

Јакобијан $\mathcal{J}$ ове смене променљивих је константан и може се апсорбовати у нормализациону константу $\mathcal{N}$, па се мера интеграла суштински не мења:

$$\mathcal{N}\int D\omega_{abc} = \mathcal{N}\int |\mathcal{J}| DA_{abc} = \mathcal{N}'\int DA_{abc}\,. \tag{214}$$

Заменом натраг у једначину (210) се добија израз који је функција поља A_{abc}:

$$\begin{aligned} Z_{3BF}\langle F\rangle_{3BF} \;=\;& \mathcal{N}\int De_\partial D\phi_\partial D\psi_\partial D\bar\psi_\partial\,\delta(\phi_\partial)\delta(\psi_\partial)\delta(\bar\psi_\partial)\delta(e_\partial)\int D\alpha DADeD\phi D\psi D\bar\psi\,\frac{|e|^{|[ab]|}}{|e|^N} \\ & \delta\left(2\pi i l_p^2\bar\psi_A\gamma_5\gamma^d\psi^A\varepsilon_{abcd}e^b{}_\mu e^c{}_\nu\varepsilon^{\mu\nu\rho\sigma} - \left(\partial_\mu e_{a\nu}\right)\varepsilon^{\mu\nu\rho\sigma} + A_{abc}|e|e_d{}^\rho e_e{}^\sigma\varepsilon^{bcde}\right) \\ & F(\phi_k)\; e^{iS_{EC}[\phi_k]}\,. \end{aligned} \tag{215}$$

Применом функционалних идентитета (199) и (202) добија се израз који се може интегралити по пољима A_{abc} праволинијски:

$$\begin{aligned} Z_{3BF}\langle F\rangle_{3BF} \;=\;& \mathcal{N}\int De_\partial D\phi_\partial D\psi_\partial D\bar\psi_\partial\,\delta(\phi_\partial)\delta(\psi_\partial)\delta(\bar\psi_\partial)\delta(e_\partial)\int D\alpha DADeD\phi D\psi D\bar\psi\,\frac{|e|^{|[ab]|}}{|e|^N} \\ & \delta\left(A_{abc} - \left(\frac{1}{2}c_{abc} - 2\pi i l_p^2\bar\psi_A\gamma_5\gamma^d\psi^A\varepsilon_{abcd}\right)\right)\frac{1}{\left|2|e|\varepsilon^{[bc][de]}e_{[d}{}^{[\rho}e_{e]}{}^{\sigma]}\right|^{|a|}} \\ & F(\phi_k)\; e^{iS_{EC}[\phi_k]}\,. \end{aligned} \tag{216}$$

Следећи корак је одређивање детерминанте производа две тетраде и тензора Леви-Чивите. Прво је потребно одредити детерминанту тензора Леви-Чивите као $\left|\varepsilon^{[ab][cd]}\right| = \left|\varepsilon^{[\mu\nu][\rho\sigma]}\right| = 1$, а потом се може конструисати идентитет:

$$\frac{1}{|e|^{|[\mu\nu]|}} = \left|\frac{1}{|e|}\varepsilon^{[\mu\nu][\rho\sigma]}\right| = \left|e_a{}^{[\mu}e_b{}^{\nu]}e_c{}^{[\rho}e_d{}^{\sigma]}\varepsilon^{abcd}\right| = \left|2e_{[a}{}^{[\mu}e_{b]}{}^{\nu]}\right|\left|2e_{[c}{}^{[\rho}e_{d]}{}^{\sigma]}\right|\left|\varepsilon^{[ab][cd]}\right| = \left|2e_{[a}{}^{[\mu}e_{b]}{}^{\nu]}\right|^2\,, \tag{217}$$

из ког следи да је

$$\left|2e_{[a}{}^{[\mu}e_{b]}{}^{\nu]}\right| = \frac{1}{|e|^{\frac{1}{2}|[\mu\nu]|}} = \frac{1}{|e|^{\frac{1}{2}|[ab]|}}\,. \tag{218}$$

Коришћењем добијене релације (218) функционални интеграл постаје

$$\begin{aligned} Z_{3BF}\langle F\rangle_{3BF} \;=\;& \mathcal{N}\int De_\partial D\phi_\partial D\psi_\partial D\bar\psi_\partial\,\delta(\phi_\partial)\delta(\psi_\partial)\delta(\bar\psi_\partial)\delta(e_\partial)\int D\alpha DADeD\phi D\psi D\bar\psi \\ & \delta\left(A_{abc} - \left(\frac{1}{2}c_{abc} - 2\pi i l_p^2\bar\psi_A\gamma_5\gamma^d\psi^A\varepsilon_{abcd}\right)\right)\frac{1}{|e|^{N+|[ab]|\left(\frac{|a|}{2}-1\right)}} \\ & F(\phi_k)\; e^{iS_{EC}[\phi_k]}\,. \end{aligned} \tag{219}$$

У случају Стандардног Модела, експонент детерминанте тетраде је дат изразом $M = N + |[ab]|\left(\frac{|a|}{2}-1\right) = N+6 = 150$, узимајући у обзир да је $|a|=4$. Интеграција по пољу A_{abc} мења поље A_{abc} у дејству са:

$$A_{abc} = \frac{1}{2}c_{abc} - 2\pi i l_p^2 \bar{\psi}_A \gamma_5 \gamma^d \psi^A \varepsilon_{abcd}\,, \tag{220}$$

што одговара замени

$$\omega_{ab\mu} = \Delta_{ab\mu} - 2\pi i l_p^2 \bar{\psi}_A \gamma_5 \gamma^d \psi^A \varepsilon_{abcd} e^c{}_\mu\,, \tag{221}$$

која је претходно описана процедура добијања дејства за ECC теорију из дејства за EC теорију (једначина (51)). Због тога коначан израз садржи дејство за ECC теорију у експоненту, а израз за очекивану вредност опсервабле у квантној 3BF теорији постаје упоредив са изразом за очекивану вредност одговарајуће опсервабле у квантној ECC теорији:

$$\begin{aligned} Z_{3BF}\langle F\rangle_{3BF} &= \mathcal{N}\int De_\partial D\phi_\partial D\psi_\partial D\bar{\psi}_\partial\, \delta(\phi_\partial)\delta(\psi_\partial)\delta(\bar{\psi}_\partial)\delta(e_\partial)\int D\alpha De D\phi D\psi D\bar{\psi} \\ &\times\ \frac{1}{|e|^M}\, F(\phi_k)\, e^{iS_{ECC}[\phi_k]} \ =\ \mathcal{N}'\, Z_{ECC}\left\langle \frac{1}{|e|^M}F\right\rangle_{ECC}\,. \end{aligned} \tag{222}$$

У специјалном случају, заменом јединичне опсервабле, $F(\phi_k)=1$, може се одредити веза између две суме по стањима

$$\begin{aligned} Z_{3BF} &= \mathcal{N}\int De_\partial D\phi_\partial D\psi_\partial D\bar{\psi}_\partial\, \delta(\phi_\partial)\delta(\psi_\partial)\delta(\bar{\psi}_\partial)\delta(e_\partial)\int D\alpha De D\phi D\psi D\bar{\psi}\ \frac{1}{|e|^M}\ e^{iS_{ECC}[\phi_k]} \\ &= \mathcal{N}'\, Z_{ECC}\left\langle \frac{1}{|e|^M}\right\rangle_{ECC}\,. \end{aligned} \tag{223}$$

У другом специјалном случају се избором опсервабле $F(\phi_k) = |e|^M$, добија друга веза између сума по стањима

$$\begin{aligned} Z_{3BF}\left\langle |e|^M\right\rangle_{3BF} &= \mathcal{N}\int De_\partial D\phi_\partial D\psi_\partial D\bar{\psi}_\partial\, \delta(\phi_\partial)\delta(\psi_\partial)\delta(\bar{\psi}_\partial)\delta(e_\partial)\int D\alpha De D\phi D\psi D\bar{\psi}\ e^{iS_{ECC}[\phi_k]} \\ &= \mathcal{N}'\, Z_{ECC}\,. \end{aligned} \tag{224}$$

Комбиновањем веза (223) и (224), се добија нормализациона релација

$$\left\langle |e|^M\right\rangle_{3BF}\left\langle \frac{1}{|e|^M}\right\rangle_{ECC} = 1\,. \tag{225}$$

Као последњи корак, заменом (223) и (224) натраг у (222), и применом дефиниција (203), коначно се добија веза између квантних 3BF и ECC теорија облику:

$$\langle F\rangle_{3BF} = \frac{\left\langle \frac{1}{|e|^M}F\right\rangle_{ECC}}{\left\langle \frac{1}{|e|^M}\right\rangle_{ECC}}\,, \qquad \langle F\rangle_{ECC} = \frac{\left\langle |e|^M F\right\rangle_{3BF}}{\left\langle |e|^M\right\rangle_{3BF}}\,. \tag{226}$$

Једначине (226) дефинишу непертурбативну везу између очекиваних вредности опсервабли у квантној 3BF и квантној ECC теорији, и представљају главни резултат у читавој глави. Постојање ове везе наглашава важност 3BF теорије јер би њеном квантизацијом сада аутоматски била добијена и квантна ECC теорија, која је важан физички модел квантне гравитације и Стандардног Модела. Експлицитна конструкција функционалног интеграла за 3BF теорију са везама ће бити изложена у следећој глави.

4.2 Примери

У овом поглављу ће бити упоређена предвиђања квантне 3BF теорије и квантне ECC теорије на примеру густине четворозапремине просторвремена, где ће бити дискутован и семикласичан лимес, и примеру гравитационих таласа.

4.2.1 Густина четворозапремине просторвермена

Као најједноставнији пример опсервабле, може се узети опсервабла густине четворозапремине просторвремена дефинисана као

$$F(\phi_k) = \rho \equiv |e| \,, \tag{227}$$

и густина четворозапремине просторвремена као очекивана вредност овог оператора у датој теорији гравитације. Ово је мотивисано чињеницом да је четворозапремина неке четвородимензионалне области $\mathcal{R}$ просторвремена дата као

$$V(\mathcal{R}) = \int_{\mathcal{R}} d^4x\, \sqrt{-g} = \int_{\mathcal{R}} d^4x\, |e| \,, \tag{228}$$

па се $|e|$ може прозвати "густином" четворозапремине просторвремена области $\mathcal{R}$.

Применом релације веза између квантних теорија (226), добија се однос између очекиваних вредности густина четворозапремина у две квантне теорије:

$$\frac{\rho_{3BF}}{\rho_{ECC}} \equiv \frac{\langle \rho \rangle_{3BF}}{\langle \rho \rangle_{ECC}} = \frac{\left\langle \frac{1}{|e|^{M-1}} \right\rangle_{ECC}}{\langle |e| \rangle_{ECC} \left\langle \frac{1}{|e|^M} \right\rangle_{ECC}} = \frac{\langle |e| \rangle_{3BF} \left\langle |e|^M \right\rangle_{3BF}}{\left\langle |e|^{M+1} \right\rangle_{3BF}} \,. \tag{229}$$

На овом месту је потребно увести дефиниције статистичких величина коваријансе и варијансе, помоћу којих се једноставно могу издвојити квантне поправке од класичних вредности:

$$\mathrm{Cov}(X,Y) = \langle XY \rangle - \langle X \rangle \langle Y \rangle \,, \qquad \mathrm{Var}(X) = \mathrm{Cov}(X,X) = (\Delta X)^2 \,. \tag{230}$$

Овде ΔX представља стандардно одступање, односно, неодређеност опсервабле X. Коваријанса и варијанса задовољавају неједнакост Коши-Шварца

$$|\mathrm{Cov}(X,Y)| \leqslant \Delta X\, \Delta Y \,, \tag{231}$$

која се може искористити за процену коваријансе у једначини. На основу дефиниције (230), јасно је да је однос густина четворозапремине у 3BF и ECC теорији једнак:

$$\frac{\rho_{3BF}}{\rho_{ECC}} = 1 + \frac{\mathrm{Cov}\left(|e|, \frac{1}{|e|^M}\right)_{ECC}}{\langle |e| \rangle_{ECC} \left\langle \frac{1}{|e|^M} \right\rangle_{ECC}} \,, \qquad \frac{\rho_{ECC}}{\rho_{3BF}} = 1 + \frac{\mathrm{Cov}\left(|e|, |e|^M\right)_{3BF}}{\langle |e| \rangle_{3BF} \langle |e|^M \rangle_{3BF}} \,. \tag{232}$$

Затим, коришћењем (231), се добија

$$\frac{\rho_{3BF}}{\rho_{ECC}} \leqslant 1 + \left(\frac{\Delta |e|}{\langle |e| \rangle} \; \frac{\Delta \frac{1}{|e|^M}}{\left\langle \frac{1}{|e|^M} \right\rangle} \right)_{ECC} \,, \qquad \frac{\rho_{ECC}}{\rho_{3BF}} \leqslant 1 + \left(\frac{\Delta |e|}{\langle |e| \rangle} \; \frac{\Delta |e|^M}{\langle |e|^M \rangle} \right)_{3BF} \,. \tag{233}$$

У класичном лимесу се може претпоставити да ове неодређености теже нули, одакле следи да је класичан лимес густине четворозапремине исти у обе теорије.

Заправо, горња једначина указује да би класичан лимес обе теорије могао да буде исти. Да би ово било демонстрирано, треба спровести претходну процедуру у случају произвољне опсервабле $F(\phi_k)$. Полазећи од (226), добија се

$$\frac{\langle F\rangle_{3BF}}{\langle F\rangle_{ECC}} = \frac{\left\langle \frac{1}{|e|^M} F \right\rangle_{ECC}}{\left\langle \frac{1}{|e|^M} \right\rangle_{ECC} \langle F\rangle_{ECC}} , \qquad \frac{\langle F\rangle_{ECC}}{\langle F\rangle_{3BF}} = \frac{\left\langle |e|^M F\right\rangle_{3BF}}{\langle |e|^M\rangle_{3BF} \langle F\rangle_{3BF}} . \tag{234}$$

Очекивана вредност производа у бројиоцу се може изразити у облику коваријансе коришћењем (230), а потом се коваријанса може проценити применом неједнакости Коши-Шварца (231), што доводи до скупа неједнакости:

$$\frac{\langle F\rangle_{3BF}}{\langle F\rangle_{ECC}} \leqslant 1 + \left(\frac{\Delta F}{\langle F\rangle} \frac{\Delta \frac{1}{|e|^M}}{\left\langle \frac{1}{|e|^M} \right\rangle} \right)_{ECC} , \qquad \frac{\langle F\rangle_{ECC}}{\langle F\rangle_{3BF}} \leqslant 1 + \left(\frac{\Delta F}{\langle F\rangle} \frac{\Delta |e|^M}{\langle |e|^M\rangle} \right)_{3BF} . \tag{235}$$

У класичном лимесу се може очекивати да неодређености опсервабли постају занемарљиве, $\Delta F \to 0$, одакле следи

$$\frac{\langle F\rangle_{3BF}}{\langle F\rangle_{ECC}} \leqslant 1 , \qquad \frac{\langle F\rangle_{ECC}}{\langle F\rangle_{3BF}} \leqslant 1 . \tag{236}$$

Комбиновањем две неједнакости се добија једнакост класичног лимеса

$$\langle F\rangle_{3BF} = \langle F\rangle_{ECC} , \tag{237}$$

која показује да све опсервабле имају исте очекиване вредности у класичном лимесу у обе теорије. Дакле, две теорије се међусобно разликују само на нивоу квантних поправки.

4.2.2 Гравитациони таласи

У примеру гравитационих таласа се појављују два (међусобно повезана) поља, конкретно, поље тетрада и метричког тензора, као и њихове ексцитације око равног просторвремена:

$$e^a{}_\mu = \delta^a_\mu + \varepsilon^a{}_\mu , \qquad g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} . \tag{238}$$

Детерминанта тетраде се може развити у ред по поправкама:

$$|e| = 1 + \varepsilon^a{}_a + \frac{1}{2}\left(\varepsilon^a{}_a\varepsilon^b{}_b - \varepsilon^a{}_b\varepsilon^b{}_a\right) + \frac{1}{6}\left(\varepsilon^a{}_a\varepsilon^b{}_b\varepsilon^c{}_c - 3\varepsilon^a{}_a\varepsilon^b{}_c\varepsilon^c{}_b + 2\varepsilon^a{}_b\varepsilon^b{}_c\varepsilon^c{}_a\right) + |\varepsilon| , \tag{239}$$

али ће њена вредност на даље бити апроксимирана до квадратног сабирка. У општем случају се могу разматрати гравитационе пертурбације и у закривљеном просторвремену. Тада постоји позадински део тетраде $\hat{e}^a{}_\mu$ и одговарајућа позадинска метрика $\hat{g}_{\mu\nu}$, а уместо (238) се може написати

$$e^a{}_\mu = \hat{e}^a{}_\mu + \varepsilon^a{}_\mu , \qquad g_{\mu\nu} = \hat{g}_{\mu\nu} + h_{\mu\nu} . \tag{240}$$

Иако је анализа овог општег случаја концептуално слична претходној, технички је захтевнија, јер је развој детерминанте тетраде (239) различит. На пример, уместо 1, водећи члан у развоју ће бити det $\hat{e}^a{}_\mu$. Због тога је једноставније не разматрати општи случај закривљеног простора. Са друге стране, сви главни закључци анализе остају исти.

Могуће је увести величину E која у себи садржи све поправке детерминанте тетраде у односу на јединицу. Ова смена је корисна за проучавање конвергенције развоја у ред $\pm M$-тог степена детерминанте тетраде по овом параметру, па су развоји одговарајућег степена детерминанте тетраде до другог реда дати као:

$$\frac{1}{|e|^M} = \frac{1}{(1+E)^M} = \sum_{n=0}^{+\infty} \binom{M+n}{n} (-E)^n = 1 - M\varepsilon^a{}_a + \frac{M}{2}\left(M\varepsilon^a{}_a\varepsilon^b{}_b + \varepsilon^a{}_b\varepsilon^b{}_a\right) + o(\varepsilon^2)\,. \tag{241}$$

Услов да би овај ред конвергирао је да је $|E| < 1$, док сабирци у реду почињу да опадају када је задовољен следећи услов:

$$|E| < \frac{n+1}{M+n+1}\,. \tag{242}$$

Одатле следи да ће допринос другог реда бити већи од доприноса трећег реда када је $|E| < 0.0196$ (уз претпоставку да је $n = 2$, $M = 150$), или, другим речима, ако се захтева да допринос другог реда буде k пута већи од доприноса трећег реда, онда је $|E| < 0.0196/k$. Такође у случају позитивног степена детерминанте тетраде, слично се добија:

$$|e|^M = (1+E)^M = \sum_{n=0}^{+\infty} \binom{M}{n} E^n = 1 + M\varepsilon^a{}_a + \frac{M}{2}\left(M\varepsilon^a{}_a\varepsilon^b{}_b - \varepsilon^a{}_b\varepsilon^b{}_a\right) + o(\varepsilon^2)\,. \tag{243}$$

Овај ред је коначан, будући да је биномни коефицијент једнак нули када је $n > M$, па не постоје проблеми са конвергенцијом. Додатно, сабирци почињу да опадају када је испуњен услов:

$$|E| < \frac{n+1}{M-n}\,, \tag{244}$$

који је слабији захтев од претходног (242) и самим тим аутоматски задовољен.

Применом једначина за везу између теорија (226) на опсерваблу пертурбације метрике, $F(\phi_k) = h_{\mu\nu}$, добија се:

$$\langle h_{\mu\nu}\rangle_{3BF} = \frac{\langle h_{\mu\nu}\rangle_{ECC} - M\,\langle \varepsilon^a{}_a h_{\mu\nu}\rangle_{ECC}}{1 - M\,\langle\varepsilon^a{}_a\rangle_{ECC} + \frac{M}{2}\,\langle M\varepsilon^a{}_a\varepsilon^b{}_b + \varepsilon^a{}_b\varepsilon^b{}_a\rangle_{ECC}}\,. \tag{245}$$

Развој имениоца у ред даје:

$$\begin{aligned}\langle h_{\mu\nu}\rangle_{3BF} &= \left(\langle h_{\mu\nu}\rangle_{ECC} - M\,\langle\varepsilon^a{}_a h_{\mu\nu}\rangle_{ECC}\right)\\ &\quad \left(1 + M\,\langle\varepsilon^a{}_a\rangle_{ECC} - \frac{M}{2}\left\langle M\varepsilon^a{}_a\varepsilon^b{}_b + \varepsilon^a{}_b\varepsilon^b{}_a\right\rangle_{ECC} + \frac{M^2}{2}\,\langle\varepsilon^a{}_a\rangle^2_{ECC}\right)\\ &= \langle h_{\mu\nu}\rangle_{ECC}\left(1 + M\,\langle\varepsilon^a{}_a\rangle_{ECC}\right) - M\,\langle\varepsilon^a{}_a h_{\mu\nu}\rangle_{ECC}\,.\end{aligned} \tag{246}$$

Такође, у супротном смеру се добија:

$$\begin{aligned}\langle h_{\mu\nu}\rangle_{ECC} &= \left(\langle h_{\mu\nu}\rangle_{3BF} + M\,\langle\varepsilon^a{}_a h_{\mu\nu}\rangle_{3BF}\right)\\ &\quad \left(1 - M\,\langle\varepsilon^a{}_a\rangle_{3BF} - \frac{M}{2}\left\langle M\varepsilon^a{}_a\varepsilon^b{}_b - \varepsilon^a{}_b\varepsilon^b{}_a\right\rangle_{3BF} - \frac{M^2}{2}\,\langle\varepsilon^a{}_a\rangle^2_{3BF}\right)\\ &= \langle h_{\mu\nu}\rangle_{3BF}\left(1 - M\,\langle\varepsilon^a{}_a\rangle_{3BF}\right) + M\,\langle\varepsilon^a{}_a h_{\mu\nu}\rangle_{3BF}\,.\end{aligned} \tag{247}$$

Ови изрази се поједностављују у следећи облик изражен преко коваријансе:

$$\langle h_{\mu\nu} \rangle_{3BF} = \langle h_{\mu\nu} \rangle_{ECC} - M\mathrm{Cov}\left(\varepsilon^a{}_a\,, h_{\mu\nu}\right)_{ECC}\,, \tag{248}$$

$$\langle h_{\mu\nu} \rangle_{ECC} = \langle h_{\mu\nu} \rangle_{3BF} + M\mathrm{Cov}\left(\varepsilon^a{}_a\,, h_{\mu\nu}\right)_{3BF}\,. \tag{249}$$

Коришћењем неједнакости Коши-Шварца (231), може се проценити потребна величина пертурбације да би се разлика између квантне 3BF теорије и квантне ECC теорије експериментално утврдила. Конкретно, до другог реда је разлика у предвиђању између теорија

$$\mathrm{Cov}\left(\varepsilon^a{}_a\,, h_{\mu\nu}\right)_{ECC} = \mathrm{Cov}\left(\varepsilon^a{}_a\,, h_{\mu\nu}\right)_{3BF} = \mathrm{Cov}\left(\varepsilon^a{}_a\,, h_{\mu\nu}\right)\,, \tag{250}$$

односно величина квантне поправке је иста у обе теорије, што се може добити комбиновањем једначина (248) и (249). Осим тога, на основу релација $h_{\mu\nu} = \eta_{\mu a}\varepsilon^a{}_\nu + \eta_{a\nu}\varepsilon^a{}_\mu$, $\varepsilon_{\mu\nu} = \eta_{\mu a}\varepsilon^a{}_\nu$ и неједнакости Коши-Шварца, уз претпоставку да је неодређеност сваке компоненте тетраде приближно иста, добија се:

$$\langle h_{\mu\nu} \rangle_{ECC} - \langle h_{\mu\nu} \rangle_{3BF} = M\mathrm{Cov}\left(\varepsilon^a{}_a\,, h_{\mu\nu}\right) \leqslant 2M\Delta\varepsilon^a{}_a\,\Delta\varepsilon_{\mu\nu} \approx 8M\left(\Delta\varepsilon_{\mu\nu}\right)^2 \approx 2M\left(\Delta h_{\mu\nu}\right)^2\,. \tag{251}$$

Неједнакост (251) се у принципу може искористити да се експериментално разликују квантна 3BF теорија и квантна ECC теорија, мерењем гравитационих таласа и упоређивањем са исходом теоријских предвиђања. У циљу стицања интуиције о потребном реду величине, може се кренути од оквирних редова величине измерених тренутно доступном технологијом, на пример ЛИГО/Вирго детекторима. На основу [56], уобичајена прецизност мерења напрезања (компонената $\langle h_{\mu\nu} \rangle$) се може проценити на 10^{-21}, што имплицира да десна страна (251) треба да буде

$$2M\left(\Delta h_{\mu\nu}\right)^2 \geqslant 10^{-21}\,. \tag{252}$$

Узимајући у обзир да је за Стандардни Модел вредност $M = 150$ (219), добија се процена минималне квантне поправке која се може детектовати:

$$\Delta h_{\mu\nu} \geqslant \sqrt{\frac{10^{-21}}{2\cdot 150}} \approx 10^{-12}\,. \tag{253}$$

Ово је велика вредност, што се може видети из чињенице да је амплитуда напрезања сигнала спајања две црне рупе у [56] реда 10^{-18}. Пошто је растојање до извора GW150914 процењено на $r_{GW} \approx 410\,\mathrm{Mpc}$, што је далеко изван наше галаксије, може се проценити интензитет напрезања у случају истог догађаја унутар Млечног Пута, односно на растојању унутар $r_{MW} \approx 34\,\mathrm{Kpc}$. Пошто амплитуда сферног таласа опада пропорционално са $1/r$ од извора, може се проценити да би спајање сличних црних рупа унутар наше галаксије створило сигнал чији је ред величине напрезања

$$h_{MW} \approx h_{GW}\frac{r_{GW}}{r_{MW}} = 10^{-18}\times\frac{4.1\cdot 10^5\,\mathrm{Kpc}}{3.4\cdot 10^1\,\mathrm{Kpc}} \approx 10^{-14}\,. \tag{254}$$

Ово је и даље два реда величине мање од потребне величине квантне поправке $\Delta h_{\mu\nu}$. Штавише, не постоји разлог у теорији због ког би систем две црне рупе које се спајају требао да има тако велику квантну неодређеност.

Другим речима, да би постојећом технологијом разлика између две теорије била опсервабилна, потребно је да извор гравитационих таласа

(a) ствара напрезање $\langle h_{\mu\nu} \rangle$ реда величине најмање 10^{-11}, и

(b) има квантну неодређеност напрезања, $\Delta h_{\mu\nu}$, најмање 10^{-12}.

Очигледно, не постоје познати кандидати за изворе таквих гравитационих таласа у природи. У сваком случају, макар у теорији, ако би такав извор постајао, било би могуће применити (251) и експериментално разликовати квантну 3BF теорију од квантне ЕСС теорије.

5 Конструкција квантне 3BF теорије са везама

У претходним главама је потврђена еквиваленција између две класичне теорије описане 3BF дејством са везама и Стандардним Моделом спрегнутим са Ајнштајн-Картановом гравитацијом. Затим је утврђена непертурбативна веза између одговарајућих квантних теорија заснованих на овим моделима без експлицитне дефиниције сваке од теорија и утврђено да обе теорије деле заједнички семикласичан лимес. У наредним поглављима ће бити формулисана квантна 3BF теорија са везама на триангулацији и биће спроведена прелиминарна анализа семикласичног лимеса добијене теорије.

5.1 Дискретизација дејства

У овом поглављу ће бити изложен поступак дефинисања дејства на триангулацији многострукости. Ова структура заправо представља део по део равну многострукост састављену од ћелија које имају геометрију 4-симплекса. Основна идеја је у придруживању чинилаца дејства елементима триангулације истог типа. Овде се могу уочити две основне могућности које се често срећу у литератури. Прва је да су чиниоци дејства који су n-форме додељени n-димензионалним ћелијама директне триангулације, и друга, да су n-форме из дејства додељене n-димензионалним ћелијама Поенкаре-дуалној триангулацији. У случају када је дејство састављено од сабирака који садрже по два чиниоца, дискретизација дејства се може вршити на комбинацији директне и Поенкаре-дуала триангулације, по неком утврђеном правилу који од два чиниоца у сабирцима у дејству треба поставити на коју триангулацију. Међутим, у случају када се сабирци у дејству састоје од више чинилаца, а поготово када број чинилаца није исти за све сабирке, постављање дејства на комбинацију директне и Поенкаре-дуалне триангулације нема превише смисла јер не постоји природан начин који указује на то која поља треба доделити којој триангулацији. Због тога је у наставку цело дејство постављено само на директну триангулацију без Поенкареовог дуала.

Сада треба увести нотацију која ће означавати елементе триангулације. Структура елементарне ћелије триангулације, 4-симплекса је следећа. Симплекс се састоји од пет вертекса означених и пребројаних малим словом v, по десет ивица и троуглова, означених са ε и Δ, редом, и пет тетраедара означених са τ. Сами симплекси се пребројавају и означавају малим словом σ. Дејство је задато као сума по елементима триангулације, а бројач у суми је ознака за елемент триангулације који се први појављује у исказу испод суме. Тај исказ може имати два типа облика $a \in b$ и $b \ni a$, где су a и b елементи триангулације било ког горе наведеног типа.

Сваки елемент триангулације осим вертекса је оријентисан. Елемент најмање димензије који има оријентацију је свакако ивица. Сви остали елементи веће димензије се могу разложити на ивице, а њихова оријентација дефинисати редоследом навођења и оријентацијом њихових ивица. Због тога је важно размотрити на колико се начина елементи веће димензије могу разложити на елементе мање димензије. Ове могућности се могу представити табеларно:

Елемент	начин за разлагање	број комбинација
△	$2 \times$ /	$3 \times 2! = 6$
◬	$3 \times$ /	$16 \times 3! = 96$
◬	/ $+$ △	12
⛤	$4 \times$ /	$125 \times 4! = 3000$
⛤	$2 \times$ / $+$ △	$150 \times 2! = 300$
⛤	/ $+$ ◬	20
⛤	$2 \times$ △	$15 \times 2! = 30$

У свакој врсти је број одређен као број начина за избор различитих елемената помножен пермутацијом редоследа елемената истог типа.

У првој врсти, у случају троугла, две од три ивице се могу одабрати на три начина, док је број пермутација редоследа њиховог набрајања 2, па је укупан број комбинација 6.

Тетраедар се може конструисати на два начина, први је узимањем три ивице које не припадају истом троуглу. Пребројавање се може извршити на два начина. Први је према топологији графа сачињеног од изабраних ивица. Постоје две топологије, прва код које све три ивице излазе из истог вертекса, и таквих има четири јер тетраедар садржи четири вертекса, и други где се три ивице надовезују једна на другу, који се може конструисати тако што је први вертекс у низу изабран на четири начина, други на три и трећи на два што је укупно 24, али пошто је граф симетричан смер бројања, има их дупло мање односно 12. Ово може да се изброји и тако што се из претходног графа са три ивице које излазе из истог вертекса који се може изабрати на 4 начина, изабере једна ивица на три начина која се извлачи из заједничког вертекса и повезује са неким од друга два вертекса. Дакле, укупно постоји 4+12=16 различитих графова са 3! пермутација, односно 96 могућности набрајања 3 ивице које чине тетраедар.

Други начин је да се прва ивица може изабрати на 6 начина, друга ивица на 5, тако да у четири случаја ове две ивице имају заједнички вертекс и у једном случају немају. Трећа ивица се сада може изабрати на три начина ако су прве две имале заједнички вертекс и на четири начина ако нису. То је укупно $6 \cdot (4 \cdot 3 + 1 \cdot 4) = 96$ начина.

Други начин за конструкцију тетраедра је троугао и ивица која му не припада. Троугао се може изабрати на 4 начина, а ивица која му не припада на три, тако да је то укупно 12 могућности.

Цео 4-симплекс се може конструисати на 4 начина. Први је помоћу 4 ивице. Бројање могућности поново може да се обави на два начина. Посматрањем топологија, уочавају се три могуће топологије. Прва је четири ивице са заједничким вертексом који се може изабрати на 5 начина. Друга је три ивице из заједничког вертекса који се може изабрати на пет начина и једна ивица која спаја пети вертекс која се може изабрати на 4 начина са неким од три вертекса ивица које вире из првог заједничког вертекса. Таквих графова има $5 \cdot 4 \cdot 3 = 60$. Трећи граф је састављен од 4 ивице које се надовезују једна на другу. Сваки вертекс у низу се може одабрати на за један мање начин од претходног. Међутим, како је цео граф независан од смера бројања, укупан број мора да се подели са два, тако да таквих графова укупно има $5 \cdot 4 \cdot 3 \cdot 2/2 = 60$. Дакле постоји 125 графова са 4! пермутација ивица, што је укупно 3000 могућности.

Други начин за бројање је тај да се прва ивица може изабрати на 10 начина. Друга ивица на 9, од чега у шест случаја постоји заједнички вертекс и у три случаја не постоји. Сада се посматрају оба случаја одвојено. Трећа ивица се у првом случају може изабрати на седам начина, од којих у шест постоји један вертекс који је заједнички са бар једном од претходне две ивице и један начин тако да не постоји заједнички вертекс. Ако је постајао један заједнички вертекс за трећу и бар једну од претходних ивица, постоје две могућности. Да је вертекс заједнички за све три ивице и тада се трећа ивица може изабрати на два, а четврта на четири начина ($10\cdot6\cdot2\cdot4{=}180$ могућности), и да трећа ивица има заједнички вертекс само са једном од прве две ивице, када се трећа ивица може изабрати на 4, а четврта на још 4 начина ($10\cdot6\cdot4\cdot4{=}960$ могућности). Ако не постоји заједнички вертекс за трећу и претходне две ивице, трећа ивица се може изабрати на само један начин, а четврта може да полази из два вертекса треће ивице и заврши се у неки од три преостала вертекса, односно може се изабрати на $2\cdot3{=}6$ начина ($10\cdot6\cdot1\cdot6{=}360$ могућности). Други случај је када прва и друга ивица немају заједнички вертекс. Тада се друга ивица може изабрати на 3 начина, а трећа ивица може да се изабере на 8 начина. Четири од осам су такви да се трећа ивица конструише тако што јој је један вертекс заједнички са првом ивицом (2 начина), а други заједнички са другом (још 2 начина), или тако што јој је један вертекс пети, који не припада ни првој ни другој ивици, а други неки од преостала 4 вертекса (преостале 4 могућности). У колико је трећа ивица добијена избором једног вертекса прве и једног вертекса друге, четврта ивица се може изабрати на четири начина. Тако да јој је један вертекс пети који не припада ни једној од прве три ивице, а други, неки од преостала четири ($10\cdot3\cdot4\cdot4{=}480$ могућности). Последњи начин је да је трећа ивица добијена спајањем петог вертекса који није припадао ни једној од прве две ивице са неким од преосталих 4 вертекса и ту се четврта ивица може изабрати на 6 начина. Пре одабирања четврте ивице, граф се састоји од два неповезана дела. Први је једна засебна ивица (два вертекса), а други су две повезане ивице (три вертекса). Четврта ивица може да се изабере тако да јој је један вертекс заједнички са једном неповезаном ивицом (два начина), а други заједнички са бар једном од друге две ивице (три начина), тако да се добије повезан граф ($10\cdot3\cdot4\cdot6{=}720$ могућности). Укупно то чини $480{+}960{+}360{+}480{+}720{=}3000$ могућности.

Други начин за конструисање целог 4-симплекса је троугао и две ивице које му не припадају, тако да све заједно садрже свих пет вертекса. Троугао се може изабрати на 10 начина, а ивице на такав начин да свака садржи бар један од преостала два вертекса која не припадају троуглу. Ако прва ивица садржи само један од преостала два вертекса (може се изабрати на 6 начина), друга ивица се може изабрати на 4 начина, а ако прва ивица садржи оба преостала вертекса (један начин), друга ивица се може изабрати на шест начина. То је укупно $10\cdot(6\cdot4{+}1\cdot6){=}300$ могућности. Алтернативно се могућности могу избројати посматрањем графова. Ако су обе ивице причвршћене једним вертексом за троугао, то је 9 могућности, а ако је само једна причвршћена за троугао, то је 6 могућности. Број пермутација ивица је 2, па је укупан број могућности $10\cdot(9{+}6)\cdot2{=}300$ могућности.

Трећи начин за конструкцију целог 4-симплекса је тетраедар и ивица која му не припада. Тетраедар се може изабрати на 5 начина, а ивица која му не припада на четири, што је укупно 20 могућности.

Четврти и последњи начин за конструкцију 4-симплекса је помоћу два троугла са само једним заједничким вертексом. Овај вертекс се може изабрати на 5 начина, а парови троуглова на три. Број пермутација троуглова је 2, па је укупно $5\cdot3\cdot2{=}30$ могућности.

Ови комбинаторни фактори су важни у тренутку преласка са суме по независним елементима 4-симплекса на суму по свим елемнтима 4-симплекса.

У наставку ће бити изложен поступак постављања дејства на директну триангулацију. Овај поступак се не може спровести са дејством у произвољном облику, већ сабирци у дејству морају да буду изражени у облику производа диференцијалних форми.

Процес постављања дејства на триангулацију је изражен помоћу контракције између вектора и форми. Пошто је многострукост део по део равна, симплекси од којих је састављена су подударни са тангентном многострукошћу, па се интеграција форме по симплексу може вршити контракцијом између форми на многострукости и $\binom{4}{0}$ тензора придруженом симплексу:

$$\sigma = \frac{1}{4!}\varepsilon_1^\mu \varepsilon_2^\nu \varepsilon_3^\rho \varepsilon_4^\sigma \partial_\mu \wedge \partial_\nu \wedge \partial_\rho \wedge \partial_\sigma \,, \tag{255}$$

где су $\varepsilon_1, \ldots, \varepsilon_4$ четири линеарно независне ивице које образују 4-симплекс. Сам тензор придружен симплексу не зависи од избора ових независнох ивица, али му је знак осетљив на редослед набрајања ивица због антисиметричног производа између њих.

Да би дејство било дискретизовано овим методом, потребно је конструисати Лагранжијан у облику 4-форме. Дискретизација дејства је онда:

$$\begin{aligned}
S[\phi] &= \int \frac{1}{4!}\mathcal{L}_{\mu\nu\rho\sigma}\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \wedge \mathrm{d}x^\rho \wedge \mathrm{d}x^\sigma = \sum_\sigma \int_\sigma \frac{1}{4!}\mathcal{L}_{\mu\nu\rho\sigma}\varepsilon^{\mu\nu\rho\sigma} d^4x = \frac{1}{4!}\sum_\sigma \mathcal{L}_{\mu\nu\rho\sigma}[\sigma]\varepsilon^{\mu\nu\rho\sigma}V[\sigma] \\
&= \frac{1}{4!}\sum_\sigma \mathcal{L}_{\mu\nu\rho\sigma}[\sigma]\varepsilon^{\mu\nu\rho\sigma}\left(-\frac{1}{4!}\varepsilon_{1a}\varepsilon_{2b}\varepsilon_{3c}\varepsilon_{4d}\varepsilon^{abcd}\right) \\
&= \frac{1}{4!}\sum_\sigma \mathcal{L}_{\mu\nu\rho\sigma}[\sigma]\varepsilon^{\mu\nu\rho\sigma}\Bigg(-\frac{1}{4!}e_{a\mu'}e_{b\nu'}e_{c\rho'}e_{d\sigma'} \\
&\times \frac{1}{4!}\varepsilon_1^\alpha\varepsilon_2^\beta\varepsilon_3^\gamma\varepsilon_4^\delta\varepsilon^{abcd}\mathrm{d}x^{\mu'} \wedge \mathrm{d}x^{\nu'} \wedge \mathrm{d}x^{\rho'} \wedge \mathrm{d}x^{\sigma'}\left[\partial_\alpha \wedge \partial_\beta \wedge \partial_\gamma \wedge \partial_\delta\right]\Bigg) \\
&= \frac{1}{4!}\sum_\sigma \mathcal{L}_{\mu\nu\rho\sigma}\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \wedge \mathrm{d}x^\rho \wedge \mathrm{d}x^\sigma \frac{1}{4!}\varepsilon_1^\alpha\varepsilon_2^\beta\varepsilon_3^\gamma\varepsilon_4^\delta\left[\partial_\alpha \wedge \partial_\beta \wedge \partial_\gamma \wedge \partial_\delta\right] \\
&= \sum_\sigma \mathcal{L}[\sigma]\,.
\end{aligned} \tag{256}$$

Сабирци у дејству које је потребно интегралити по елементима триангулације могу бити у укупно четири облика. Резултат интеграције у сва четири случаја је следећи.

Први случај, 4-симплекс састављен од четири ивице:

$$\begin{aligned}
\int \alpha_a \wedge \alpha_b \wedge \alpha_c \wedge \alpha_d &= \sum_\sigma \alpha_a \wedge \alpha_b \wedge \alpha_c \wedge \alpha_d[\sigma] \\
&= \sum_\sigma \alpha_{a\mu}\alpha_{b\nu}\alpha_{c\rho}\alpha_{d\sigma}\frac{1}{4!}\varepsilon_1^\lambda\varepsilon_2^\chi\varepsilon_3^\varphi\varepsilon_4^\xi \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \wedge \mathrm{d}x^\rho \wedge \mathrm{d}x^\sigma\left[\partial_\lambda \wedge \partial_\chi \wedge \partial_\varphi \wedge \partial_\xi\right] \\
&= \frac{1}{4!}\sum_\sigma \alpha_{a\mu}\alpha_{b\nu}\alpha_{c\rho}\alpha_{d\sigma}\varepsilon_1^\lambda\varepsilon_2^\chi\varepsilon_3^\varphi\varepsilon_4^\xi\left(\delta_\lambda^\mu\delta_\chi^\nu\delta_\varphi^\rho\delta_\xi^\sigma + (23)\right) \\
&= \frac{1}{3000}\sum_\sigma \sum_{\varepsilon_1,\ldots,\varepsilon_4\in\sigma} \alpha_a[\varepsilon_1]\alpha_b[\varepsilon_2]\alpha_c[\varepsilon_3]\alpha_d[\varepsilon_4]W[\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4]\,.
\end{aligned} \tag{257}$$

Сабирак $+(23)$ у загради означава преосталих 23 сабирака који садрже пермутацију индекса са одговарајућим предзнаком, док је функција $W[\ldots]$ функција знака која ће бити дефинисана касније. Други случај, 4-симплекс састављен од две ивице и троугла:

$$\begin{aligned}\int \alpha_a\wedge\alpha_b\wedge\beta_c &= \sum_\sigma \alpha_a\wedge\alpha_b\wedge\beta_c[\sigma]=\sum_\sigma \alpha_{a\mu}\alpha_{b\nu}\frac{1}{2}\beta_{c\rho\sigma}\frac{1}{4!}\varepsilon_1^\lambda\varepsilon_2^\chi\varepsilon_3^\varphi\varepsilon_4^\xi\left(\delta^\mu_\lambda\delta^\nu_\chi\delta^\rho_\varphi\delta^\sigma_\xi+(23)\right)\\ &= \frac{1}{3000}\sum_\sigma\sum_{\varepsilon_1,\ldots,\varepsilon_4\in\sigma}\alpha_a[\varepsilon_1]\alpha_b[\varepsilon_2]\frac{1}{2}\beta_{c\rho\sigma}\varepsilon_3^\rho\varepsilon_4^\sigma W[\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4]\\ &= \frac{1}{300}\sum_\sigma\sum_{\varepsilon_1,\varepsilon_2,\Delta\in\sigma}\alpha_a[\varepsilon_1]\alpha_b[\varepsilon_2]\beta_c[\Delta]W[\varepsilon_1,\varepsilon_2,\Delta]\,. \end{aligned}\tag{258}$$

Трећи случај, 4-симплекс састављен од два троугла:

$$\begin{aligned}\int\beta_a\wedge\beta_b &= \sum_\sigma\frac{1}{2}\beta_{a\mu\nu}\frac{1}{2}\beta_{b\rho\sigma}\frac{1}{4!}\varepsilon_1^\lambda\varepsilon_2^\chi\varepsilon_3^\varphi\varepsilon_4^\xi\left(\delta^\mu_\lambda\delta^\nu_\chi\delta^\rho_\varphi\delta^\sigma_\xi+(23)\right)\\ &= \frac{1}{30}\sum_\sigma\sum_{\Delta_1,\Delta_2\in\sigma}\beta_a[\Delta_1]\beta_b[\Delta_2]W[\Delta_1,\Delta_2]\,.\end{aligned}\tag{259}$$

Четврти случај, 4-симплекс састављен од тетраедра и ивице:

$$\int\alpha_a\wedge\gamma_b=\sum_\sigma\alpha_{a\mu}\frac{1}{3!}\gamma_{b\nu\rho\sigma}\frac{1}{4!}\varepsilon_1^\lambda\varepsilon_2^\chi\varepsilon_3^\varphi\varepsilon_4^\xi\left(\delta^\mu_\lambda\delta^\nu_\chi\delta^\rho_\varphi\delta^\sigma_\xi+(23)\right)=\frac{1}{20}\sum_\sigma\sum_{\varepsilon,\tau\in\sigma}\alpha_a[\varepsilon]\gamma_b[\tau]W[\varepsilon,\tau]\,. \tag{260}$$

Осим 1, 2, 3 и 4-форми, у дејству се појављују и нула-форме које стављањем на триангулацију живе у вертексима, па је њихова интеграција по елементу триангулације тривијална (вредност поља на вертексу је једнака вредности поља у тачки у којој се налази вертекс). Додавање ових поља на триангулацију се врши независно од интеграције форми, па производ скаларних ϕ_i поља постаје:

$$\int\phi_1\ldots\phi_n\,F=\sum_\sigma F[\sigma]\,\frac{1}{5^n}\sum_{v_1,\ldots,v_n\in\sigma}\phi_1[v_1]\ldots\phi_n[v_n]\,, \tag{261}$$

где је F произвољна 4-форма.

У локално инерцијалном референтном систему се тетрада своди на општу координатну трансформацију заједничку за све елементе појединачног 4-симплекса. Деловање форме тетраде на тензор 4-симплекса и последица захтева да компоненте тетраде представљају компоненте матрице опште координатне трансформације $GL(4,\mathbb{R})$ даје везу:

$$e^a[\varepsilon]=e^a{}_\mu\varepsilon^\mu=\Gamma^a{}_\mu\varepsilon^\mu=\varepsilon^a\,, \tag{262}$$

где је $\Gamma^\mu{}_a$ степен слободе фиксирања калибрације.

Вредности квадрата запремина елемената триангулације $l^2[\varepsilon]$, $A^2[\Delta]$, ${}^{(3)}V^2[\tau]$ и ${}^{(4)}V^2[\sigma]$ су реалне функције компонената вектора ивица триангулације дате као:

$${}^{(0)}V^2[v] = -\frac{1}{4!}\frac{1}{(0!)^2}\varepsilon_{\mu\nu\rho\sigma}\varepsilon_{\lambda\xi\theta\phi}\eta^{\mu\lambda}\eta^{\nu\xi}\eta^{\rho\theta}\eta^{\sigma\phi}=1\,, \tag{263}$$

$$l^2[\varepsilon] = -\frac{1}{3!}\frac{1}{(1!)^2}\varepsilon_{\mu\nu\rho\sigma}\varepsilon_1^\mu\varepsilon_{\lambda\xi\theta\phi}\varepsilon_1^\lambda\eta^{\nu\xi}\eta^{\rho\theta}\eta^{\sigma\phi}=\varepsilon_1^\mu\varepsilon_1^\nu\eta_{\mu\nu}=\varepsilon_1^2\,, \tag{264}$$

$$A^2[\Delta] = -\frac{1}{2!}\frac{1}{(2!)^2}\varepsilon_{\mu\nu\rho\sigma}\varepsilon_1^\rho\varepsilon_2^\sigma\varepsilon_{\lambda\xi\theta\phi}\varepsilon_1^\theta\varepsilon_2^\phi\eta^{\mu\lambda}\eta^{\nu\xi}=\frac{1}{4}\left(\varepsilon_1^2\varepsilon_2^2-(\varepsilon_1^\mu\varepsilon_2^\nu\eta_{\mu\nu})^2\right)\,, \tag{265}$$

$${}^{(3)}V^2[\tau] = -\frac{1}{1!}\frac{1}{(3!)^2}\varepsilon_{\mu\nu\rho\sigma}\varepsilon_1^\nu\varepsilon_2^\rho\varepsilon_3^\sigma\varepsilon_{\lambda\xi\theta\phi}\varepsilon_1^\xi\varepsilon_2^\theta\varepsilon_3^\phi\eta^{\mu\lambda}\,, \tag{266}$$

$${}^{(4)}V^2[\sigma] = -\frac{1}{0!}\frac{1}{(4!)^2}\left(\varepsilon_{\mu\nu\rho\sigma}\varepsilon_1^\mu\varepsilon_2^\nu\varepsilon_3^\rho\varepsilon_4^\sigma\right)^2\,. \tag{267}$$

Пошто су вредности $l^2[\varepsilon]$, $A^2[\Delta]$, ${}^{(3)}V^2[\tau]$ и ${}^{(4)}V^2[\sigma]$ независне од овог степена слободе, вектори ивица се могу заменити векторима ивица у сопственом референтном систему ε^a и запремине се неће променити. Због тога сабирак који садржи производ четири тетраде одговара оријентисаној запремини 4-симплекса:

$$\int \frac{1}{4!}\varepsilon_{abcd}e^a \wedge e^b \wedge e^c \wedge e^d = \sum_\sigma \frac{1}{4!}\varepsilon_{abcd}\varepsilon_1^a\varepsilon_2^b\varepsilon_3^c\varepsilon_4^d = -\sum_\sigma V[\sigma]\,. \tag{268}$$

Функција $W[\dots]$ се може дефинисати коришћењем претходне особине. Пошто је триангулација оријентисан граф, сваком троуглу се могу доделити две ивице чијим је редоследом он оријентисан, $\Delta = (\varepsilon_{\Delta 1}, \varepsilon_{\Delta 2})$, сваком тетраедру три ивице чијим је редоследом навођења задата његова оријентација $\tau = (\varepsilon_{\tau 1}, \varepsilon_{\tau 2}, \varepsilon_{\tau 3})$ и сваком 4-симплексу четири ивице $\sigma = (\varepsilon_{\sigma 1}, \varepsilon_{\sigma 2}, \varepsilon_{\sigma 3}, \varepsilon_{\sigma 4})$. Функција $W[\dots]$ је сада количник оријентисане 4-запремине над елементима 4-симплекса који су аргументи функције подељени оријентисаном запремином тог 4-симплекса. Случајеви ове функције су:

$$W[\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4] = -\frac{1}{4!V[\sigma]}\varepsilon_{abcd}\varepsilon_1^a\varepsilon_2^b\varepsilon_3^c\varepsilon_4^d\,, \tag{269}$$

$$W[\Delta, \varepsilon_1, \varepsilon_2] = -\frac{1}{4!V[\sigma]}\varepsilon_{abcd}\varepsilon_{\Delta 1}^a\varepsilon_{\Delta 2}^b\varepsilon_1^c\varepsilon_2^d\,, \tag{270}$$

$$W[\Delta, \tilde{\Delta}] = -\frac{1}{4!V[\sigma]}\varepsilon_{abcd}\varepsilon_{\Delta 1}^a\varepsilon_{\Delta 2}^b\varepsilon_{\tilde{\Delta} 1}^c\varepsilon_{\tilde{\Delta} 2}^d\,, \tag{271}$$

$$W[\varepsilon, \tau] = -\frac{1}{4!V[\sigma]}\varepsilon_{abcd}\varepsilon^a\varepsilon_{\tau 1}^b\varepsilon_{\tau 2}^c\varepsilon_{\tau 3}^d\,. \tag{272}$$

Пошто ова функција аутоматски враћа вредност нула када су јој аргументи линеарно зависни, није потребно стављати никакве додатне услове у дефиницији суме осим захтева да аргументи припадају истом 4-симплексу.

Први део дефиниције функционалног интеграла је његова мера. У наставку ће бити изложена два метода дискретизације мере. Мера интеграла на триангулацији се дефинише као интеграл по вредности поља на елементу триангулације. Експлицитна дефиниција је на први начин задата као:

$$D\phi = \prod_v d\phi[v]\,, \quad D\alpha = \prod_\varepsilon d\alpha[\varepsilon]\,, \quad D\beta = \prod_\Delta d\beta[\Delta]\,, \quad D\gamma = \prod_\tau d\gamma[\tau]\,, \quad D\delta = \prod_\sigma d\delta[\sigma]\,. \tag{273}$$

Алтернативно је уместо екстензивних варијабли могуће изабрати интензивне. Ако се произвољна n-форма запише као:

$$F = \frac{1}{n!}F_{a_1\dots a_n}e^{a_1} \wedge \dots \wedge e^{a_n}\,, \tag{274}$$

контракција форме $e^{a_1} \wedge \dots \wedge e^{a_n}$ са n-симплекс тензором σ_n даје:

$$e^{a_1} \wedge \dots \wedge e^{a_n}[\sigma_n] = V[\sigma_n]E^{a_1\dots a_n}\,, \tag{275}$$

па је контракција полазне n-форме F једнака:

$$F[\sigma_n] = \frac{1}{n!}F_{a_1\dots a_n}[\sigma_n]E^{a_1\dots a_n}V[\sigma_n] = \tilde{F}[\sigma_n]V[\sigma_n]\,. \tag{276}$$

Мера интеграла је у овим промењљивим једнака:

$$D\phi = \prod_v d\phi[v]\,, \quad D\alpha = \prod_\varepsilon l[\varepsilon] d\alpha[\varepsilon]\,, \quad D\beta = \prod_\Delta A[\Delta] d\beta[\Delta]\,,$$
$$D\gamma = \prod_\tau V[\tau] d\gamma[\tau]\,, \quad D\delta = \prod_\sigma V[\sigma] d\delta[\sigma]\,. \tag{277}$$

Будући да је веза између ове две дефиниције смена промењљивих, резултат интеграције не би требало да зависи од конкретног избора, па ће у наставку бити коришћена прва дефиниција задата у (273).

Следећи корак у триангулацији дејства представља дефинисање извода на триангулацији. Пошто је многострукост састављена од елемената као што су вертекси, ивице, површине итд. извод на објекту n-тог реда треба дефинисати преко објеката $(n-1)$-ог реда, односно извод на објекту триангулације је функција над границом тог објекта. Због тога је извод поља на триангулацији дефинисан применом Стоксове теореме:

$$\int_{\mathcal{M}_n} \mathrm{d}A = \oint_{\partial\mathcal{M}_n} A \implies \sum_{\sigma_n \in T(\mathcal{M})} \mathrm{d}A\,[\sigma_n] = \sum_{\sigma_n \in T(\mathcal{M})} \oint\!\!\!\!\!\!\sum_{\sigma_{n-1}\in\partial\sigma_n} A[\sigma_{n-1}]\,. \tag{278}$$

Заокружена сума представља суму елемената по оријентисаној граници, тако да се води рачуна о предзнаку сабирака који сада зависи и од усклађености оријентација сегмената границе $\partial\sigma_n$ и самог σ_n. За ту потребу се уводи функција знака $z[\sigma_n, \sigma_{n-1} \in \partial\sigma_n]$ која враћа вредност ± 1, па заокружена сума постаје:

$$\oint\!\!\!\!\!\!\sum_{\sigma_{n-1}\in\partial\sigma_n} A[\sigma_{n-1}] = \sum_{\sigma_{n-1}\in\partial\sigma_n} z[\sigma_n, \sigma_{n-1}] A[\sigma_{n-1}]\,. \tag{279}$$

Вредност функције знака зависи од оријентација елемената триангулације над којима се врши. Аргументи нису произвољне комбинације два елемента триангулације, већ други мора да буде део границе првог и самим тим за једну димензију мањи. Свака ивица се може представити преко два вертекса $\varepsilon = (v_{\varepsilon 1}, v_{\varepsilon 2})$, тако да је оријентација ивице усмерена од другог ка првом вертексу. Функција знака онда враћа вредност:

$$z[\varepsilon, v] = \delta_{v_{\varepsilon 1}, v} - \delta_{v_{\varepsilon 2}, v}\,. \tag{280}$$

Троугао је задат као уређени пар ивица које полазе из истог вертекса $\Delta = (\varepsilon_{\Delta 1}, \varepsilon_{\Delta 2})$. Његова оријентација је задата редоследом набрајања ивица, па је вредност функције знака:

$$z[\Delta, \varepsilon] = \frac{\varepsilon^a_{\Delta 1}\varepsilon_a}{\varepsilon^a\varepsilon_a}\delta_{\varepsilon_{\Delta 1},\varepsilon} - \frac{\varepsilon^a_{\Delta 2}\varepsilon_a}{\varepsilon^a\varepsilon_a}\delta_{\varepsilon_{\Delta 2},\varepsilon} + \frac{(\varepsilon_{\Delta 2} - \varepsilon_{\Delta 1})^a\varepsilon_a}{\varepsilon^a\varepsilon_a}\delta_{\varepsilon_{\Delta 2} - \varepsilon_{\Delta 1},\varepsilon}\,. \tag{281}$$

Тетраедар је задат као уређена тројка ивица које полазе из истог вертекса и може се посматрати као уређени пар троуглова са заједничком ивицом $\tau = (\varepsilon_{\tau 1}\varepsilon_{\tau 2}, \varepsilon_{\tau_3}) = (\Delta_{\tau 1}, \Delta_{\tau_2})$, таквих да је $\Delta_{\tau 1} = (\varepsilon_{\tau 1}, \varepsilon_{\tau 2})$ и $\Delta_{\tau 2} = (\varepsilon_{\tau 2}, \varepsilon_{\tau 3})$. Додатно, зарад једноставнијег записа, потребно је увести дефиницију и преостала два троугла $\Delta_{\tau 3} = (\varepsilon_{\tau 3}, \varepsilon_{\tau 1})$ и $\Delta_{\tau 4} = (\varepsilon_{\tau 1} - \varepsilon_{\tau 2}, \varepsilon_{\tau 3} - \varepsilon_{\tau 2})$. Функција знака је онда:

$$z[\tau, \Delta] = \sum_{j=1}^{4} \frac{\Delta^{ab}_{\tau j}\Delta_{ab}}{\Delta^{ab}\Delta_{ab}}\delta_{\Delta_{\tau j},\Delta}\,, \tag{282}$$

где је:

$$\Delta^{ab} = \frac{1}{2}\left(\varepsilon^a_{\Delta 1}\varepsilon^b_{\Delta 2} - \varepsilon^b_{\Delta 1}\varepsilon^a_{\Delta 2}\right) . \tag{283}$$

Коначно, 4-симплекс је дефинисан као уређена четворка четири ивице које излазе из истог вертекса. Он садржи пет позитивно оријентисаних тетраедара $\tau_{\sigma 1} = (\varepsilon_{\sigma 1}, \varepsilon_{\sigma 2}, \varepsilon_{\sigma 3})$, $\tau_{\sigma 2} = (\varepsilon_{\sigma 1}, \varepsilon_{\sigma 4}, \varepsilon_{\sigma 2})$, $\tau_{\sigma 3} = (\varepsilon_{\sigma 1}, \varepsilon_{\sigma 3}, \varepsilon_{\sigma 4})$, $\tau_{\sigma 4} = (\varepsilon_{\sigma 2}, \varepsilon_{\sigma 4}, \varepsilon_{\sigma 3})$ и $\tau_{\sigma 5} = (\varepsilon_{\sigma 1} - \varepsilon_{\sigma 2}, \varepsilon_{\sigma 2} - \varepsilon_{\sigma 3}, \varepsilon_{\sigma 4} - \varepsilon_{\sigma 3})$. Функција знака је у овом случају:

$$z[\sigma, \tau] = \sum_{j=1}^{5} \frac{\tau^{abc}_{\sigma j}\tau_{abc}}{\tau^{abc}\tau_{abc}} \delta_{\tau_{\sigma j}, \tau} , \tag{284}$$

где је:

$$\tau^{abc} = \frac{1}{3!}\varepsilon^{abcd}\varepsilon_{defg}\varepsilon^e_{\tau 1}\varepsilon^f_{\tau 2}\varepsilon^g_{\tau 3} . \tag{285}$$

Делта између елемената симплекса је дефинисана када су оба елемента истог типа. Ти елементи су дати уређеним n-торкама ивица, а свака ивица је дата као уређени пар вертекса, одакле се сваком елементу симплекса може придружити скуп вертекса над којим је разапет. Делта између два елемента симплекса A и B, разапетих над вертексима v_{Ai} и v_{Bj} је онда дата као перманента матрице делти над вертексима:

$$\delta_{A,B} = \mathrm{perm}\left(\delta_{v_{Ai}, v_{Bj}}\right) . \tag{286}$$

Постојање ове делте у дефиницији функције знака обезбеђује да функција врати вредност нула у колико њен други аргумент није део границе првог аргумента. Због тога је могуће проширити домен сумације по другом аргументу на цео 4-симплекс без промене резултата суме.

Интеграција целе 4-форме која у себи садржи диференцијал је:

$$\int \gamma \wedge \mathrm{d}\phi = \frac{1}{20}\sum_{\sigma}\sum_{\tau,\varepsilon\in\sigma}\gamma[\tau]\sum_{v\in\varepsilon}\phi[v]z[\varepsilon, v]W[\tau,\varepsilon] = \frac{1}{20}\sum_{\sigma}\sum_{\tau,\varepsilon,v\in\sigma}\gamma[\tau]\phi[v]z[\varepsilon, v]W[\tau,\varepsilon] , \tag{287}$$

$$\begin{aligned}\int \beta \wedge \mathrm{d}\alpha &= \frac{1}{30}\sum_{\sigma}\sum_{\Delta_1,\Delta_2\in\sigma}\beta[\Delta_1]\sum_{\varepsilon\in\Delta_2}\alpha[\varepsilon]z[\Delta_2,\varepsilon]W[\Delta_1,\Delta_2] \\ &= \frac{1}{30}\sum_{\sigma}\sum_{\Delta_1,\Delta_2,\varepsilon\in\sigma}\beta[\Delta_1]\alpha[\varepsilon]z[\Delta_2,\varepsilon]W[\Delta_1,\Delta_2] ,\end{aligned} \tag{288}$$

$$\begin{aligned}\int \alpha \wedge \mathrm{d}\beta &= \frac{1}{20}\sum_{\sigma}\sum_{\varepsilon,\tau\in\sigma}\alpha[\varepsilon]\sum_{\Delta\in\tau}\beta[\Delta]z[\tau,\Delta]W[\varepsilon,\tau] \\ &= \frac{1}{20}\sum_{\sigma}\sum_{\varepsilon,\tau,\Delta\in\sigma}\alpha[\varepsilon]\beta[\Delta]z[\tau,\Delta]W[\varepsilon,\tau] ,\end{aligned} \tag{289}$$

$$\int \mathrm{d}\gamma = \sum_{\sigma}\sum_{\tau\in\sigma}\gamma[\tau]z[\sigma,\tau] . \tag{290}$$

Након дефинисања свих правила и случајева који се јављају у процесу дискретизације дејства, могуће је дефинисати очекивану вредност произвољне опсервабле на триангулацији по угледу на (205):

$$
\begin{aligned}
\langle F \rangle &= \frac{\mathcal{N}}{Z} \left(\prod_v \int d\phi[v] d\psi[v] d\bar{\psi}[v] d\Lambda[v] d\zeta[v] dH[v] dM[v] \right. \\
& \prod_\varepsilon \int d\alpha[\varepsilon] d\omega[\varepsilon] de[\varepsilon] d\tilde{\lambda}[\varepsilon] d\lambda[\varepsilon] d\bar{\lambda}[\varepsilon] \\
& \left. \prod_\Delta \int dB[\Delta] d\beta[\Delta] d\lambda[\Delta] \prod_\tau \int d\tilde{\gamma}[\tau] d\gamma[\tau] d\bar{\gamma}[\tau] \right) \\
& \prod_\sigma \left(\exp \left(\frac{i}{30} \sum_{\Delta_1 \in \sigma} B_\alpha[\Delta_1] \left(\sum_{\Delta_2, \varepsilon \in \sigma} \alpha^\alpha[\varepsilon] z[\Delta_2, \varepsilon] W[\Delta_1, \Delta_2] \right. \right. \right. \\
& \left. \left. + \frac{1}{10} \sum_{\varepsilon_1, \varepsilon_2 \in \sigma} \alpha^\beta[\varepsilon_1] \alpha^\gamma[\varepsilon_2] f_{\beta\gamma}{}^\alpha W[\Delta_1, \varepsilon_1, \varepsilon_2] \right) \right) \\
& \exp \left(\frac{i}{30} \sum_{\Delta_1 \in \sigma} B_{[ab]}[\Delta_1] \left(\sum_{\Delta_2, \varepsilon \in \sigma} \omega^{[ab]}[\varepsilon] z[\Delta_2, \varepsilon] W[\Delta_1, \Delta_2] \right. \right. \\
& \left. \left. + \frac{1}{10} \sum_{\varepsilon_1, \varepsilon_2 \in \sigma} \omega^{[cd]}[\varepsilon_1] \omega^{[ef]}[\varepsilon_2] f_{[cd][ef]}{}^{[ab]} W[\Delta_1, \varepsilon_1, \varepsilon_2] \right) \right) \\
& \exp \left(\frac{i}{20} \sum_{\varepsilon_1 \in \sigma} e_a[\varepsilon_1] \left(\sum_{\tau, \Delta \in \sigma} \beta^a[\Delta] z[\tau, \Delta] W[\varepsilon_1, \tau] \right. \right. \\
& \left. \left. + \frac{1}{15} \sum_{\varepsilon_2, \Delta \in \sigma} \omega^{[cd]}[\varepsilon_2] \beta^b[\Delta] \rhd_{[cd]b}{}^a W[\varepsilon_1, \varepsilon_2, \Delta] \right) \right) \\
& \exp \left(\frac{i}{5} \sum_{v \in \sigma} \phi^A[v] \left(\sum_{\tau \in \sigma} \tilde{\gamma}_A[\tau] z[\sigma, \tau] + \frac{1}{20} \sum_{\varepsilon, \tau \in \sigma} W[\varepsilon, \tau] \tilde{\gamma}^B[\tau] \alpha^\alpha[\varepsilon] \rhd_{\alpha BA} \right) \right) \\
& \exp \left(\frac{i}{5} \sum_{v \in \sigma} \bar{\psi}_A[v] \left(\sum_{\tau \in \sigma} \gamma^A[\tau] z[\sigma, \tau] \right. \right. \\
& \left. \left. + \frac{1}{20} \sum_{\varepsilon, \tau \in \sigma} W[\varepsilon, \tau] \gamma^B[\tau] \left(\alpha^\alpha[\varepsilon] \rhd_{\alpha B}{}^A + \omega^{[ab]}[\varepsilon] \rhd_{[ab]B}{}^A \right) \right) \right) \\
& \exp \left(\frac{i}{5} \left(\sum_{\tau \in \sigma} \bar{\gamma}_A[\tau] z[\sigma, \tau] \right. \right. \\
& \left. \left. - \frac{1}{20} \sum_{\varepsilon, \tau \in \sigma} W[\varepsilon, \tau] \bar{\gamma}^B[\tau] \left(\alpha^\alpha[\varepsilon] \rhd_{\alpha BA} + \omega^{[ab]}[\varepsilon] \rhd_{[ab]BA} \right) \right) \sum_{v \in \sigma} \psi^A[v] \right) \\
& \exp \left(\frac{i}{30} \sum_{\Delta_1 \in \sigma} \lambda^\alpha[\Delta_1] \left(\sum_{\Delta_2 \in \sigma} B_\alpha[\Delta_2] W[\Delta_1, \Delta_2] \right. \right. \\
& \left. \left. - \frac{12}{50} C_{\alpha\beta} \sum_{v \in \sigma} M^\beta{}_{ab}[v] \sum_{\varepsilon_1, \varepsilon_2 \in \sigma} \varepsilon_1^a \varepsilon_2^b W[\Delta_1, \varepsilon_1, \varepsilon_2] \right) \right)
\end{aligned}
$$

$$
\begin{aligned}
&\exp\left(\frac{i}{30}\sum_{\Delta_1\in\sigma}\lambda^{[ab]}[\Delta_1]\left(\frac{\varepsilon_{[ab]cd}}{80\pi l_p^2}\sum_{\varepsilon_1,\varepsilon_2\in\sigma}\varepsilon_1^c\varepsilon_2^d W[\Delta_1,\varepsilon_1,\varepsilon_2]-\sum_{\Delta_2\in\sigma}B_{[ab]}[\Delta_2]W[\Delta_1,\Delta_2]\right)\right)\\
&\exp\left(\frac{i}{20}\sum_{\varepsilon_1\in\sigma}\tilde{\lambda}^A[\varepsilon_1]\left(\sum_{\tau\in\sigma}\tilde{\gamma}_A[\tau]W[\varepsilon_1,\tau]\right.\right.\\
&\left.\left.-\frac{1}{750}\sum_{v\in\sigma}H_{abcA}[v]\sum_{\varepsilon_2,\varepsilon_3,\varepsilon_4\in\sigma}\varepsilon_2^a\varepsilon_3^b\varepsilon_4^c W[\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4]\right)\right)\\
&\exp\left(\frac{i}{20}\sum_{\varepsilon_1\in\sigma}\bar{\lambda}_A[\varepsilon_1]\left(\sum_{\tau\in\sigma}\gamma^A[\tau]W[\varepsilon_1,\tau]\right.\right.\\
&\left.\left.+\frac{i\varepsilon_{abcd}}{4500}\sum_{v\in\sigma}\sum_{\varepsilon_2,\varepsilon_3,\varepsilon_4\in\sigma}\varepsilon_2^a\varepsilon_3^b\varepsilon_4^c(\gamma^d\psi[v])^A W[\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4]\right)\right)\\
&\exp\left(-\frac{i}{20}\sum_{\varepsilon_4\in\sigma}\left(\sum_{\tau\in\sigma}\bar{\gamma}_A[\tau]W[\tau,\varepsilon_4]\right.\right.\\
&\left.\left.-\frac{i\varepsilon_{abcd}}{4500}\sum_{v\in\sigma}\sum_{\varepsilon_1,\varepsilon_2,\varepsilon_3\in\sigma}\varepsilon_1^a\varepsilon_2^b\varepsilon_3^c(\bar{\psi}[v]\gamma^d)_A W[\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4]\right)\lambda^A[\varepsilon_4]\right)\\
&\exp\left(\frac{-2\pi l_p^2\varepsilon_{abcd}}{7500}\sum_{v_1,v_2,\varepsilon_1,\varepsilon_2,\Delta\in\sigma}\bar{\psi}_A[v_1]\gamma_5\gamma^a\psi^A[v_2]\varepsilon_1^b\varepsilon_2^c\beta^d[\Delta]W[\varepsilon_1,\varepsilon_2,\Delta]\right)\\
&\exp\left(\frac{i}{1500}\sum_{v_1\in\sigma}\zeta_\alpha{}^{ab}[v_1]\sum_{\varepsilon_1,\varepsilon_2\in\sigma}\varepsilon_1^c\varepsilon_2^d\left(\sum_{\varepsilon_3,\varepsilon_4\in\sigma}W[\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4]\left(\frac{1}{50}\varepsilon_{cdef}\varepsilon_3^e\varepsilon_4^f\sum_{v_2\in\sigma}M^\alpha{}_{ab}[v_2]\right.\right.\right.\\
&\left.\left.\left.-\frac{1}{10}f_{\beta\gamma}{}^\alpha\alpha^\beta[\varepsilon_3]\alpha^\gamma[\varepsilon_4]\right)-\sum_{\Delta,\varepsilon_3\in\sigma}\alpha^\alpha[\varepsilon_3]z[\Delta,\varepsilon_3]W[\Delta,\varepsilon_1,\varepsilon_2]\eta_{ac}\eta_{bd}\right)\right)\\
&\exp\left(\frac{i}{3000}\sum_{\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4,v\in\sigma}\Lambda^{abA}[\varepsilon_1]W[\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4]\left(\frac{1}{5}H_{abcA}[v]\varepsilon^{cdef}\varepsilon_{2d}\varepsilon_{3e}\varepsilon_{4f}\right.\right.\\
&\left.\left.+\left(\phi_A[v]z[\varepsilon_2,v]+\frac{1}{5}\phi^B[v]\alpha^\alpha[\varepsilon_2]\rhd_{\alpha BA}\right)\varepsilon_{3a}\varepsilon_{4b}\right)\right)\\
&\exp\left(\frac{-i}{36000}\sum_{\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4\in\sigma}\varepsilon_{abcd}\varepsilon_1^a\varepsilon_2^b\varepsilon_3^c\varepsilon_4^d\left(\frac{\Lambda}{8\pi l_p^2}+\chi v^4-\frac{2\chi v^2}{25}\sum_{v_1,v_2\in\sigma}\phi_A[v_1]\phi^A[v_2]\right.\right.\\
&\left.\left.+\frac{1}{125}Y_{ABC}\sum_{v_1,v_2,v_3\in\sigma}\bar{\psi}^A[v_1]\psi^B[v_2]\phi^C[v_3]+\frac{\chi}{625}\sum_{v_1,v_2,v_3,v_4\in\sigma}\phi_A[v_1]\phi^A[v_2]\phi_B[v_3]\phi^B[v_4]\right)\right)\\
&F\left(\phi_k[\sigma]\right)\,. \qquad (291)
\end{aligned}
$$

Овиме је сваком пољу у квантној 3BF теорији са везама придружен одговарајући елемент триангулације и постављена је једначина за очекивану вредност произвољне опсервабле. Да би дефиниција квантне теорије била комплетна, потребно је још дефинисати домен интеграције по сваком пољу.

5.2 Прелазак са интеграције по пољима на интеграцију по групама

Да би домени интеграције по пољима били дефинисани, интеграл по алгебри ће бити замењен интегралом по групи. Ови интеграли се дефинишу увођењем појма инваријантне интеграције. Овде треба направити разлику између инваријантне интеграције и инваријантног интеграла. Инваријантни интеграл је интеграл чија је вредност оператор који комутира са свим елементима групе. Инваријантна интеграција подразумева интеграцију по елементима групе, на такав начин да резултат интеграције не зависи од преуређења елемената групе. Преуређење елемената групе се задаје деловањем једног елемента групе $h \in G$ на цео домен интеграције (све елементе групе) с лева или с десна. У општем случају се дефиниције леве и десне инваријантне интеграције разликују:

$$\int_G f(g)d\mu_L(g) = \int_G f(hg)d\mu_L(g)\,, \quad \int_G f(g)d\mu_D(g) = \int_G f(gh)d\mu_D(g)\,, \qquad d\mu_L(g) \neq d\mu_D(g)\,. \tag{292}$$

Величине $d\mu_L(g)$ и $d\mu_D(g)$ су лева и десна Хаарова мера. На основу Јакобијеве теореме, јасно је да су тежинске функције у Хааровој мери Јакобијани:

$$\mu_L(hg) = \frac{\mu_L(g)}{\left|\det \frac{\partial\theta_i(hg)}{\partial\theta_j(g)}\right|}\,, \qquad \mu_D(gh) = \frac{\mu_D(g)}{\left|\det \frac{\partial\theta_i(gh)}{\partial\theta_j(g)}\right|}\,, \tag{293}$$

где су $\theta_i(g)$ парамтери алгебре који одговарају елементу g групе G. Пошто исказ важи за свако $g, h \in G$, може се узети $g = e$, па је резултат:

$$\mu_L(h) = \frac{\mu_L(e)}{\left|\det \frac{\partial\theta_i(hg)}{\partial\theta_j(g)}\right|_{g\to e}}\,, \qquad \mu_D(h) = \frac{\mu_D(e)}{\left|\det \frac{\partial\theta_i(gh)}{\partial\theta_j(g)}\right|_{g\to e}}\,. \tag{294}$$

Заправо, фиксирањем константе нормирања у мери $\mu_L(e) = \mu_D(e) = 1$, може се конструисати веза између мера и придружене репрезентације:

$$|\det(\mathrm{ad}(g))| = \left|\det \frac{\partial\theta_i(ghg^{-1})}{\partial\theta_j(h)}\right|_{h\to e} = \left|\det \frac{\partial\theta_i(ghg^{-1})\partial\theta_k(gh)}{\partial\theta_k(ghg^{-1}g)\partial\theta_j(h)}\right|_{h\to e} = \frac{\mu_L(g)}{\mu_D(g)}\,. \tag{295}$$

Коначно, мера у интегралу по групи је:

$$d\mu_L(h) = \frac{\mu_L(e)}{\left|\det \frac{\partial\theta_i(hg)}{\partial\theta_j(g)}\right|_{g\to e}} d^n\theta\,, \qquad d\mu_D(h) = \frac{\mu_D(e)}{\left|\det \frac{\partial\theta_i(gh)}{\partial\theta_j(g)}\right|_{g\to e}} d^n\theta\,. \tag{296}$$

Главне особине Хаарове мере потребне за даљу анализу се могу изразити помоћу следеће две теореме:

Теорема 2. *Лијева група G има једнаку леву и десну Хаарову меру ако и само ако је за свако $g \in G$:*

$$|\det(\mathrm{ad}(g))| = 1\,. \tag{297}$$

Овај исказ је екввалентан захтеву да је траг придружене репрезентације елемента Лијеве алгебре једнак нули.

Доказ. Директно следи из (295). □

Директна последица теореме је да су лева и десна Хаарова мера Абелових Лијевих група једнаке. Пошто се n-параметарска Абелова група факторише на производ n једнопараметарских, тежинска функција n-параметарске Абелове групе се такође факторише на производ n једнопараметарских тежинских функција. Коначно, за једнопараметарску Абелову групу је тежинска функција:

$$\mu(g) = \frac{1}{\left|\frac{\partial\theta(ge^t)}{\partial t}\right|_{t\to 0}} = \frac{1}{|\tau|} = \mu(e)\,, \tag{298}$$

где је τ генератор Абелове групе. У случају групе Грасманових бројева $\mathbb{G}$, тежинска функција у Хааровој мери мора да буде константа, јер линеарни допринос мора да дође од подинтегралне функције.

Теорема 3. *Лева и десна Хаарова мера специјалних $SL(n,\mathbb{F})$ група, над пољем $\mathbb{F}$ су једнаке, јер њихове структурне константе задовољавају услов ${f_{ab}}^b = 0$, што је траг придружене репрезентације њихове алгебре.*

Доказ. Пошто структурне константе не зависе од избора репрезентације, докле год је репрезентација верна, репрезентација $(\tau^{\alpha\beta})_{ab}$ генератора $SL(n,\mathbb{F})$ групе се може изабрати као:

$$(\tau^{\alpha\beta})_{ab} = \delta^\alpha_a\delta^\beta_b - \delta^n_a\delta^n_b\delta^{\alpha\beta}\,. \tag{299}$$

Матрична репрезентација комутатора $[\tau^{\alpha\beta},\tau^{\gamma\delta}]_{ab}$ је антисиметрична (по ab), па је додатно множење репрезентацијом генератора $(\tau^{ab})_{cd}$ не мења, одакле се закључује да она мора бити једнака структурној константи:

$$[\tau^{\alpha\beta},\tau^{\gamma\delta}]_{ab} = \delta^{\alpha\beta}\left(\delta^\gamma_a\delta^{n\delta}\delta^n_b - \delta^\delta_b\delta^n_a\delta^{\gamma n}\right) + \delta^{\gamma\delta}\left(\delta^\beta_b\delta^n_a\delta^{n\alpha} - \delta^\alpha_a\delta^n_b\delta^{n\beta}\right) + \delta^\alpha_a\delta^\delta_b\delta^{\beta\gamma} - \delta^\gamma_a\delta^\beta_b\delta^{\alpha\delta}\,, \tag{300}$$

$$[\tau^{\alpha\beta},\tau^{\gamma\delta}]_{ab}(\tau^{ab})_{cd} = [\tau^{\alpha\beta},\tau^{\gamma\delta}]_{cd} \implies {f^{\alpha\beta\,\gamma\delta}}_{ab} = [\tau^{\alpha\beta},\tau^{\gamma\delta}]_{ab}\,. \tag{301}$$

Директном провером се добија да је ${f^{\alpha\beta\,\gamma\delta}}_{\gamma\delta} = 0$. □

Овим теоремама су обухваћене све групе које се појављују у 3-групи за Стандардни Модел (12). То значи да све групе које се појављују теорији имају исту леву и десну Хаарову меру, док је тежинска функција у Хааровој мери Абелових група константа.

Интеграл по групи је онда дефинисан као:

$$I_G[f] = \int_G f(g)d\mu(g) \tag{302}$$

и параметризован је параметрима алгебре θ_i, па се интеграција врши по овим параметрима. Тада облик интеграла постаје:

$$I_G[f] = \int_{\mathbb{F}^n} \Pi(\theta)f(g(\theta))\mu(g(\theta))d^n\theta\,, \tag{303}$$

где је $\Pi(\theta)$ правоугаона прозорска функција-компактни носач скупа вредности параметара алгебре. Дефиниција се може проширити и на некомпактне групе узимањем константне прозорске функције $\Pi(\theta) = 1$. Скуп $\mathbb{F}$ ком припадају параметри алгебре по којима се интеграли је, у зависности од врсте поља, скуп реалних или Грасманових бројева (у случају бозонских и фермионских поља, редом).

Сва поља која се јављају у дефиницији 3BF теорије са везама су елементи једне или више алгебра група из структуре 3-групе (12) и зарад лакше прегледности их је могуће приказати таблично. Пошто теорија садржи фермионе, потребно је групу $SO(3,1)$ заменити универзално наткривајућом $SL(2,\mathbb{C})$, јер $SO(3,1)$ не садржи репрезентације по којима се фермионска поља трансформишу (ротације са основним периодом 4π). Прва таблица садржи поља која припадају само једној алгебри, истовремено показујући којој алгебри поље припада и на ком симплексу живи:

симплекс	$SL(2,\mathbb{C})$	$SU(3)\times SU(2)\times U(1)$	$\mathbb{R}^4$	$\mathbb{C}^4$	$\mathbb{G}^{64}\times\mathbb{G}^{64}\times\mathbb{G}^{64}$
·				ϕ^A	ψ^A, $\bar{\psi}_A$
/	$\omega^{[ab]}$	α^{α}	e^a	$\tilde{\lambda}^A$	λ^A, $\bar{\lambda}_A$
△	$\lambda_{[ab]}$, $B_{[ab]}$	λ_α, B_α	β^a		
◬				$\tilde{\gamma}^A$	γ^A, $\bar{\gamma}_A$
⬠					

док друга садржи множитеље који носе индексе више различитих алгебри:

$(SU(3)\times SU(2)\times U(1))\times\mathbb{R}^4\times\mathbb{R}^4$	$M_{\alpha ab}$, $\zeta^{\alpha ab}$
$\mathbb{R}^4\times\mathbb{R}^4\times\mathbb{C}^4$	Λ^{abA}
$\mathbb{R}^4\times\mathbb{R}^4\times\mathbb{R}^4\times\mathbb{C}^4$	H^{abcA}

Сви множитељи у другој таблици живе само у вертексима.

5.2.1 Делта на групи и Питер-Вејлова теорема

У теорији коначних група (група са коначним бројем елемената $|G|$) је дефинисан скуп ортогоналних функција на групи, карактера $\chi^{(\Lambda)}(g)$, пребројаних иредуцибилним репрезентацијама групе, као и функција на групи $\delta(g)$ која се назива и њеном регуларном репрезентацијом, тако да важи:

$$\delta(g) = \begin{cases} 0 & g \neq e\,, \\ |G| & g = e\,. \end{cases} \tag{304}$$

Разлагање функције $\delta(g)$ по базису карактера иредуцибилних репрезентација групе даје:

$$\delta(g) = \sum_{\Lambda} \mathrm{dim}\Lambda\, \chi^{(\Lambda)}(g)\,. \tag{305}$$

Уопштење ових исказа и доказ егзистенције ортогоналног скупа карактера у случају компактних Лијевих група је дат Питер-Вејловом теоремом. Разлог употребе ове теореме је да би се интеграција Диракове делте добијене у процесу квантизације могла извршити у сваком иредуцибилном подпростору засебно, а потом се интеграл по групи заменио сумом по иредуцибилним репрезентацијама. Овај део поступка квантизације је уобичајен у моделима спинске пене. У случају 3BF теорије, чија је симетрија описана 3-групом, би уопштење овог приступа требало да Диракову делту замени делтом на 3-групи како би се интеграција по пољима вршила у ортогоналним компонентама разлагања репрезентације 3-групе. Међутим, како теорија репрезентација 3-група није развијена у довољној мери, резултати Питер-Вејлове теореме се још увек не могу уопштити на структуру 3-групе. Због тога је основна идеја да се уместо у ортогоналним подпросторима репрезентација 3-групе, интеграција делти врши унутар иредуцибилних подпростора група од којих је 3-група састављена.

Због тога је потребно одредити параметре алгебре, Хаарове мере, прозорске функције и регуларне репрезентације (делте) група које се јављају у теорији:

- $U(1)$ група

  Алгебра је параметризована једним параметром θ, а тежинска функција у Хааровој мери је константа:

$$d\mu(g) = \frac{1}{2\pi} d\theta \,. \tag{306}$$

  Домен интеграције је задат прозорском функцијом облика:

$$\Pi(\theta) = \mathrm{H}(\theta + \pi)\mathrm{H}(\pi - \theta) \,, \tag{307}$$

  где је $\mathrm{H}(s)$ Хевисајдова тета функција. Делта на групи је:

$$\delta(g) = 2\pi \sum_{k=-\infty}^{+\infty} \delta(\theta + 2k\pi) \,. \tag{308}$$

- $SU(2)$ група

  Алгебра је параметризована са три параметра, али мера, прозорска функција и делта на групи зависе само од инваријантног растојања које одговара растојању у сферним координатама $r = (\theta_\alpha \theta^\alpha)^{\frac{1}{2}}$. Мера, прозорска функција и делта на групи су:

$$d\mu(g) = \frac{1}{2\pi} \sin^2\left(\frac{1}{2} r\right) dr \,, \tag{309}$$

$$\Pi(r) = \mathrm{H}(r)\mathrm{H}(2\pi - r) \,, \tag{310}$$

$$\delta(g) = -\frac{4\pi}{\sin\left(\frac{1}{2} r\right)} \sum_{k=-\infty}^{+\infty} \delta^{(1)}(r + 4k\pi) \,, \tag{311}$$

  што након парцијалне интеграције постаје:

$$\delta(g) = \frac{4\pi}{\sin^2\left(\frac{1}{2} r\right)} \sum_{k=-\infty}^{+\infty} \delta(r + 4k\pi) \,. \tag{312}$$

- $SU(3)$ група

  Алгебра $SU(3)$ групе је осмопараметарска која има $S^3 \times S^5$ топологију, па се у складу с тим може извршити параметризација простора [57]. Инваријантна растојања од којих зависи мера, прозорска функција и делта на групи су:

$$\varphi = \frac{2}{\sqrt{3}} \sqrt{\sigma_1} \sin\left(\frac{1}{3} \mathrm{arctg}\left(\frac{\sqrt{3}\sigma_2}{\sqrt{\sigma_1^3 - 3\sigma_2^2}}\right)\right) \,, \qquad \xi = \sqrt{\sigma_1} \cos\left(\frac{1}{3} \mathrm{arctg}\left(\frac{\sqrt{3}\sigma_2}{\sqrt{\sigma_1^3 - 3\sigma_2^2}}\right)\right) \,, \tag{313}$$

$$\sigma_1 = \theta_\alpha \theta^\alpha \,, \qquad \sigma_2 = d_{\alpha\beta\gamma} \theta^\alpha \theta^\beta \theta^\gamma \,. \tag{314}$$

  Коефицијенти $d_{\alpha\beta\gamma}$ су коефицијенти из антикомутационих релација:

$$\{\tau_\alpha, \tau_\beta\} = d_{\alpha\beta\gamma} \tau^\gamma \,. \tag{315}$$

  Мера, прозорска функција и делта на групи су:

$$d\mu(g) = \frac{2}{3\pi^2} \sin^2\left(\frac{1}{2}\varphi\right) \sin^2\left(\frac{1}{4}(\varphi + 3\xi)\right) \sin^2\left(\frac{1}{4}(\varphi - 3\xi)\right) d\varphi d\xi \,, \tag{316}$$

$$\Pi(\xi,\varphi) = \mathrm{H}(\xi+2\pi)\mathrm{H}(2\pi-\xi)\mathrm{H}(\varphi+2\pi)\mathrm{H}(2\pi-\varphi)\,, \tag{317}$$

$$\delta(g) = 4\pi^2 \frac{\left(\partial_\varphi^2 + 3\partial_\xi^2 - 4\partial_\varphi\partial_\xi\right)\partial_\varphi}{\sin\left(\frac{\varphi}{2}-\frac{3\xi}{2}\right)+\sin\left(\frac{\varphi}{2}+\frac{3\xi}{2}\right)-\sin(\varphi)} \sum_{k,l=-\infty}^{+\infty} \delta(\xi+2k\pi)\delta(\xi+\varphi+4l\pi)\,. \tag{318}$$

Поново, након парцијалне интеграције се добија:

$$\delta(g) = \frac{3\pi^2}{2\sin^2\left(\frac{1}{2}\varphi\right)\sin^2\left(\frac{1}{4}\left(\varphi+3\xi\right)\right)\sin^2\left(\frac{1}{4}(\varphi-3\xi)\right)} \sum_{k,l=-\infty}^{+\infty} \delta(\xi+2k\pi)\delta(\xi+\varphi+4l\pi)\,. \tag{319}$$

- $SL(2,\mathbb{C})$ група

  Алгебра је параметризована са шест параметара и може се идентификовати са комплексификованом $su(2)$ алгебром. Ова алгебра се може разложити на директни збир две коњуговано комплексне $su(2)$ алгебре чији су параметри $\theta_a + i\rho_a$ и $\theta_a - i\rho_a$, где су $\theta_a = \varepsilon_{abc}\omega^{[bc]}$ и $\rho_a = \omega_{[0a]}$. Мера, прозорска функција и делта на групи (иако је $SL(2,\mathbb{C})$ некомпактна група, може се дефинисати делта на групи као функција параметара алгебре) су по аналогији са $SU(2)$ групом, функције од инваријантних растојања $(\theta_a\theta^a - \rho_a\rho^a + 2i\theta_a\rho^a)^{\frac{1}{2}}$ и $(\theta_a\theta^a - \rho_a\rho^a - 2i\theta_a\rho^a)^{\frac{1}{2}}$. Изрази се додатно могу поједноставити увођењем додатних смена $a = \theta_i\theta^i - \rho_i\rho^i$ и $b = 2\theta_i\rho^i$, па су мера, прозорска функција и делта на групи једнаке:

$$d\mu(g) = \frac{\left|\sin^2\left(\frac{1}{2}\sqrt{a-ib}\right)\right|^2}{8\pi^2\sqrt{a^2+b^2}} da db\,, \tag{320}$$

$$\Pi(a,b) = \mathrm{H}\left(2\pi - \left|\mathrm{Re}\left(\sqrt{a-ib}\right)\right|\right)\,, \tag{321}$$

$$\delta(g) = 4\pi^2 \frac{\delta\left(\mathrm{Im}\left(\sqrt{a-ib}\right)\right)}{\left|\sin^2\left(\frac{1}{2}\sqrt{a-ib}\right)\right|^2} \sum_{k=-\infty}^{+\infty} \delta\left(\mathrm{Re}\left(\sqrt{a-ib}\right)+4k\pi\right)\,. \tag{322}$$

- групе $\mathbb{R}$, $\mathbb{C}$ и $\mathbb{G}$

  Прво треба нагласити да су ове групе адитивне и да се њихова алгебра поклапа са групом. Алгебре су једнопараметарске са реалним, комплексним или Грасмановим бројем као параметром. Тежинска функција у мери је једнака јединици, прозорска функција обухвата цео скуп бројева, а делта је стандардна Диракова делта над параметром θ:

$$d\mu(g) = d\theta\,, \qquad \Pi(\theta) = 1\,, \qquad \delta(g) = \delta(\theta)\,. \tag{323}$$

  У случају Грасманових параметара, да би интеграл био ненулти, Диракова делта од параметра је једнака самом параметру, јер за Грасманове бројеве важе правила за интеграцију:

$$\int_{\mathbb{G}} d\theta = 0\,, \qquad \int_{\mathbb{G}} \theta d\theta = 1 \implies \delta(\theta) = \theta\,. \tag{324}$$

Овде треба издвојити неколико општих особина делте на групи. Прва, представља јединични оператор у простору функција на групи:

$$\int_G f(g)\delta(gh^{-1})d\mu(g) = f(h)\,. \tag{325}$$

У специјалном случају, када се за функцију на групи изабере баш функција која враћа елемент групе у репрезентацији D, добија се репрезентација инверзног елемента:

$$D\left(h^{-1}\right) = \int_G D(g)\delta(gh)d\mu(g)\,. \tag{326}$$

Ако се овај интеграл напише по параметрима алгебре, у случају компактних група је могуће да више различитих избора параметара у алгебри одговара истом елементу у групи, што може да поквари јединственост инверзног елемента групе до на мултипликативни фактор. Због тога је важно да прозорска функција одсеца простор параметара алгебре тако да се сваком елементу групе додељује тачно један скуп вредности параметара у алгебри који му одговара. На тај начин је обезбеђена јединственост инверза у алгебри, као и јединственост неутралног елемента у алгебри. То даље имплицира да ће производ делте на групи и прозорске функције дати само једну Диракову делту у тачки која одговара неутралном елементу алгебре и јединичном елементу у групи. Коначно, множењем са тежинском функцијом у Хааровој мери, добија се једнакост:

$$\Pi(\theta)\delta(g(\theta))\mu(g(\theta))d^n\theta = \delta(\theta)d^n\theta\,. \tag{327}$$

На овај начин је одређено како се Диракове делте добијене интеграцијом по пољима мењају делтама на групи и како прозорске функције дефинишу домен интеграције по пољима. Међутим, како квантна 3BF теорија са везама није тополошка, осим Дираковиx делти у дефиницији очекиване вредности опсервабле (207), стоји и сама опсервабла али и остали чиниоци који нису облика делте и који у свом разлагању садрже све иредуцибилне компоненте група. Због тога би разлагање дефиниције очекиване вредности опсервабле на иредуцибилне компоненте група подразумевало разлагање производа делти и осталих израза у дефиницији што је крајње нетривијално (због постојања ненултих Клебш-Гордонових коефицијената у разлагању производа репрезентација).

На овом месту се мора направити отклон од стандардне процедуре квантизације по угледу на моделе спинске пене и добијене Диракове делте се морају интегралити по дефиницији, јер замена делтом на групи не олакшава израчунавања у теорији.

5.3 Подела дејства на тополошки део и везе

Дејство 3BF теорије са везама се састоји од тополошког дела и веза. Даљи поступак конструкције квантне 3BF теорије са везама подразумева интеграцију по одређеним пољима по угледу на конструкцију тополошке теорије (66). Сума по стањима за тополошку теорију има облик производа делти на групи, па се поља по којима се врши интеграција бирају тако да се добије што је могуће више делти. Интеграли по множитељима који као резултат интеграције дају Диракову делту су облика:

$$\int \exp\left(i\sum_{\sigma}\sum_{k\in\sigma} f[k]F[k,\sigma]\right)\prod_k df[k] = \mathcal{N}\prod_k \delta\left(\sum_{\sigma\ni k}F[k,\sigma]\right)\,. \tag{328}$$

Поља по којима се интеграли су множитељи $B_\alpha[\Delta]$, $B_{[ab]}[\Delta]$, $\beta^a[\Delta]$, $\tilde{\gamma}^A[\tau]$, $\bar{\gamma}_A[\tau]$, $\gamma^A[\tau]$, $\Lambda^{abA}[\varepsilon]$ и $\zeta_{\alpha ab}[v]$. Да би резултат интеграције по овим пољима био Диракова делта, потребно је да опсервабла не зависи од ових поља. Добијен резултат се може поделити на део који одговара тополошкој инваријанти, интеракциони део и део са јаким везама.

$$\langle F\rangle = \frac{\mathcal{N}}{Z}\int\prod_{\sigma,\tau,\Delta,\varepsilon,v} d\varphi[\tau,\Delta,\varepsilon,v]\mathrm{TI}\Big[\mathcal{F}[\varphi]-\lambda[\varphi]\Big]\,\mathrm{INT}\Big[\lambda[\varphi]-G[\varphi]\Big]\,\mathrm{SC}[\varphi]\,F(\varphi_k)\,. \tag{329}$$

За разлику од тополошке теорије, где је интеграција вршена по пољима B, C и D, која одговарају множитељима B, тетрадама и фермионима и скаларним пољима, у случају конструкције квантне 3BF теорије са везама се не може интегралити по тетрадама и пољима материје, јер произвољна опсервабла може нетривијално да зависи од њих. Међутим, како је вредност суме по стањима у тополошкој 3BF теорији (66) независна од редоследа интеграције, може се добити и алтернативна тополошка инваријанта интеграцијом по пољима B, β и γ. Ова алтернативна тополошка инваријанта је истог облика као први чинилац у изразу (329) који је једнак:

$$
\begin{aligned}
\mathrm{TI}[\mathcal{F}[\varphi]-\lambda[\varphi]] = \prod_{\Delta_1} \delta\Bigg(\sum_{\sigma\ni\Delta_1} \Bigg(\sum_{\Delta_2,\varepsilon\in\sigma} \alpha^{\alpha}[\varepsilon] z[\Delta_2,\varepsilon] W[\Delta_1,\Delta_2] \\
+ \frac{1}{10} \sum_{\varepsilon_1,\varepsilon_2\in\sigma} \alpha^{\beta}[\varepsilon_1]\alpha^{\gamma}[\varepsilon_2] f_{\beta\gamma}{}^{\alpha} W[\Delta_1,\varepsilon_1,\varepsilon_2] + \sum_{\Delta_2\in\sigma} \lambda^{\alpha}[\Delta_2] W[\Delta_1,\Delta_2] \Bigg)\Bigg) \\
\prod_{\Delta_1} \delta\Bigg(\sum_{\sigma\ni\Delta_1} \Bigg(\sum_{\Delta_2,\varepsilon\in\sigma} \omega^{[ab]}[\varepsilon] z[\Delta_2,\varepsilon] W[\Delta_1,\Delta_2] - \sum_{\Delta_2\in\sigma} \lambda^{[ab]}[\Delta_2] W[\Delta_1,\Delta_2] \\
+ \frac{1}{10} \sum_{\varepsilon_1,\varepsilon_2\in\sigma} \omega^{[cd]}[\varepsilon_1]\omega^{[ef]}[\varepsilon_2] f_{[cd][ef]}{}^{[ab]} W[\Delta_1,\varepsilon_1,\varepsilon_2] \Bigg)\Bigg) \\
\prod_{\Delta} \delta\Bigg(\sum_{\sigma\ni\Delta} \Bigg(\sum_{\varepsilon,\tau\in\sigma} \varepsilon_a z[\tau,\Delta] W[\varepsilon,\tau] + \frac{1}{15} \sum_{\varepsilon_1,\varepsilon_2\in\sigma} \varepsilon_{1b}\, \omega^{[cd]}[\varepsilon_2] \rhd_{[cd]a}{}^{b} W[\varepsilon_1,\varepsilon_2,\Delta] \\
+ \frac{4\pi i l_p^2}{750} \sum_{v_1,v_2,\varepsilon_1,\varepsilon_2\in\sigma} \bar{\psi}_A[v_1]\gamma_5\gamma^d\psi^A[v_2] \varepsilon_1^b \varepsilon_2^c \varepsilon_{dbca} W[\varepsilon_1,\varepsilon_2,\Delta] \Bigg)\Bigg) \\
\prod_{\tau} \delta\Bigg(\sum_{\sigma\ni\tau} \Bigg(\sum_{v\in\sigma} \phi^A[v] \Bigg(\delta^B_A z[\sigma,\tau] + \frac{1}{20} \sum_{\varepsilon\in\sigma} W[\varepsilon,\tau] \alpha^{\alpha}[\varepsilon] \rhd_{\alpha}{}^{B}{}_{A} \Bigg) \\
+ \frac{1}{4} \sum_{\varepsilon\in\sigma} \tilde{\lambda}^B[\varepsilon] W[\varepsilon,\tau] \Bigg)\Bigg) \\
\prod_{\tau} \delta\Bigg(\sum_{\sigma\ni\tau} \Bigg(\sum_{v\in\sigma} \bar{\psi}_B[v] \Bigg(\delta^B_A z[\sigma,\tau] + \frac{1}{20} \sum_{\varepsilon\in\sigma} W[\varepsilon,\tau] \left(\alpha^{\alpha}[\varepsilon] \rhd_{\alpha A}{}^{B} \right.\\
\left. + \; \omega^{[ab]}[\varepsilon] \rhd_{[ab]A}{}^{B} \right) \Bigg) + \frac{1}{4} \sum_{\varepsilon\in\sigma} \bar{\lambda}_A[\varepsilon] W[\varepsilon,\tau] \Bigg)\Bigg) \\
\prod_{\tau} \delta\Bigg(\sum_{\sigma\ni\tau} \Bigg(\Bigg(\delta^A_B z[\sigma,\tau] + \frac{1}{20} \sum_{\varepsilon\in\sigma} W[\varepsilon,\tau] \left(\alpha^{\alpha}[\varepsilon] \rhd_{\alpha B}{}^{A} \right.\\
\left. + \; \omega^{[ab]}[\varepsilon] \rhd_{[ab]B}{}^{A} \right) \Bigg) \sum_{v\in\sigma} \psi^B[v] - \frac{1}{4} \sum_{\varepsilon\in\sigma} \lambda^A[\varepsilon] W[\varepsilon,\tau] \Bigg)\Bigg) \,. \qquad (330)
\end{aligned}
$$

Део који одговара јаким везама је једнак:

$$\begin{aligned}
\mathrm{SC}[\varphi] = \prod_{\varepsilon_1} \delta \Bigg(\sum_{\sigma \ni \varepsilon_1} \Bigg(\sum_{v, \varepsilon_2, \varepsilon_3, \varepsilon_4 \in \sigma} W[\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4] \left(\frac{1}{5} H_{abcA}[v] \varepsilon^{cdef} \varepsilon_{2d} \varepsilon_{3e} \varepsilon_{4f} \right. \\
+ \left. \left(\phi_A[v] z[\varepsilon_2, v] + \frac{1}{5} \phi^B[v] \alpha^\alpha[\varepsilon_2] \rhd_{\alpha BA} \right) \varepsilon_{3a} \varepsilon_{4b} \right) \Bigg) \Bigg) \\
\prod_v \delta \Bigg(\sum_{\sigma \ni v} \Bigg(\sum_{\varepsilon_1, \varepsilon_2, \varepsilon_3 \in \sigma} \varepsilon_1^c \varepsilon_2^d \Bigg(\sum_{\varepsilon_4 \in \sigma} W[\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4] \left(\frac{1}{50} \varepsilon_{cdef} \varepsilon_3^e \varepsilon_4^f \sum_{\tilde{v} \in \sigma} M^\alpha{}_{ab}[\tilde{v}] \right. \\
- \left. \frac{1}{10} f_{\beta\gamma}{}^\alpha \alpha^\beta[\varepsilon_3] \alpha^\gamma[\varepsilon_4] \eta_{ac} \eta_{bd} \right) - \sum_{\Delta \in \sigma} \alpha^\alpha[\varepsilon_3] z[\Delta, \varepsilon_3] W[\Delta, \varepsilon_1, \varepsilon_2] \eta_{ac} \eta_{bd} \Bigg) \Bigg) \Bigg), \qquad (331)
\end{aligned}$$

док је део који одговара интеракцијама једнак:

$$\begin{aligned}
\mathrm{INT}[\lambda[\varphi] - G[\varphi]] = \prod_\sigma \Bigg[\exp \left(\frac{-i}{125} \sum_{\Delta, \varepsilon_1, \varepsilon_2, v \in \sigma} C_{\alpha\beta} \lambda^\alpha[\Delta] M^\beta{}_{ab}[v] \varepsilon_1^a \varepsilon_2^b W[\Delta, \varepsilon_1, \varepsilon_2] \right) \\
\exp \left(\frac{i}{2400 \pi l_p^2} \sum_{\Delta, \varepsilon_1, \varepsilon_2 \in \sigma} \varepsilon_{[ab]cd} \lambda^{[ab]}[\Delta] \varepsilon_1^c \varepsilon_2^d W[\Delta, \varepsilon_1, \varepsilon_2] \right) \\
\exp \left(\frac{-i}{15000} \sum_{\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, v \in \sigma} \tilde{\lambda}^A[\varepsilon_1] H_{abcA}[v] \varepsilon_2^a \varepsilon_3^b \varepsilon_4^c W[\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4] \right) \\
\exp \left(\frac{-i}{90000} \sum_{\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, v \in \sigma} \bar{\lambda}_A[\varepsilon_1] \left(\gamma^d \psi[v] \right)^A \varepsilon_2^a \varepsilon_3^b \varepsilon_4^c \varepsilon_{abcd} W[\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4] \right) \\
\exp \left(\frac{-i}{90000} \sum_{\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, v \in \sigma} \bar{\psi}_A[v] \left(\gamma^d \lambda[\varepsilon_4] \right)^A \varepsilon_1^a \varepsilon_2^b \varepsilon_3^c \varepsilon_{abcd} W[\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4] \right) \\
\exp \Bigg(\frac{-i}{36000} \sum_{\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4 \in \sigma} \varepsilon_{abcd} \varepsilon_1^a \varepsilon_2^b \varepsilon_3^c \varepsilon_4^d \Bigg(\frac{\Lambda}{8 \pi l_p^2} + \chi v^4 - \frac{2 \chi v^2}{25} \sum_{v_1, v_2 \in \sigma} \phi_A[v_1] \phi^A[v_2] \\
+ \frac{1}{125} Y_{ABC} \sum_{v_1, v_2, v_3 \in \sigma} \bar{\psi}^A[v_1] \psi^B[v_2] \phi^C[v_3] \\
+ \frac{\chi}{625} \sum_{v_1, v_2, v_3, v_4 \in \sigma} \phi_A[v_1] \phi^A[v_2] \phi_B[v_3] \phi^B[v_4] \Bigg) \Bigg) \Bigg]. \qquad (332)
\end{aligned}$$

Даљу интеграцију по Лагранжевим множитељима није једноставно спровести по дефиницији јер облик израза зависи од облика триангулације, међутим, интеграли су добро дефинисани. Због тога се предвиђања теорије могу израчунати нумерички коришћењем прилагођеног софтвера за конкретан избор триангулације. Овиме је у потпуности задата дефиниција квантне 3BF теорије са везама која одговара квантној Ајнштајн-Картановој теорији гравитације са спин-спин контактном интеракцијом спрегнутом са Стандардним Моделом према релацији (226). У наредном поглављу ће, комплетности ради, бити извршена прелиминарна анализа семикласичног лимеса ове квантне теорије.

5.4 Прелиминарна анализа семикласичног лимеса

Семикласични лимес представља испитивање предвиђања теорије у стању у ком се квантне поправке могу занемарити. Један од начина за испитивање семикласичног лимеса је већ представљен у поглављу 4.2 узимањем лимеса у ком све неордђености опсервабли теже нули. Овакав приступ је тешко поновити у случају произвољне опсервабле јер би то подразумевало да постоји општа формула за очекивану вредност произвољне опсервабле у којој су сви интеграли из дефиниције (329) решени. Како ово није случај, семикласичан лимес се испитује рачунањем ефективног дејства. Након одређивања ефективног дејства се над добијеним резултатом спроводи двоструки лимес:

$$\lambda_{\varphi_k} \gg l_\varepsilon \gg l_p \,, \tag{333}$$

где су λ_{φ_k} придружене таласне дужине компонената поља φ_k, које су обрнуто пропорционале брзини промене поља кроз просторвреме, l_ε је дужина ивице 4-симплекса на триангулацији и l_p је Планкова дужина. Другим речима, поља су споропроменљива у односу на период ћелија триангулације, тако да се конфигурација поља на многострукости може описати глатко са део-по-део константним вредностима на елементима триангулације, јер су ћелије довољно мале и у довољно великом броју. Са друге стране димензија ћелије много већа од Планкове дужине, па је таквом триангулацијом немогуће описати ефекте на скалама на којима се традиционално очекују доминантни ефекти квантне гравитације.

У овом поглављу, ефективно дејство неће бити рачунато, већ ће двоструки лимес бити примењен директно на дефиницију очекиване вредности опсервабле. Применом горње стране лимеса, да су поља споропроменљива у односу на учестаност дискретизације (број симплекса је велики унутар периода промене поља), вредност поља на једном симплексу се може апроксимирати средњом вредношћу поља на околним симплексима. Последица тога је да се делте у тополошкој инваријанти и јаким везама, чији је аргумент сума по симплексима, могу заменити производом делти по симплексима, чији аргумент зависи само од једног симплекса. Овај апроксимативни облик чиниоца са јаким везама је:

$$\begin{aligned}
\mathrm{SC}[\varphi] \;=\; & \prod_{\sigma} \Bigg[\delta\Bigg(\sum_{\varepsilon_1,\varepsilon_2,\varepsilon_3\in\sigma} \varepsilon_1^c \varepsilon_2^d \Bigg(\sum_{\varepsilon_4\in\sigma} W[\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4] \Bigg(\frac{1}{50} \varepsilon_{cdef} \varepsilon_3^e \varepsilon_4^f \sum_{v\in\sigma} M^{\alpha}{}_{ab}[v] \\
& - \frac{1}{10} f_{\beta\gamma}{}^{\alpha} \alpha^\beta[\varepsilon_3] \alpha^\gamma[\varepsilon_4] \eta_{ac}\eta_{bd} \Bigg) - \sum_{\Delta\in\sigma} \alpha^\alpha[\varepsilon_3] z[\Delta,\varepsilon_3] W[\Delta,\varepsilon_1,\varepsilon_2] \eta_{ac}\eta_{bd} \Bigg)\Bigg) \\
& \prod_{\varepsilon_1\in\sigma} \delta\Bigg(\sum_{v,\varepsilon_2,\varepsilon_3,\varepsilon_4\in\sigma} W[\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4] \Bigg(\frac{1}{5} H_{abcA}[v] \varepsilon^{cdef} \varepsilon_{2d}\varepsilon_{3e}\varepsilon_{4f} \\
& + \Bigg(\phi_A[v] z[\varepsilon_2, v] + \frac{1}{5}\phi^B[v] \alpha^\alpha[\varepsilon_2] \triangleright_{\alpha BA} \Bigg) \varepsilon_{3a}\varepsilon_{4b} \Bigg)\Bigg)\Bigg] \,,
\end{aligned} \tag{334}$$

док је део који одговара тополошкој инваријанти једнак:

$$
\begin{aligned}
\mathrm{TI}[\mathcal{F}[\varphi]-\lambda[\varphi]] = \prod_{\sigma} \Bigg[& \prod_{\Delta_1\in\sigma} \delta \Bigg(\sum_{\Delta_2,\varepsilon\in\sigma} \alpha^{\alpha}[\varepsilon] z[\Delta_2,\varepsilon] W[\Delta_1,\Delta_2] \\
& + \frac{1}{10} \sum_{\varepsilon_1,\varepsilon_2\in\sigma} \alpha^{\beta}[\varepsilon_1]\alpha^{\gamma}[\varepsilon_2] f_{\beta\gamma}{}^{\alpha} W[\Delta_1,\varepsilon_1,\varepsilon_2] + \sum_{\Delta_2\in\sigma} \lambda^{\alpha}[\Delta_2] W[\Delta_1,\Delta_2] \Bigg) \\
& \prod_{\Delta_1\in\sigma} \delta \Bigg(\sum_{\Delta_2,\varepsilon\in\sigma} \omega^{[ab]}[\varepsilon] z[\Delta_2,\varepsilon] W[\Delta_1,\Delta_2] \\
& + \frac{1}{10} \sum_{\varepsilon_1,\varepsilon_2\in\sigma} \omega^{[cd]}[\varepsilon_1]\omega^{[ef]}[\varepsilon_2] f_{[cd][ef]}{}^{[ab]} W[\Delta_1,\varepsilon_1,\varepsilon_2] - \sum_{\Delta_2\in\sigma} \lambda^{[ab]}[\Delta_2] W[\Delta_1,\Delta_2] \Bigg) \\
& \prod_{\Delta\in\sigma} \delta \Bigg(\sum_{\varepsilon,\tau\in\sigma} \varepsilon_a z[\tau,\Delta] W[\varepsilon,\tau] + \frac{1}{15} \sum_{\varepsilon_1,\varepsilon_2\in\sigma} \varepsilon_{1b}\, \omega^{[cd]}[\varepsilon_2] \rhd_{[cd]a}{}^{b} W[\varepsilon_1,\varepsilon_2,\Delta] \\
& + \frac{4\pi i l_p^2}{750} \sum_{v_1,v_2,\varepsilon_1,\varepsilon_2\in\sigma} \bar{\psi}_A[v_1]\gamma_5\gamma^d\psi^A[v_2]\varepsilon_1^b\varepsilon_2^c\varepsilon_{dbca} W[\varepsilon_1,\varepsilon_2,\Delta] \Bigg) \\
& \prod_{\tau\in\sigma} \delta \Bigg(\sum_{v\in\sigma} \phi^A[v] \Bigg(\delta_A^B z[\sigma,\tau] + \frac{1}{20} \sum_{\varepsilon\in\sigma} W[\varepsilon,\tau] \alpha^{\alpha}[\varepsilon] \rhd_{\alpha}{}^{B}{}_{A} \Bigg) \\
& + \frac{1}{4} \sum_{\varepsilon\in\sigma} \tilde{\lambda}^B[\varepsilon] W[\varepsilon,\tau] \Bigg) \\
& \prod_{\tau\in\sigma} \delta \Bigg(\sum_{v\in\sigma} \bar{\psi}_B[v] \Bigg(\delta_A^B z[\sigma,\tau] + \frac{1}{20} \sum_{\varepsilon\in\sigma} W[\varepsilon,\tau] \left(\alpha^{\alpha}[\varepsilon] \rhd_{\alpha A}{}^{B} + \omega^{[ab]}[\varepsilon] \rhd_{[ab]A}{}^{B} \right) \Bigg) \\
& + \frac{1}{4} \sum_{\varepsilon\in\sigma} \bar{\lambda}_A[\varepsilon] W[\varepsilon,\tau] \Bigg) \\
& \prod_{\tau\in\sigma} \delta \Bigg(\Bigg(\Bigg(\delta_B^A z[\sigma,\tau] + \frac{1}{20} \sum_{\varepsilon\in\sigma} W[\varepsilon,\tau] \left(\alpha^{\alpha}[\varepsilon] \rhd_{\alpha B}{}^{A} + \omega^{[ab]}[\varepsilon] \rhd_{[ab]B}{}^{A} \right) \Bigg) \sum_{v\in\sigma} \psi^B[v] \\
& - \frac{1}{4} \sum_{\varepsilon\in\sigma} \lambda^A[\varepsilon] W[\varepsilon,\tau] \Bigg) \Bigg] \,.
\end{aligned}
\tag{335}
$$

Аргументи делти добијених у чиниоцу са јаким везама зависе од поља $M^{\alpha}{}_{ab}$ и H_{abcA}. Међутим, у колико опсервабла чија се очекивана вредност рачуна не зависи од ових поља, остатак дефиниције не зависи од вредности ових поља у сваком вертексу симплекса појединачно, већ само од средње вредности ових поља на целом 4-симплексу, па се увођењем смена:

$$
\Xi^{\alpha}{}_{ab}[\sigma] = \frac{1}{5}\sum_{v\in\sigma} M^{\alpha}{}_{ab}[v]\,, \qquad \Theta_{abcA}[\sigma] = \frac{1}{5}\sum_{v\in\sigma} H_{abcA}[v]\,, \tag{336}
$$

ове делте могу свести на облик погодан за интеграцију. Интеграција по осталим вредностима поља у појединачним вертексима симплекса не даје допринос очекиваној вредности опсервабле, већ само мења вредност нормализационе константе $\mathcal{N}$.

Чинилац са јаким везама добијен овим сменама је облика:

$$\begin{aligned}
\mathrm{SC}[\varphi] \;=\; & \prod_{\sigma}\Bigg[\frac{1}{V^{|\alpha||[ab]|}[\sigma]}\delta\Bigg({\Xi^{\alpha}}_{ab}[\sigma]+\frac{1}{4!V[\sigma]}\Bigg(\frac{1}{300}\sum_{\varepsilon_1,\varepsilon_2,\varepsilon_3,\Delta\in\sigma}\alpha^{\alpha}[\varepsilon_3]z[\Delta,\varepsilon_3]W[\Delta,\varepsilon_1,\varepsilon_2]\varepsilon_{1a}\varepsilon_{2b} \\
& + \frac{1}{3000}\sum_{\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4\in\sigma}{f_{\beta\gamma}}^{\alpha}\alpha^{\beta}[\varepsilon_3]\alpha^{\gamma}[\varepsilon_4]\varepsilon_{1a}\varepsilon_{2b}W[\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4]\Bigg)\Bigg) \\
& \frac{1}{V^{|A|(|a|-1)|[ab]|}[\sigma]}\delta\Bigg(\Theta_{abcA}[\sigma]+\frac{1}{3!V[\sigma]}\frac{1}{3000}\sum_{v,\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4\in\sigma}W[\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4]\Bigg(\phi_A[v]z[\varepsilon_1,v] \\
& + \frac{1}{5}\phi^{B}[v]\alpha^{\alpha}[\varepsilon_1]\triangleright_{\alpha BA}\Bigg)\varepsilon_{2c}\varepsilon_{3a}\varepsilon_{4b}\Bigg)\Bigg]\,. \qquad (337)
\end{aligned}$$

У случају тополошке инваријанте, постоји више врста смена које треба увести у циљу свођења делти на облик погодан за интеграцију. Прву врсту чине смене за поља $\lambda^{[ab]}$ и λ^{α}:

$$Q^{[ab]}[\Delta,\sigma]=\frac{1}{3}\sum_{\tilde{\Delta}\in\sigma}\lambda^{[ab]}[\tilde{\Delta}]W[\Delta,\tilde{\Delta}]\,,\qquad Q^{\alpha}[\Delta,\sigma]=\frac{1}{3}\sum_{\tilde{\Delta}\in\sigma}\lambda^{\alpha}[\tilde{\Delta}]W[\Delta,\tilde{\Delta}]\,. \qquad (338)$$

Израз у интеракционом чиниоцу који садржи поља $\lambda^{[ab]}$ и λ^{α} је облика:

$$\begin{aligned}
\frac{1}{300}\sum_{\tilde{\Delta},\varepsilon_1,\varepsilon_2\in\sigma}\lambda[\tilde{\Delta}]W[\tilde{\Delta},\varepsilon_1,\varepsilon_2]\varepsilon_1^a\varepsilon_2^b &= \frac{1}{300}\sum_{\tilde{\Delta},\Delta\in\sigma}\Bigg(6\lambda[\tilde{\Delta}]\Delta^{ab}W[\tilde{\Delta},\Delta]+\sum_{\varepsilon_1,\varepsilon_2\notin\Delta}\lambda[\tilde{\Delta}]\varepsilon_1^a\varepsilon_2^bW[\tilde{\Delta},\varepsilon_1,\varepsilon_2]\Bigg) \\
&= \frac{1}{30}\sum_{\tilde{\Delta},\Delta\in\sigma}\lambda[\tilde{\Delta}]W[\tilde{\Delta},\Delta]\Delta^{ab}=\frac{1}{10}\sum_{\Delta\in\sigma}Q[\Delta,\sigma]\Delta^{ab}\,. \qquad (339)
\end{aligned}$$

Допринос додатна четири сабирка која долазе од случаја када две ивице не формирају троугао се такође може апроксимирати коришћењем лимеса споропроменљивих поља доприносом осталих сабирака код којих избор ивица чини троугао. Другу врсту смена чине смене за поља $\tilde{\lambda}^A$, λ^A и $\bar{\lambda}_A$:

$$Q^{A}[\tau,\sigma]=\frac{1}{4}\sum_{\varepsilon\in\sigma}\lambda^{A}[\varepsilon]W[\varepsilon,\tau]\,. \qquad (340)$$

Слично као у претходном случају, израз који се појављује у интеракционом чиниоцу, а који садржи поља $\tilde{\lambda}^A$, λ^A и $\bar{\lambda}_A$ је облика:

$$\begin{aligned}
&\frac{1}{3000}\sum_{\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4\in\sigma}\lambda^{A}[\varepsilon_1]\varepsilon_2^a\varepsilon_3^b\varepsilon_4^cW[\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4] \\
&= \frac{1}{3000}\sum_{\varepsilon_1,\tau\in\sigma}\lambda^{A}[\varepsilon_1]\Bigg(96W[\varepsilon_1,\tau]\tau^{abc}+\sum_{\varepsilon_2,\varepsilon_3,\varepsilon_4\notin\tau}\varepsilon_2^a\varepsilon_3^b\varepsilon_4^cW[\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4]\Bigg) \\
&= \frac{1}{3000}\sum_{\varepsilon_1,\tau\in\sigma}\lambda^{A}[\varepsilon_1]\left(150W[\varepsilon_1,\tau]\tau^{abc}\right)=\frac{1}{5}\sum_{\tau\in\sigma}Q^{A}[\tau,\sigma]\tau^{abc}\,. \qquad (341)
\end{aligned}$$

Овде је поново, као у претходном случају, допринос преостала 54 сабирка који потичу од избора ивица које не чине тетраедар симплекса апроксимиран на основу лимеса споропроменљивих поља доприносом осталих сабирака који одговарају тетраедрима.

Трећу врсту чини скуп смена за поље $\omega^{[ab]}$. Смена се уводи у два корака:

$$\omega^{[ab]}[\varepsilon] = \Omega^{[ab]e}[\sigma]\varepsilon_e\,, \quad A_{abc}[\sigma] = \frac{1}{2}\left(\Omega_{abc}[\sigma] - \Omega_{acb}[\sigma]\right)\,. \tag{342}$$

Пошто је смена $\Omega^{[ab]e}[\sigma]$ уведена за цео 4-симплекс, постоји избор 4 независне ивице ε на основу којих се може одредити њена вредност у зависности од $\omega^{[ab]}[\varepsilon]$ и скупа четири независних ивица ε_a. Јакобијан смене је нетривијалан и исти као у (211):

$$d\omega^{[ab]}[\varepsilon] = \mathcal{N}' V^{|[ab]|}[\sigma] \prod_c d\Omega[\sigma]^{[ab]c}\,, \tag{343}$$

где нормализациони фактор потиче од чињенице да са десне стране једнакости постоји већи број независних компонената поља него на левој. Међутим, како подинтеграна функција не зависи од ових додатних компонената (јер се у изразима појављује само производ $\Omega^{[ab]e}[\sigma]\varepsilon_e$), интеграција по овим додатним компонентама даје само додатни нормализациони фактор $\frac{1}{\mathcal{N}'}$. Јакобијан друге смене је константан, а инверзна смена промењљивих

$$\omega^{ab}[\varepsilon] = \left(A^{abc}[\sigma] - A^{bac}[\sigma] - A^{cab}[\sigma]\right)\varepsilon_c\,. \tag{344}$$

Ова смена је, због једноставности записа, експлицитно уведена само у делти која је настала интеграцијом везе за торзију, док је у осталим делтама подразумевано да имплицитно важи.

Тополошка инваријанта након увођења ових смена постаје:

$$\begin{aligned}
\mathrm{TI}[\mathcal{F}[\varphi] - \lambda[\varphi]] = \prod_\sigma \Bigg[&\prod_{\Delta_1\in\sigma} \delta\Bigg(\sum_{\Delta_2,\varepsilon\in\sigma} \alpha^\alpha[\varepsilon] z[\Delta_2,\varepsilon] W[\Delta_1,\Delta_2] \\
&+ \frac{1}{10}\sum_{\varepsilon_1,\varepsilon_2\in\sigma} \alpha^\beta[\varepsilon_1]\alpha^\gamma[\varepsilon_2] f_{\beta\gamma}{}^\alpha W[\Delta_1,\varepsilon_1,\varepsilon_2] + 3Q^\alpha[\Delta_1,\sigma]\Bigg) \\
&\prod_{\Delta_1\in\sigma} \delta\Bigg(\sum_{\Delta_2,\varepsilon\in\sigma} \omega^{[ab]}[\varepsilon] z[\Delta_2,\varepsilon] W[\Delta_1,\Delta_2] \\
&+ \frac{1}{10}\sum_{\varepsilon_1,\varepsilon_2\in\sigma} \omega^{[cd]}[\varepsilon_1]\omega^{[ef]}[\varepsilon_2] f_{[cd][ef]}{}^{[ab]} W[\Delta_1,\varepsilon_1,\varepsilon_2] - 3Q^{[ab]}[\Delta_1,\sigma]\Bigg) \\
&\frac{1}{V^{|[ab]|\left(\frac{|a|}{2}-1\right)}[\sigma]} \delta\Bigg(A_{abc}[\sigma] - \frac{1}{20V[\sigma]} \sum_{\Delta,\varepsilon,\tau\in\sigma} \varepsilon_a z[\tau,\Delta] W[\varepsilon,\tau] \Delta^{ef}\varepsilon_{efbc} \\
&+ \frac{2\pi i l_p^2}{25} \sum_{v_1,v_2\in\sigma} \bar\psi_A[v_1]\gamma_5\gamma^d\psi^A[v_2]\varepsilon_{dbca}\Bigg) \\
&\prod_{\tau\in\sigma} \delta\Bigg(\sum_{v\in\sigma} \phi^A[v] \Bigg(\delta^B_A z[\sigma,\tau] + \frac{1}{20}\sum_{\varepsilon\in\sigma} W[\varepsilon,\tau]\alpha^\alpha[\varepsilon] \triangleright_\alpha{}^B{}_A \Bigg) + \tilde Q^B[\tau,\sigma]\Bigg) \\
&\prod_{\tau\in\sigma} \delta\Bigg(\sum_{v\in\sigma} \bar\psi_B[v] \Bigg(\delta^B_A z[\sigma,\tau] + \frac{1}{20}\sum_{\varepsilon\in\sigma} W[\varepsilon,\tau]\left(\alpha^\alpha[\varepsilon] \triangleright_{\alpha A}{}^B + \omega^{[ab]}[\varepsilon]\triangleright_{[ab]A}{}^B\right) \Bigg) + \bar Q_a[\tau,\sigma]\Bigg) \\
&\prod_{\tau\in\sigma} \delta\Bigg(\Bigg(\delta^A_B z[\sigma,\tau] + \frac{1}{20}\sum_{\varepsilon\in\sigma} W[\varepsilon,\tau]\left(\alpha^\alpha[\varepsilon] \triangleright_{\alpha B}{}^A + \omega^{[ab]}[\varepsilon]\triangleright_{[ab]B}{}^A\right) \Bigg) \sum_{v\in\sigma}\psi^B[v] - Q^A[\tau,\sigma]\Bigg)\Bigg]\,,
\end{aligned} \tag{345}$$

Интеракциони чинилац након увођења смена постаје:

$$
\begin{aligned}
\mathrm{INT}[\lambda[\varphi]-G[\varphi]] = \prod_{\sigma} \Bigg[& \exp\left(\frac{-6i}{5}\sum_{\tilde{\Delta},\Delta\in\sigma} C_{\alpha\beta}Q^{\alpha}[\tilde{\Delta},\sigma]W[\Delta,\tilde{\Delta}]\Xi^{\beta}{}_{ab}[\sigma]\Delta^{ab}\right) \\
& \exp\left(\frac{i}{80\pi l_p^2}\sum_{\Delta,\tilde{\Delta}\in\sigma}\varepsilon_{[ab]cd}Q^{[ab]}[\tilde{\Delta},\sigma]W[\Delta,\tilde{\Delta}]\Delta^{cd}\right) \\
& \exp\left(-\frac{i}{5}\sum_{\tau\in\sigma}\tilde{Q}^{A}[\tau,\sigma]\Theta_{abcA}[\sigma]\tau^{abc}\right) \\
& \exp\left(\frac{-i}{150}\sum_{v,\tau\in\sigma}\bar{Q}_{A}[\tau,\sigma]\left(\gamma^{d}\psi[v]\right)^{A}\varepsilon_{abcd}\tau^{abc}\right) \\
& \exp\left(\frac{-i}{150}\sum_{v,\tau\in\sigma}\bar{\psi}_{A}[v]\left(\gamma^{d}Q[\tau,\sigma]\right)^{A}\varepsilon_{abcd}\tau^{abc}\right) \\
& \exp\left(\frac{-i}{36000}\sum_{\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4\in\sigma}\varepsilon_{abcd}\varepsilon_1^a\varepsilon_2^b\varepsilon_3^c\varepsilon_4^d\left(\frac{\Lambda}{8\pi l_p^2}+\chi v^4-\frac{2\chi v^2}{25}\sum_{v_1,v_2\in\sigma}\phi_A[v_1]\phi^A[v_2]\right.\right. \\
& +\frac{1}{125}Y_{ABC}\sum_{v_1,v_2,v_3\in\sigma}\bar{\psi}^{A}[v_1]\psi^{B}[v_2]\phi^{C}[v_3] \\
& \left.\left.+\ \frac{\chi}{625}\sum_{v_1,v_2,v_3,v_4\in\sigma}\phi_A[v_1]\phi^A[v_2]\phi_B[v_3]\phi^B[v_4]\right)\right)\Bigg]\,. \qquad (346)
\end{aligned}
$$

Након интеграције свих делти, резултујући сабирци у експоненту одговарају лимесу споро-промењљивих поља за дејство за Стандардни Модел на триангулацији у закривљеном простору са додатним сабирком за контактну спин-спин интеракцију и сабирак који одговара скаларној кривини, који треба посебно анализирати. Тај сабирак је једнак:

$$
iS_{\mathrm{grav}} = \frac{i}{80\pi l_p^2}\sum_{\sigma}\sum_{\Delta\in\sigma}\tilde{Q}^{[ab]}[\Delta,\sigma]\Delta^{cd}\varepsilon_{[ab]cd} = \frac{i}{8\pi l_p^2}\sum_{\sigma}\frac{1}{10}\sum_{\Delta\in\sigma}\frac{\varepsilon_{[ab]cd}\tilde{Q}^{[ab]}[\Delta,\sigma]\Delta^{cd}}{A[\Delta]}A[\Delta]\,. \qquad (347)
$$

Величина $\tilde{Q}^{[ab]}[\Delta]$ је бездимензионална и представља угаони фактор једнак:

$$
\tilde{Q}^{[ab]}[\Delta,\sigma] = \frac{1}{3}\sum_{\varepsilon,\tilde{\Delta}\in\sigma}\tilde{\omega}^{[ab]}[\varepsilon]z[\tilde{\Delta},\varepsilon]W[\tilde{\Delta},\Delta] + \frac{1}{30}\sum_{\varepsilon_1,\varepsilon_2\in\sigma}\tilde{\omega}^{[a|c}[\varepsilon_1]\tilde{\omega}_{c}{}^{|b]}[\varepsilon_2]W[\Delta,\varepsilon_1,\varepsilon_2]\,, \qquad (348)
$$

где је

$$
\tilde{\omega}^{[ab]}[\varepsilon] = \frac{1}{V[\sigma]}\frac{1}{20}\sum_{\tilde{\varepsilon},\Delta,\tau\in\sigma}z[\tau,\Delta]W[\tilde{\varepsilon},\tau]\Delta^{ef}\varepsilon_c\tilde{\varepsilon}^q\left(\delta_q^a\varepsilon_{ef}{}^{bc}-\delta_q^b\varepsilon_{ef}{}^{ac}-\delta_q^c\varepsilon_{ef}{}^{ab}\right)\,. \qquad (349)
$$

Производ функција знака у овом изразу намеће захтев да ивица $\tilde{\varepsilon}$ и троугао Δ чине тетраедар ненулте запремине. (За фиксан избор троугла постоје два тетраедра која га садрже и сваки од њих се добија тако што се ивица $\tilde{\varepsilon}$ може одабрати на три начина.) Због тога допринос суми даје само компонента производа тотално антисиметрична по индексима ових елемената, па се израз поједностављује и постаје:

$$
\tilde{\omega}^{[ab]}[\varepsilon] \;=\; -\frac{1}{V[\sigma]}\frac{1}{10}\sum_{\tilde{\varepsilon},\Delta,\tau\in\sigma}W[\tilde{\varepsilon},\tau]z[\tau,\Delta]\Delta^{ef}\tilde{\varepsilon}^q\varepsilon_q\varepsilon_{ef}{}^{[ab]} = -\frac{3}{20}\frac{1}{V[\sigma]}\sum_{\tau\in\sigma}\tau^{efq}\varepsilon_q\varepsilon_{ef}{}^{[ab]}\,. \qquad (350)
$$

Други сабирак у изразу (348) се може сматрати поправком вишег реда, док први сабирак даје допринос облика:

$$
\begin{aligned}
iS_{grav} &\approx \frac{i}{8\pi l_p^2} \sum_{\sigma} \frac{1}{10} \sum_{\Delta\in\sigma} \frac{\varepsilon_{[ab]cd}\Delta^{cd}}{A[\Delta]} A[\Delta] \frac{1}{3} \sum_{\varepsilon,\tilde{\Delta},\Delta_1\in\sigma} z[\tilde{\Delta},\varepsilon] W[\Delta,\Delta_1] W[\tilde{\Delta},\Delta_1] \frac{-3}{20} \frac{1}{V[\sigma]} \sum_{\tau\in\sigma} \tau^{efq} \varepsilon_q {\varepsilon_{ef}}^{[ab]} \\
&= \frac{i}{8\pi l_p^2} \sum_{\sigma} \sum_{\Delta\in\sigma} \frac{\Delta_{ef}}{A[\Delta]} A[\Delta] \frac{1}{10} \sum_{\Delta_1\in\sigma} W[\Delta,\Delta_1] \frac{1}{10} \sum_{\varepsilon,\tilde{\Delta}\in\sigma} z[\tilde{\Delta},\varepsilon] W[\tilde{\Delta},\Delta_1] \frac{1}{V[\sigma]} \sum_{\tau\in\sigma} \tau^{efq} \varepsilon_q \,. \qquad (351)
\end{aligned}
$$

Разломак $\frac{\Delta_{ef}}{A[\Delta]}$ одговара редукованом тензору Леви-Чивите на троуглу Δ, а производ три суме даје диедарски угао између два тетраедра над троуглом Δ, означен са $\theta_{\Delta,\sigma}$. Због тога је дејство за гравитационо поље (у апроксимацији доприноса првог сабирка) облика:

$$
S_{\text{grav}} \approx \frac{1}{8\pi l_p^2} \sum_{\sigma} \sum_{\Delta\in\sigma} \theta_{\Delta,\sigma} A[\Delta] = \frac{1}{8\pi l_p^2} \sum_{\Delta} A[\Delta] \sum_{\sigma\ni\Delta} \theta_{\Delta,\sigma} = \frac{1}{8\pi l_p^2} \sum_{\Delta} \delta_\Delta A[\Delta] = S_R \,, \qquad (352)
$$

где је δ_Δ дефицит угао (сума свих диедарских углова). У колико је простор раван, са сигнатуром простора Минковског, дефицит угао је једнак нули, док одступање овог угла од нуле указује на присуство ненулте кривине.

Добијено дејство S_R одговара дејству за Реџе гравитацију која описује класичну гравитацију на део по део равној многострукости. То значи да је у прелиминарној анализи семикласичног лимеса добијен резултат за очекивану вредност опсервабле облика:

$$
\langle F \rangle = \frac{\mathcal{N}}{Z} \prod_{v} \int d\phi[v] d\psi[v] d\bar{\psi}[v] \int \prod_{\varepsilon} d\alpha[\varepsilon] d\varepsilon \prod_{\sigma} |V[\sigma]|^{-M} \; F(\varphi_k) \exp\left(iS_R + iS_{ECC}^{(M)} \right) \,, \qquad (353)
$$

где је дејство $S_{ECC}^{(M)}$ дејство за Стандардни Модел са додатним сабирком који описује контактну спин-спин интеракцију, а $M = 150$.

Овиме је завршена прелиминарна анализа семикласичног лимеса у којој су добијена дејства класичних теорија у експоненту. Ова анализа указује да квантна 3BF теорија са везама има добар семикласичан лимес, јер је за сличне теорије које садрже само дејство за гравитационо поље у овом облику већ показано да имају добар семикласичан лимес [58, 59].

6 Закључак

Конструкција реалистичне квантне теорије гравитације и материје и даље представља отворен проблем у савременој теоријској физици.

У овом раду је дат предлог модела обједињене квантне теорије заснован на концепту тополошког 3BF дејства са везама које дају физичке степене слободе и математичког појма 3-групе, коришћеног за опис симетрије. Први резултат изложен у глави 3 се бавио испитивањем симетрије класичне теорије задате 3BF дејством са веземa, испитивањем механизама за њено нарушење, било експлицитно, додавањем веза у процесу конструкције реалистичне теорије, или спонтано Хигсовим механизмом. Добијени резултати су потврдили да је класична теорија добијена на овакав начин потпуно еквивалентна класичној теорији задатој Стандардним Моделом спрегнутим са Ајнштајн-Картановом гравитацијом. Ова еквиваленција је успостављена и на нивоу симетрије целог дејства и на нивоу једначина кретања.

Потврда ове еквиваленције је значила да је било могуће приступити испитивању еквиваленције између ових теорија на квантном нивоу. Ово питање је обрађено у следећој глави 4. Резултати анализе су потврдили да у извесној мери постоји еквиваленција између квантне и класичне теорије под одређеним условима. Први услов је да се мере у функционалним интегралима дефинишу на исти начин и други, да се посматрају само опсервабле које се могу истовремено дефинисати у обе теорије. Под тим условима су конструисане везе (226) између парова опсервабли са истом очекиваном вредношћу у квантној 3BF теорији са везама и квантној Ајнштајн-Картановој контактној теорији спрегнутој са Стандардним Моделом (ECC). Осим ове везе, дефиниција теорије садржи чланове у дејству који намећу нулте граничне услове на поља тетраде, Хигсова и Диракова поља. Ови гранични услови се могу променити додавањем одговарајућих површинских чланова у дејству, али је сама анализа тог поступка изван предмета истраживања изложеног овде. На крају, дати су илустративни примери на којима се концептуално може тестирати разлика у предвиђањима ове две квантне теорије у конкретном експерименту.

Коначно, након успостављања везе између две квантне теорије, последња глава 5 се бави конструкцијом квантне 3BF теорије и прелиминарном анализом њеног семикласичног лимеса. Конструкција квантне теорије је извршена триангулацијом многострукости и постављањем једначине за очекивану вредност произвољне опсервабле на таквој део-по-део равној многострукости. Дефинисана процедура је применљива на произвољно дејство изражено помоћу диференцијалних форми, што је важан резултат за конструкцију квантних теорија на триангулацији уопште. Главни резултат главе представља сама дефиниција квантне теорије (329) јер се комбиновањем ове дефиниције и везе (226) може репродуковати дефиниција квантне Ајнштајн-Картанове теорије спрегнуте са Стандардним Моделом на триангулацији. Испитивање семикласичног лимеса је спроведено применом апроксимације споропроменљивих поља над дефиницијом очекиване вредности произвољне опсервабле, чиме је добијено дејство класичне теорије у експоненту. На тај начин је резултат сведен на претходно проверене моделе са добрим семикласичним лимесом и оправдана је исправност конструкције квантне теорије (329). Међутим, за праву анализу семикласичног лимеса ове теорије, потребно је одредити одговарајуће ефективно дејство и испитати његов лимес према (333), што је ван предмета овог истраживања.

A Додаци

A.1 Доказ Теореме 1

Доказ. Особина G-инваријантности билинеарне форме $\langle \_ \, , \_ \rangle_{\mathfrak{h}}$ је дефинисана као

$$\langle g \triangleright h_1, g \triangleright h_2 \rangle_{\mathfrak{h}} = \langle h_1, h_2 \rangle_{\mathfrak{h}} \,, \tag{354}$$

за свако $g \in G$ и свако $h_1, h_2 \in \mathfrak{h}$. Записивањем компонената g, h_1 и h_2 у одговарајућем базису, једноставно је уочити да се лева страна једначине може записати као:

$$\langle g \triangleright h_1, g \triangleright h_2 \rangle_{\mathfrak{h}} = h_1^a h_2^c \Big[g_{ac} + g^\alpha \big(\triangleright_{\alpha c}{}^d g_{ad} + \triangleright_{\alpha a}{}^d g_{dc} \big) \Big] + \mathcal{O}(g^2) \,. \tag{355}$$

Изједначавањем са десном страном, очигледно је да се десна страна поништава са првим сабирком у загради. Затим се применом особине симетричности билинеарне форме остатак израза у загради своди на:

$$\triangleright_{\alpha ca} + \triangleright_{\alpha ac} = 0 \,, \tag{356}$$

како је и тврђено у теореми. Антисиметричност преостала два дејства $\triangleright_{\alpha\beta\gamma}$ и $\triangleright_{\alpha AB}$ се доказује аналогно.

Додатно, недегенерисаност билинеарне форме g_{ab} имплицира да постоји њен инверз, означен са g^{ab}. Затим, контракцијом (356) са g^{ac} добија се базис-независан исказ да је дејство $\triangleright$ бестражно

$$\triangleright_{\alpha a}{}^a = 0. \tag{357}$$

Поред тога, може се изабрати базис у Лијевој алгебри $\mathfrak{h}$ такав да је билинеарна форма g_{ab}, па самим тим и њен инверз, дијагонална. На основу идентитета

$$\triangleright_{\alpha a}{}^b = \triangleright_{\alpha ac} g^{cb} \tag{358}$$

се може закључити да, у овом конкретном базису, $\triangleright_{\alpha a}{}^b$ мора да буде пропорционално $\triangleright_{\alpha ab}$ јер је $g^{cb} = 0$ за $c \neq b$. Због тога, пошто је $\triangleright_{\alpha ab}$ антисиметричан према (356), мора бити једнак нули за $a = b$, што даље имплицира да је $\triangleright_{\alpha a}{}^b$ такође нула у том случају. Другим речима, $\triangleright_{\alpha a}{}^b$ има нулте елементе на дијагонали у односу на други и трећи индекс у овом базису, као што је и тврђено теоремом. Исте особине се за остала два дејства $\triangleright_{\alpha\beta}{}^\gamma$ и $\triangleright_{\alpha A}{}^B$ доказују аналогно. ☐

A.2 Произвољност κ-билинеарне форме

У одељку 3.2.4 је уведена билинеарна форма $\kappa^{\alpha\beta}$ у (167) као и нове променљиве (171) које задовољавају листу идентитета (172). Да би била обезбеђена један на један веза између старих и нових променљивих, у овим идентитетима је потребно да број независних променљивих са леве и десне стране једнакости буде исти. Штавише, ни једна од ових променљивих (171) не сме бити експлицитно помножена нулом током конструкције идентитета (172). Ови захтеви имају нетривијалне последице на произвољност избора билинеарне форме $\kappa^{\alpha\beta}$. На пример, први идентитет са листе (172) је

$$\theta^\alpha \wedge \rho_\alpha = \lambda^{\bar{A}} \wedge \gamma_{\bar{A}}. \tag{359}$$

Из дефиниције променљиве ρ_α у (171), може се закључити да дејство пројектора (168) не мења ρ_α

$$\rho_\alpha = P_\alpha{}^\beta \rho_\beta. \tag{360}$$

Ово имплицира да овај пројектор не мења леву страну идентитета (359) па важи:

$$\theta^{\alpha} \wedge \rho_{\alpha} = \theta^{\alpha} \wedge P_{\alpha}{}^{\beta} \rho_{\beta}. \tag{361}$$

Коришћењем захтева да ни једна променљива не сме да буде помножена нулом у току конструкције идентитета (359), закључује се да дејство пројектора мора да чува све променљиве θ^{α}. Даљим коришћењем дефиниције променљивих θ^{α} добија се нетривијални услов на билинеарну форму $\kappa^{\alpha\beta}$:

$$\theta^{\alpha} = \theta^{\beta} P_{\beta}{}^{\alpha} \qquad \Rightarrow \qquad -2\kappa^{\alpha\gamma} \triangleright_{\gamma}{}^{H}{}_{\bar{A}} \lambda^{\bar{A}} \equiv -2\kappa^{\alpha\gamma} P_{\gamma}{}^{\delta} \triangleright_{\delta}{}^{H}{}_{\bar{A}} \lambda^{\bar{A}} = -2 P_{\beta}{}^{\alpha} \kappa^{\beta\gamma} P_{\gamma}{}^{\delta} \triangleright_{\delta}{}^{H}{}_{\bar{A}} \lambda^{\bar{A}} \,, \tag{362}$$

који захваљујући произвољности поља $\lambda^{\bar{A}}$ имплицира да важи

$$\kappa^{\alpha\gamma} P_{\gamma}{}^{\delta} = P_{\beta}{}^{\alpha} \kappa^{\beta\gamma} P_{\gamma}{}^{\delta} \,. \tag{363}$$

Пошто је $\kappa^{\alpha\beta}$ симетрична билинеарна форма, транспоновањем (363) се добија да пројектор и $\kappa^{\alpha\beta}$ комутирају

$$\kappa^{\alpha\gamma} P_{\gamma}{}^{\delta} = P_{\gamma}{}^{\alpha} \kappa^{\gamma\delta} \,. \tag{364}$$

Сада се овај услов може искористити тако да се редукује произвољност билинеарне форме $\kappa^{\alpha\beta}$. Комбиновањем (169) и (364) се добија

$$P_{\gamma}{}^{\alpha} A^{\gamma\delta} \Big[\delta_{\delta}^{\beta} - P_{\delta}{}^{\beta} \Big] = \Big[\delta_{\gamma}^{\alpha} - P_{\gamma}{}^{\alpha} \Big] A^{\gamma\delta} P_{\delta}{}^{\beta} \,, \tag{365}$$

који је задовољен само ако је

$$A^{\alpha\delta} \Big[\delta_{\delta}^{\beta} - P_{\delta}{}^{\beta} \Big] = \Big[\delta_{\gamma}^{\alpha} - P_{\gamma}{}^{\alpha} \Big] A^{\gamma\beta} = \Big[\delta_{\gamma}^{\alpha} - P_{\gamma}{}^{\alpha} \Big] A^{\gamma\delta} \Big[\delta_{\delta}^{\beta} - P_{\delta}{}^{\beta} \Big] \,. \tag{366}$$

Будући да је димензија подпростора ортогоналног пројектора једнака један, произвољност билинеарне форме $\kappa^{\alpha\beta}$ је редукована на један слободан параметар, а облик ове произвољности је дат у (173):

$$\kappa^{\alpha\beta} \to \kappa^{\alpha\beta} + \Big[\delta_{\gamma}^{\alpha} - P_{\gamma}{}^{\alpha} \Big] A^{\gamma\delta} \Big[\delta_{\delta}^{\beta} - P_{\delta}{}^{\beta} \Big] \,. \tag{367}$$

Произвољност овог параметра гарантује постојање инвертибилне билинеарне форме $\kappa^{-1}_{\alpha\beta}$, која такође комутира са пројектором, што је последица (364).

Сада треба испитати утицај ове произвољности на дејство (174). Коришћењем чињенице да инверзна билинеарна форма делује на променљиве (171) у (174) и да су променљиве (171) непромењене при дејству пројектора, дејство (174) и матрица квадрата маса (176) зависе само од пројекције инверзне билинеарне форме $\kappa^{-1}_{\alpha\beta} P_{\gamma}{}^{\beta}$. Произвољност инверза билинеарне форме $\kappa^{-1}_{\alpha\beta}$ се може изразити као Тејлоров ред по произвољној билинеарној форми $A^{\gamma\delta}$ као

$$\kappa^{-1}_{\alpha\beta} \to \kappa^{-1}_{\alpha\gamma} \sum_{n=0}^{\infty} \Big[(-1)^n \left([\delta - P]\, A \, [\delta - P]\, \kappa^{-1} \right)^n \Big]^{\gamma}{}_{\beta} \,, \tag{368}$$

одакле се може уочити да пројекција инверзне билинеарне форме не зависи од избора произвољне билинеарне форме $A^{\gamma\delta}$. Ово даље имплицира да су дејство (174) и матрица квадрата маса (176) једнозначно дефинисани.

A.3 Докази идентитета који садрже Диракову делту

Докази идентитета који садрже Диракову делту над реалним и Грасмановим пољима се могу извести праволинијским рачуном:

- идентитет (199):

$$
\begin{aligned}
&\int D\varphi D\phi_k\, \delta\left(\varphi^{aB} F_B{}^A(\phi_k) - G^{aA}(\phi_k)\right) H(\varphi, \phi_k) \\
&= \int D\left(\varphi^{aB} F_B{}^A(\phi_k)\right) D\phi_k \begin{vmatrix} \frac{\delta\left(\varphi F(\phi_k) F^{-1}(\phi_k)\right)^{aA}}{\delta \phi_n} & \frac{\delta \phi_m}{\delta \phi_n} \\ \frac{\delta\left(\varphi F(\phi_k) F^{-1}(\phi_k)\right)^{aA}}{\delta(\varphi F(\phi_k))^{bB}} & 0 \end{vmatrix} \delta\left(\varphi^{aB} F_B{}^A(\phi_k) - G^{aA}(\phi_k)\right) H(\varphi, \phi_k) \\
&= \int D\hat{\varphi} D\phi_k \left|\delta^a_b F^{-1}{}_B{}^A(\phi_k)\right| \delta\left(\hat{\varphi}^{aA} - G^{aA}(\phi_k)\right) H\left(\hat{\varphi}^{aB} F^{-1}{}_B{}^A(\phi_k), \phi_k\right) \\
&= \int D\phi_k \frac{1}{|F(\phi_k)|^{|a|}} H\left(G^{aB}(\phi_k) F^{-1}{}_B{}^A(\phi_k), \phi_k\right)\ ,
\end{aligned} \tag{369}
$$

- идентитет (201):

$$
\begin{aligned}
&\int_{\mathbb{G}^n} d\theta_1 d\theta_2 d\theta_3 \dots d\theta_n e^{i\theta_1(\theta_2 - \theta_3 - \dots - \theta_k)} F(\theta_2, \theta_3, \dots, \theta_n) \\
&= \int_{\mathbb{G}^n} d\theta_1 d\theta_2 d\theta_3 \dots d\theta_n \left(1 + i\theta_1\left(\theta_2 - \theta_3 - \dots - \theta_k\right)\right) \\
&\quad \times \left(\theta_3 \dots \theta_k f_{01\dots1}(\theta_{k+1}, \dots, \theta_n) + \dots + \theta_2 \dots \theta_{k-1} f_{1\dots10}(\theta_{k+1}, \dots, \theta_n)\right) \\
&= i \int_{\mathbb{G}^n} d\theta_1 d\theta_2 d\theta_3 \dots d\theta_n\, \theta_1\theta_2\theta_3 \dots \theta_k \left(f_{01\dots1}(\theta_{k+1}, \dots, \theta_n)\right. \\
&\quad \left. + \dots + (-1)^l f_{11\dots10_l1\dots1}(\theta_{k+1}, \dots, \theta_n)\right) \\
&= -i \int_{\mathbb{G}^n} d\theta_3 \dots d\theta_n d\theta_2 d\theta_1\, \theta_1\theta_2\theta_3 \dots \theta_k \left(f_{01\dots1}(\theta_{k+1}, \dots, \theta_n)\right. \\
&\quad \left. + \dots + (-1)^l f_{11\dots10_l1\dots1}(\theta_{k+1}, \dots, \theta_n)\right) \\
&= -i \int_{\mathbb{G}^{n-2}} d\theta_3 \dots d\theta_n\, \theta_3 \dots \theta_k \left(f_{01\dots1}(\theta_{k+1}, \dots, \theta_n) + \dots + (-1)^l f_{11\dots10_l1\dots1}(\theta_{k+1}, \dots, \theta_n)\right) \\
&= -i \int_{\mathbb{G}^{n-2}} d\theta_3 \dots d\theta_n\, F(\theta_3 + \dots + \theta_k, \theta_3, \dots, \theta_n) \\
&= -i \int_{\mathbb{G}^{n-1}} d\theta_3 \dots d\theta_n d\theta_2\, \delta(\theta_2 - \theta_3 - \dots - \theta_k) F(\theta_2, \theta_3, \dots, \theta_n) \\
&= (-1)^{n-1} i \int_{\mathbb{G}^{n-1}} d\theta_2 \dots d\theta_n\, \delta(\theta_2 - \theta_3 - \dots - \theta_k) F(\theta_2, \theta_3, \dots, \theta_n)\,,
\end{aligned} \tag{370}
$$

• идентитет (202):

$$
\begin{aligned}
&\int_{\mathbb{R}^m} d^m y \int_{\mathbb{R}^m} d^m x \int_{\mathbb{G}^k} d\theta_1 d\theta_2 \dots d\theta_k \int_{\mathbb{G}^{n-k}} d\theta_{k+1} \dots d\theta_n \, e^{iy_a\left(x^a - M^{aij}\theta_i\theta_j\right)} F(x, \theta_1, \dots, \theta_n) \\
&= \int_{\mathbb{R}^m} d^m y \int_{\mathbb{R}^m} d^m x \int_{\mathbb{G}^k} d\theta_1 d\theta_2 \dots d\theta_k \int_{\mathbb{G}^{n-k}} d\theta_{k+1} \dots d\theta_n \, e^{iy_a x^a} \\
&\qquad \times \sum_{b=0}^{+\infty} \frac{1}{b!} \left(-iy_a M^{aij}\theta_i\theta_j\right)^b F(x, \theta_1, \dots, \theta_n) \\
&= \int_{\mathbb{R}^m} d^m y \int_{\mathbb{R}^m} d^m x \int_{\mathbb{G}^k} d\theta_1 d\theta_2 \dots d\theta_k \int_{\mathbb{G}^{n-k}} d\theta_{k+1} \dots d\theta_n \, e^{iy_a x^a} \\
&\qquad \times \sum_{b=0}^{+\infty} \frac{1}{b!} \left(M^{aij}\theta_i\theta_j \frac{\partial}{\partial x^a}\right)^b F(x, \theta_1, \dots, \theta_n) \\
&= (2\pi)^m \int_{\mathbb{R}^m} d^m x \int_{\mathbb{G}^k} d\theta_1 d\theta_2 \dots d\theta_k \int_{\mathbb{G}^{n-k}} d\theta_{k+1} \dots d\theta_n \prod_{a=1}^{m} \delta(x^a) \\
&\qquad \times \sum_{b=0}^{+\infty} \frac{1}{b!} \left(M^{aij}\theta_i\theta_j \frac{\partial}{\partial x^a}\right)^b F(x, \theta_1, \dots, \theta_n) \\
&= (2\pi)^m \int_{\mathbb{G}^k} d\theta_1 d\theta_2 \dots d\theta_k \int_{\mathbb{G}^{n-k}} d\theta_{k+1} \dots d\theta_n \, \sum_{b=0}^{+\infty} \frac{1}{b!} \left(M^{aij}\theta_i\theta_j \frac{\partial}{\partial x^a}\right)^b F(x, \theta_1, \dots, \theta_n)\Big|_{x=0} \\
&= (2\pi)^m \int_{\mathbb{G}^k} d\theta_1 d\theta_2 \dots d\theta_k \int_{\mathbb{G}^{n-k}} d\theta_{k+1} \dots d\theta_n \, F\left(M^{aij}\theta_i\theta_j, \theta_1, \dots, \theta_n\right) \\
&= (2\pi)^m \int_{\mathbb{R}^m} d^m x \int_{\mathbb{G}^k} d\theta_1 d\theta_2 \dots d\theta_k \int_{\mathbb{G}^{n-k}} d\theta_{k+1} \dots d\theta_n \prod_{a=1}^{m} \delta\left(x^a - M^{aij}\theta_i\theta_j\right) F(x, \theta_1, \dots, \theta_n) \, .
\end{aligned}
\tag{371}
$$

Литература

[1] J. Polchinski, *String Theory Vol. I: An Introduction to the Bosonic String*, Cambridge University Press, Cambridge (1998).

[2] J. Polchinski, *String Theory Vol. II: Superstring Theory and Beyond*, Cambridge University Press, Cambridge (1998).

[3] C. Rovelli, *Quantum Gravity*, Cambridge University Press, Cambridge (2004).

[4] T. Thiemann, *Modern Canonical Quantum General Relativity*, Cambridge University Press, Cambridge (2007).

[5] C. Rovelli and F. Vidotto, *Covariant Loop Quantum Gravity*, Cambridge University Press, Cambridge (2014).

[6] E. Bianchi, M. Han, E. Magliaro, C. Perini, C. Rovelli and W. Wieland, "Spinfoam fermions", *Class. Quant. Grav.* **30**, 235023 (2013), `arXiv:1012.4719`.

[7] M. Vojinović, "Cosine problem in EPRL/FK spin foam model", *Gen. Relativ. Gravit.* **46**, 1616 (2014), `arXiv:1307.5352`.

[8] A. Miković and M. Vojinović, *State-Sum Models of Piecewise Linear Quantum Gravity*, World Scientific, Singapore (2023).

[9] J. C. Baez and J. Huerta, "An Invitation to Higher Gauge Theory", *Gen. Relativ. Gravit.* **43**, 2335 (2011), `arXiv:1003.4485`.

[10] L. Crane and M. D. Sheppeard, "2-Categorical Poincaré representations and state sum applications", `arXiv:math/0306440`.

[11] A. Miković and M. Vojinović, "Poincaré 2-group and quantum gravity", *Class. Quant. Grav.* **29**, 165003 (2012), `arXiv:1110.4694`.

[12] Z. Li, "A global geometric approach to parallel transport of strings in gauge theory", `arxiv:1910.14230`.

[13] C, Saemann and M. Wolf, "Non-abelian tensor multiplet equations from twistor space", *Commun. Math. Phys.* **328**, 527 (2014), `arxiv:1205.3108`.

[14] B. Jurco, C. Saemann and M. Wolf, "Semistrict Higher Gauge Theory", *JHEP* **87** (2005), `arxiv:1403.7185`.

[15] Y. Hidaka, M. Nitta and R. Yokokura, "Higher-form symmetries and 3-group in axion electrodynamics", *Phys. Lett. B* **808**, 135672 (2020), `arxiv:2006.12532`.

[16] Y. Hidaka, M. Nitta and R. Yokokura, "Global 3-group symmetry and 't Hooft anomalies in axion electrodynamics", *JHEP* **1**, 173 (2021), `arxiv:2009.14368`.

[17] C. Saemann and M. Wolf, "Six-Dimensional Superconformal Field Theories from Principal 3-Bundles over Twistor Space", *Lett. Math. Phys.* **104**, 1147 (2014), `arxiv:1305.4870`.

[18] D. Song, K. Wu and J. Yang, "3-form Yang-Mills based on 2-crossed modules", *Jour. Geom. Phys.* **178**, 104537 (2022), arxiv:2108.12852.

[19] D. Song, K. Wu and J. Yang, "Higher form Yang-Mills as higher BFYM theories", *Eur. Phys. Jour. C* **82**, 1034 (2022), arxiv:2109.13443.

[20] Y. Hidaka, M. Nitta and R. Yokokura, "Topological axion electrodynamics and 4-group symmetry", *Phys. Lett. B* **823**, 136762 (2021), arxiv:2107.08753.

[21] Y. Hidaka, M. Nitta and R. Yokokura, "Global 4-group symmetry and 't Hooft anomalies in topological axion electrodynamics", *Prog. Theor. Exper. Phys.* **2022**, 04A109 (2022), arxiv:2108.12564.

[22] B. Jurco, C. Saemann and M. Wolf, "Higher Groupoid Bundles, Higher Spaces, and Self-Dual Tensor Field Equations", *Fortsch. Phys.* **64**, 674 (2016), arxiv:1604.01639.

[23] C. Saemann and M. Wolf, "Supersymmetric Yang-Mills Theory as Higher Chern-Simons Theory", *JHEP* **07**, 111 (2017), arxiv:1702.04160.

[24] B. Jurco, T. Macrelli, L. Raspollini, C. Saemann and M. Wolf, "L-infinity Algebras, the BV Formalism, and Classical Fields", *Fortsch. Phys.* **67**, 1910025 (2019), arxiv:1903.02887.

[25] T. Radenković and M. Vojinović, "Higher Gauge Theories Based on 3-groups", *JHEP* **10**, 222 (2019), arXiv:1904.07566.

[26] T. Radenković and M. Vojinović, "Topological invariant of 4-manifolds based on a 3-group", *JHEP* **7**, 105 (2022), arXiv:2201.02572.

[27] J. F. Plebanski, "On the separation of Einsteinian substructures", *Jour. Math. Phys.* **18**, 2511 (1977).

[28] J. C. Baez, "An Introduction to Spin Foam Models of Quantum Gravity and BF Theory", *Lect. Notes Phys.* **543**, 25 (2000), arXiv:gr-qc/9905087.

[29] M. Celada, D. González and M. Montesinos, "BF gravity", *Class. Quant. Grav.* **33**, 213001 (2016), arXiv:1610.02020.

[30] J. Baez, "4-Dimensional BF Theory as a Topological Quantum Field Theory", *Lett. Math. Phys.* **38**, 129 (1996), arxiv:q-alg/9507006.

[31] F. Girelli, H. Pfeiffer and E. M. Popescu, "Topological higher gauge theory – from BF to BFCG theory", *Jour. Math. Phys.* **49**, 032503 (2008), arXiv:0708.3051.

[32] J. F. Martins and A. Miković, "Lie crossed modules and gauge-invariant actions for 2-BF theories", *Adv. Theor. Math. Phys.* **15**, 1059 (2011), arXiv:1006.0903.

[33] A. Miković and M. A. Oliveira, "Canonical formulation of Poincaré BFCG theory and its quantization", *Gen. Relativ. Gravit.* **47**, 58 (2015), arXiv:1409.3751.

[34] A. Miković, M. A. Oliveira and M. Vojinović, "Hamiltonian analysis of the $BFCG$ theory for the Poincaré 2-group", *Class. Quant. Grav.* **33**, 065007 (2016), arxiv:1508.05635.

[35] A. Miković, M. A. Oliveira and M. Vojinović, "Hamiltonian analysis of the BFCG theory for a generic Lie 2-group", *Adv. Theor. Math. Phys.* **26**, 3783 (2022), `arXiv:1610.09621`.

[36] A. Miković, M. A. Oliveira and M. Vojinović, "Hamiltonian analysis of the BFCG formulation of General Relativity", *Class. Quant. Grav.* **36**, 015005 (2019), `arXiv:1807.06354.`

[37] S. K. Asante, B. Dittrich, F. Girelli, A. Riello and P. Tsimiklis, "Quantum geometry from higher gauge theory", *Class. Quant. Grav.* **37**, 205001 (2020), `arxiv:1908.05970`.

[38] F. Girelli, M. Laudonio and P. Tsimiklis, "Polyhedron phase space using 2-groups: kappa-Poincare as a Poisson 2-group", `arxiv:2105.10616`.

[39] A. Miković and M. Vojinović, "Standard Model and 4-groups", *Europhys. Lett.* **133**, 61001 (2021), `arXiv:2008.06354.`

[40] T. Porter, "Topological quantum field theories from homotopy n-types", *J. London Math. Soc.* **58**, 723 (1998), `MR 1678163`.

[41] T. Radenković, M. Vojinović, "The 3BF theory as a TQFT", *J. Phys. A: Math. Theor.* **58**, 345203 (2025), `arXiv:2412.21032`.

[42] M. Blagojević, *Gravitation and Gauge Symmetries*, Institute of Physics Publishing, Bristol (2002).

[43] T. Radenković and M. Vojinović, "Hamiltonian Analysis for the Scalar Electrodynamics as $3BF$ Theory", *Symmetry* **12**, 620 (2020), `arXiv:2004.06901.`

[44] T. Radenković and M. Vojinović, "Gauge symmetry of the $3BF$ theory for a generic Lie 3-group", *Class. Quant. Grav.* **39**, 135009 (2022), `arXiv:2101.04049.`

[45] J. F. Martins and R. Picken, "The fundamental Gray 3-groupoid of a smooth manifold and local 3-dimensional holonomy based on a 2-crossed module", *Differ. Geom. Appl. Jour.* **29**, 179-206 (2011), `arXiv:0907.2566.`

[46] W. Wang, "On 3-gauge transformations, 3-curvatures and Gray-categories", *Jour. Math. Phys.* **55**, 043506 (2014).

[47] P. Stipsić and M. Vojinović, "Symmetry breaking mechanisms of the 3BF action for the Standard Model coupled to gravity", *Jour. Phys. A: Math. Theor.* **58**, 215201 (2025), `arXiv:2402.17675`.

[48] P. Stipsić and M. Vojinović, "Correspondence between 3BF and Einstain-Cartan formulations of quantum gravity", *Class. Quantum Grav.* **42**, 195009 (2025), `arXiv:2506.17722`.

[49] J. Baez and J. Dolan, "Higher dimensional algebra and Topological Quantum Field Theory", *Jour. Math. Phys.* **36**, 6073 (1995), `arxiv:q-alg/9503002`.

[50] T. Radenković and M. Vojinović, "Quantum gravity and elementary particles from higher gauge theory", *Ann. Univ. Craiova Phys.* **30**, 74 (2020), `arXiv:2103.08037`.

[51] M. Vasilić and M. Vojinović, "Spinning branes in Riemann-Cartan spacetime", *Phys. Rev. D* **78**, 104002 (2008), `arXiv:1010.1861`.

[52] M. Đorđević, T. Radenković etal, "Henneaux-Teitelboim gauge symmetry and its applications to higher gauge theories", *Universe* **9**, 281 (2023), arXiv:2305.00117.

[53] J. W. York, "Role of Conformal Three-Geometry in the Dynamics of Gravitation", *Phys. Rev. Lett.* **28**, 1082 (1972).

[54] G. W. Gibbons and S. W. Hawking, "Action integrals and partition functions in quantum gravity", *Phys. Rev. D* **15**, 2752 (1977).

[55] J. Erdmenger, B. Heß, I. Matthaiakakis and R. Meyer, "Universal Gibbons-Hawking-York term for theories with curvature, torsion and non-metricity", *SciPost Phys.* **14**, 099 (2023), arXiv:2211.02064.

[56] B. P. Abbott etal, "Observation of Gravitational Waves from a Binary Black Hole Merger", *Phys. Rev. Lett.* **116**, 061102 (2016). arXiv:1602.03837.

[57] V. Gerdt, R. Horan etal, "On the Hamiltonian reduction of geodesic motion on SU(3) to SU(3)/SU(2)", *J. Math. Phys.* **47**, 112902 (2006), arXiv:hep-th/0511245.

[58] A. Miković, M. Vojinović, "Effective action and semiclassical limit of spin foam models", *Class. Quant. Grav.* **28**, 225004 (2011), arXiv:1104.1384.

[59] A. Miković, "Effective Actions for Regge Piecewise Flat Quantum Gravity", *Universe* **8(5)**, 268 (2022), arXiv:2203.12313.

Биографија

Павле Стипсић је рођен 19.10.1996. године у Алексинцу, где је завршио основну школу и основну музичку школу. Септембра 2011. године је уписао и током наредне четири године завршио Математичку гимназију и теоријски одсек средње музичке школе Мокрањац у Београду. Основне академске студије на Физичком факултету универзитета у Београду, смер Теоријска и експериментална физика и основне академске студије на Електротехничком факулету универзитета у Београду, смер Електротехника и рачунарство, модул Електроника, уписао је паралелно 2015. године. Основне студије физике је завршио септембра 2019. године, са просечном оценом 9,51, а на Електротехничком факултету основне студије је завршио у марту 2023. године, са просечном оценом 8,61.

Мастер академске студије је завршио на Физичком факултету у периоду од октобра 2019. године до септембра 2020. године, са просечном оценом 10,0, одбранивши мастер рад на тему "Утицај магнетног поља на проводност у Хабардовом моделу", под руководством др Јакше Вучичевића, вишег научног сарадника Института за физику у Београду. Уписао је докторске студије физике на Физичком факултету у Београду, у ужој научној области Квантна поља, честице и гравитација октобра 2020. године.

Од априла 2021. године је запослен као истраживач на Институту за физику у Београду, у групи за Гравитацију, честице и поља, чији је руководилац др Бранислав Цветковић, научни саветник Института за физику у Београду. Бави се научним истраживањем под руководством др Марка Војиновића, научног саветника Института за физику у Београду, радећи на темама везаним за испитивање и конструкцију модела квантне гравитације.

Досадашњи научно-истраживачки рад Павла Стипсића се може класификовати у три области:

(1) изучавање фононски индукованих спинских релаксационих процеса у спинсим кубитима,

(2) изучавање проводности у Хабардовом моделу,

(3) изучавање симетрија и особина квантне теорије гравитације формулисане преко $3BF$ модела у контексту виших градијентних теорија.

Прве две области спадају у његов истраживачки рад везан за мастер тезу, док је трећа област везана за истраживачки рад у оквиру рада на докторској тези.

У оквиру изучавања фононски индукованих спинских релаксационих процеса у спинским кубитима, посматран је утицај симетрије квантне тачке на Рабијеву фреквенцију и фононски индуковане спинске релаксационе процесе у једноелектронском спинском кубиту и утицај избора контролног потенцијала на зависност времена живота од правца магнетног поља за неколико дискретних симетрија кубита на конкретним примерима троугаоног, квадратног и правоугаоног потенцијала јаме [1, 2].

У оквиру изучавања проводности материјала који се могу моделовати Хабардовим моделом, нумеричким симулацијама је испитивана зависност проводности материјала од примењеног спољашњег магнетног поља у режиму слабих електрон-електрон интеракција на чворовима решетке на различитим температурама. Такође, аналитички је решена инверзна Фуријеова трансформација по Мацубара фреквенцијама Лорановог развоја Гринових функција на коначној температури за фермионе и бозоне [3].

Коначно, у оквиру изучавања квантне теорије гравитације преко виших градијентних теорија, проучаван је механизам експлицитног и спонтаног нарушења симетрије у $3BF$ теоријама са везама, и разматран је Хигсов механизам. Такође је дат допринос изучавању Ено-Тајтелбом симетрије. Затим, детаљно је проучена веза између квантне $3BF$ теорије са везама и квантне Ајнштајн-Картанове теорије са контактном интеракцијом. Коначно, формулисан је експлицитан поступак квантизације $3BF$ теорије са везама, дефинисањем интеграла по трајекторијама за целу теорију и урађена је прелиминарна анализа семикласичног лимеса теорије [4, 5, 6].

Списак публикација

Радови у међународним часописима

[1] S. Miladić, P. Stipsić, E. Dobardžić i M. Milivojević,
"Electrical control of a spin qubit in InSb nanowire quantum dots: Strongly suppressed spin relaxation in high magnetic field",
Phys. Rev. B **101**, 155307 (2020).
[DOI: 10.1103/PhysRevB.101.155307] [arXiv:1911.07942] [M21, **IF=4.036**]

[2] P. Stipsić i M. Milivojević,
"Control of a spin qubit in a lateral GaAs quantum dot based on symmetry of gating potential",
Phys. Rev. B **101**, 165302 (2020).
[DOI: 10.1103/PhysRevB.101.165302] [arXiv:2001.03900] [M21, **IF=4.036**]

[3] J. Vučičević, P. Stipsić i M. Ferrero,
"Analytical solution for time integrals in diagrammatic expansions: Application to real-frequency diagrammatic Monte Carlo",
Phys. Rev. Research **3**, 023082 (2021).
[DOI: 10.1103/PhysRevResearch.3.023082] [arXiv:2011.08226] [M21, **IF=N/A**]

[4] M. Đorđević, T. Radenković, P. Stipsić i M. Vojinović,
"Henneaux–Teitelboim gauge symmetry and its applications to higher gauge theories",
Universe **9**, 281 (2023).
[DOI: 10.3390/universe9060281] [arXiv:2305.00117] [M22, **IF=2.9**]

[5] P. Stipsić i M. Vojinović,
"Symmetry breaking mechanisms of the 3BF action for the Standard Model coupled to gravity",
Jour. Phys. A: Math. Theor. **58**, 215201 (2025).
[DOI: 10.1088/1751-8121/add7a5] [arXiv:2402.17675] [M21, **IF=2.1**]

[6] P. Stipsić i M. Vojinović,
"Correspondence between 3BF and Einstein-Cartan formulations of quantum gravity",
Class. Quant. Grav. **42**, 195009 (2025).
[DOI: 10.1088/1361-6382/ae06e9] [arXiv:2506.17722] [M21, **IF=3.7**]

Конференцијски радови

[7] P. Stipsić,
"Metod modelovanja rezonatora sa dve bliske frekvencije primenom narušenja simetrije",
Program i zbornik apstrakta LXIX konferencije ETRAN (2025).
[DOI: 10.69994/69E25021] [N/A, **IF=N/A**]

образац изјаве о ауторству

Изјава о ауторству

Име и презиме аутора Павле Стипсић

Број индекса 2020/8004

Изјављујем

да је докторска дисертација под насловом

Симетрије у вишим градијентним теоријама

- резултат сопственог истраживачког рада;
- да дисертација у целини ни у деловима није била предложена за стицање друге дипломе према студијским програмима других високошколских установа;
- да су резултати коректно наведени и
- да нисам кршио/ла ауторска права и користио/ла интелектуалну својину других лица.

Потпис аутора

У Београду, 27. 2. 2026.

Павле Стипсић

Изјава о истоветности штампане и електронске верзије докторског рада

Име и презиме аутора Павле Стипсић

Број индекса 2020/8004

Студијски програм Квантна поља, честице и гравитација

Наслов рада Симетрије у вишим градијентним теоријама

Ментор др Марко Војиновић

Изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/ла ради похрањивања у **Дигиталном репозиторијуму Универзитета у Београду.**

Дозвољавам да се објаве моји лични подаци везани за добијање академског назива доктора наука, као што су име и презиме, година и место рођења и датум одбране рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Универзитета у Београду.

Потпис аутора

У Београду, 27. 2. 2026.

Павле Стипсић

образац изјаве о коришћењу

Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић" да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

Симетрије у вишим градијентним теоријама

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигиталном репозиторијуму Универзитета у Београду и доступну у отвореном приступу могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио/ла.

1. Ауторство (CC BY)

2. Ауторство – некомерцијално (CC BY-NC)

3. Ауторство – некомерцијално – без прерада (CC BY-NC-ND)

4. Ауторство – некомерцијално – делити под истим условима (CC BY-NC-SA)

5. Ауторство – без прерада (CC BY-ND)

6. Ауторство – делити под истим условима (CC BY-SA)

(Молимо да заокружите само једну од шест понуђених лиценци.
Кратак опис лиценци је саставни део ове изјаве).

Потпис аутора

У Београду, 27. 2. 2026.

[signature]

1. **Ауторство**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце, чак и у комерцијалне сврхе. Ово је најслободнија од свих лиценци.

2. **Ауторство – некомерцијално**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела.

3. **Ауторство – некомерцијално – без прерада**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела. У односу на све остале лиценце, овом лиценцом се ограничава највећи обим права коришћења дела.

4. **Ауторство – некомерцијално – делити под истим условима**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца не дозвољава комерцијалну употребу дела и прерада.

5. **Ауторство – без прерада**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца дозвољава комерцијалну употребу дела.

6. **Ауторство – делити под истим условима**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца дозвољава комерцијалну употребу дела и прерада. Слична је софтверским лиценцама, односно лиценцама отвореног кода.