

Supersymmetry and trace formulas II. Selberg trace formula

CHANGHA CHOI AND LEON A. TAKHTAJAN

By extending the new supersymmetric localization principle introduced in [2], we present a path integral derivation of the Selberg trace formula on arbitrary compact Riemann surfaces, including the case of vector-valued automorphic forms of arbitrary half-integer weight corresponding to Maass Laplacian. We also generalize the method to formulate the Selberg trace formula on generic compact locally symmetric space.

1	Introduction	871
2	Extended new supersymmetric localization principle	873
3	Propagator on non-compact semisimple G	875
4	Selberg trace formula on compact Riemann surfaces	880
5	Selberg trace formula on general compact locally symmetric space	901
	Appendix A Laplace operators on $SL(2, \mathbb{R})$ and $SL(2, \mathbb{R})/SO(2)$	911
	References	914

1. Introduction

Selberg trace formula [3] exhibits deep and highly non-trivial spectral/geometric duality on Riemann surfaces of constant negative curvature, which relates the spectrum of Laplace operator and closed geodesics. The information on the eigenvalues of the Laplacian is only accessible by numerical methods for

generic Riemann surfaces, and the Selberg trace formula is the only available analytic tool to uncover the detailed structure of the spectrum.

The possible implication in physics is extremely tantalizing, since the trace formula can be reformulated as a non-trivial identity regarding the thermal partition function of the purely bosonic system. The Selberg trace formula can be thought of as a chaotic version of the trace formula as opposed to the Eskin trace formula on compact Lie groups, since the particle moving on compact Riemann surfaces is chaotic both classically and quantum mechanically. Note that most physical systems of utmost interest (from Yang-Mills theory to black holes) are not integrable nor having exactly solvable spectrum. Therefore, it is desirable to understand the precise physical origin of the Selberg trace formula from the quantum mechanical point of view, which might give some salient physical lessons.

It is important to highlight the remarkable observation made by Gutzwiller [1], who discovered that the part of the Selberg trace formula related to the hyperbolic elements coincides with the exact semiclassical contribution to the partition function by closed geodesics, up to DeWitt term. This led him to formulate the ‘Gutzwiller trace formula’, which expresses, at the semiclassical level, the spectrum of a quantum system in terms of periodic orbits of the corresponding classical chaotic system. However, understanding in this approach the appearance of equally important remaining terms in the Selberg trace formula was a mystery.

The goal of the present paper is to provide a physical origin of the Selberg trace formula in the light of the path integral formalism. We will demonstrate that the Selberg trace formula, in its precise form, can be derived from the new supersymmetric localization principle, formulated in [2] and successfully applied to the Eskin trace formula for compact Lie groups.

The organization of the paper is the following. In Section 2, we streamline and extend the new supersymmetric localization principle presented in [2]. In Section 3, we study the propagator on non-compact semisimple Lie group by applying the new localization principle to certain supersymmetric sigma model, which will be the stepping stone for the rest of the paper. In Section 4, we carefully examine the relevant model and the corresponding observables that give the Selberg trace formula on compact Riemann surface, and successfully derive it using the localization principle. Finally, in Section 5 we use localization principle to formulate the Selberg trace formula on general compact locally symmetric space, with a non-trivial check for the case of compact hyperbolic 3-manifolds.

Acknowledgments. The first author (C.C.) thanks J. Gomis, Z. Komarogodski, K. Lee, N. Lee, and P. Yi for discussions. The research of C.C. was supported by the Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development and by the Province of Ontario through the Ministry of Colleges and Universities. The second author (L.T.) thanks A. Alekseev, S. Shatashvili, and D. Sullivan for discussions.

2. Extended new supersymmetric localization principle

Here we formulate a generalized version of the new supersymmetric localization principle, first presented in [2], which is applicable to a large class of non-supersymmetric correlation functions of supersymmetric QFT in any spacetime dimension.

Specifically, consider a supersymmetric QFT which has at least one Hermitian supercharge \hat{Q} . We are interested in general non-supersymmetric correlation function made out of operators (or defects) with arbitrary codimensions. This situation is quite natural, since usual supersymmetric correlation functions with $(-1)^F$ can be thought of as insertions of codimension 1 (topological) defects.

Let

$$(1) \quad \langle \hat{\mathcal{O}}(\{\mathcal{M}_0, \mathcal{M}_1, \dots, \mathcal{M}_d\}) \rangle_E = \int \mathcal{O}(\{\mathcal{M}_0, \mathcal{M}_1, \dots, \mathcal{M}_d\}) e^{-S_E} \mathcal{D}\mu,$$

be generic correlation function of certain d -dimensional reflection positive Euclidean QFT, where M_j denotes a union of the domain of all codimension $d - j$ operators and $\mathcal{D}\mu$ is some ‘integration measure’. We assume that the theory is supersymmetric, $\delta S_E = 0$, where δ is the supersymmetry generator, which in the operator formalism is encoded into the definition of the bracket, and $\delta(\mathcal{D}\mu) = 0$.

A correlation function (1) is called supersymmetric if the integrand is δ -closed,

$$\delta \mathcal{O}(\{\mathcal{M}_0, \mathcal{M}_1, \dots, \mathcal{M}_d\}) = 0.$$

The standard localization principle is the statement that supersymmetric correlation function is invariant under the change $S_E \mapsto S_E + \lambda \delta V$ with an arbitrary positive parameter λ , where a deformation V is such $\delta^2 V = 0$ and δV satisfies the standard positivity property.

The main idea of [2] was to extend this deformation invariance to the case of non-supersymmetric observables and to formulate a new supersymmetric localization principle. This was achieved for systems having fermionic zero modes in the action, and to have a non-vanishing correlation function, we must require that \mathcal{O} saturates all fermionic zero modes. We formulate it as a statement that non-supersymmetric path integral (1) is invariant under the change $S_E \rightarrow S_E + \lambda\delta V$ where, in addition to $\delta^2 V = 0$, the deformation V has the property that both $V\delta\mathcal{O}$ and δV do not saturate fermionic zero modes. It is easy to establish this generalized localization principle along the lines in [2, Section 2].

It is important to observe that the localization argument outlined above works not only for the Euclidean path integral, but is also applicable to the Lorentzian path integral, where the weighting function is e^{iS_L} with Minkowski (real time) action S_L . This framework is useful when the observable \mathcal{O} doesn't make sense in Euclidean time, or even if it does, there is no natural Euclidean path integral associated with it (without using some analytic continuation). Here we consider an example of the Lorentzian path integral

$$(2) \quad \langle \hat{\mathcal{O}}(\{\mathcal{M}_0, \mathcal{M}_1, \dots, \mathcal{M}_d\}) \rangle_L = \int \mathcal{O}(\{\mathcal{M}_0, \mathcal{M}_1, \dots, \mathcal{M}_d\}) e^{iS_L} \mathcal{D}\mu,$$

which is invariant under the change $S_L \rightarrow S_L + \lambda\delta V$, provided the deformation V satisfies the conditions stated above.

We remark that this approach is quite natural from both physics and mathematics perspectives. Namely, the path integral was first formulated by Feynman in terms of the physical time, where an integrand has an oscillatory nature. This path integral, in semi-classical approximation, can be evaluated by the stationary phase method, based on the mathematical idea that rapid oscillations of an integrand sum up to a negligible contribution.¹ In mathematics, it was first observed by Duistermaat and Heckman [4] that for a certain class of finite-dimensional oscillatory integrals the stationary phase method gives an exact answer. This idea was subsequently generalized by Atiyah and Bott [5] and by Berline and Vergne [6], and led to a mathematical notion of a localization.

¹In principle, there could also be a contribution from 'complex' critical points which needs the method of steepest descent whose path integral analogue has been discussed in [7]. Following the standard logic, for all examples considered in the paper it can be argued that only 'real' critical points contribute for our choice of the 'real' contour, since $\delta V = 0$ at all possible critical points. It would be interesting to explore cases where such analytic continuation is necessary.

Presented extension of the localization principle to non-supersymmetric observables is quite encouraging. We believe it can be successfully applied to even larger class of non-supersymmetric observables, and it may be based on different conditions than the ones used in this paper. The main lesson is that one can still find more surprises in the supersymmetric path integral.

3. Propagator on non-compact semisimple G

The main focus of this section is on quantum system describing a bosonic particle moving on a group manifold of a non-compact real semisimple Lie group G . Throughout this paper, we endow G by a bi-invariant pseudo-Riemannian metric, determined by² the Cartan-Killing form B on the Lie algebra \mathfrak{g} of G ,

$$(3) \quad \langle u, v \rangle_g \equiv B((L_{g^{-1}})_*u, (L_{g^{-1}})_*v) \quad \text{for all } u, v \in T_gG,$$

where L_g are the left translations on G . Note that this pseudo-Riemannian metric G induces left-invariant Riemannian metric on the homogeneous space G/K , where K is a maximal compact subgroup of G . Indeed, restriction of B to the Lie algebra \mathfrak{k} of K is negative-definite, while its restriction to the orthogonal complement \mathfrak{p} of \mathfrak{k} in \mathfrak{g} (with respect to B) is positive-definite. We denote $n = \dim G$ and $d_{\mathfrak{k}} = \dim \mathfrak{k}$, $d_{\mathfrak{p}} = \dim \mathfrak{p}$, so $n = d_{\mathfrak{p}} + d_{\mathfrak{k}}$.

Consider the Hilbert space $\mathcal{H}_B = L^2(G, dg)$, with respect to the Haar measure dg on G . The Lie algebra \mathfrak{g} acts on functions on G by the left-invariant vector fields,

$$(4) \quad (\hat{u}f)(g) = \left. \frac{d}{dt} \right|_{t=0} f(ge^{tu}) \quad \text{for all } u \in \mathfrak{g},$$

and the universal enveloping algebra $U\mathfrak{g}$ of \mathfrak{g} acts on \mathcal{H}_B by the left-invariant differential operators. The Laplace operator Δ of G is a negative of a quadratic Casimir element C_2 in $U\mathfrak{g}$. Explicitly (see Appendix A for the case $G = \text{SL}(2, \mathbb{R})$),

$$(5) \quad \Delta = - \sum_{a=1}^{d_{\mathfrak{p}}} \hat{e}_a^2 + \sum_{a=d_{\mathfrak{p}}+1}^n \hat{e}_a^2,$$

²Note that in the case of a compact Lie group, considered in [2], a bi-invariant Riemannian metric is determined by the negative of the Cartan-Killing form.

where according to the decomposition $\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}$, an orthogonal basis e_a of \mathfrak{g} satisfies $B(e_a, e_a) = 1$ for $a = 1, \dots, d_p$ and $B(e_a, e_a) = -1$ for $a = d_p + 1, \dots, n$.

The corresponding real time quantum Hamiltonian $\hat{H} = \Delta/2$ is an unbounded self-adjoint operator on \mathcal{H}_B having an absolutely continuous spectrum of infinite multiplicity filling the real line \mathbb{R} . The evolution operator $U(iT) = e^{-i\hat{H}T}$ makes a perfect sense as unitary operator with a distributional kernel, while in Euclidean time the operator $e^{-\hat{H}T}$ still exists, but is an unbounded self-adjoint operator on \mathcal{H}_B . It was shown by Krausz and Marinov [8] that the real time propagator, a distributional kernel of $U(iT)$, exhibits a similar ‘almost’ semi-classical exactness as in the case of compact Lie groups.

Here we demonstrate that the propagator of a quantum mechanical bosonic particle on a non-compact semisimple Lie group G can be understood by the localization principle, exactly as in the case of compact Lie group [2]. The setup is similar to the compact case; the main difference is that because Cartan-Killing metric is indefinite, we need to work with the real time and not with the Euclidean time as in [2]. Relatedly, we cannot naively use the fermionic Lagrangian $\frac{i}{2}\langle\psi, \dot{\psi}\rangle$ with $\psi = \psi^a T_a \in \Pi L\mathfrak{g}$ in the supersymmetric action of [2]. The reason is quite simple: canonical (graded) commutation relations $[\hat{\psi}^a, \hat{\psi}^b] = g^{ab}$ can not be represented in a positive-norm Hilbert space, because the metric on the subspace \mathfrak{k} of the real Lie algebra \mathfrak{g} is negative-definite.

However, this can be easily overcome by setting $(\hat{\psi}^a)^\dagger = \hat{\psi}^a$ for $a = 1, \dots, d_p$ and $(\hat{\psi}^a)^\dagger = -\hat{\psi}^a$ along $a = d_p + 1, \dots, n$. On the path integral side, this corresponds to a choice of integration contour for fermions such that $\psi = \psi^a T_a \in \Pi L\mathfrak{p} \oplus i\Pi L\mathfrak{k}$.

After straightening out this subtlety, we are allowed to use the supersymmetric action on G generalizing the compact case [2] which is

$$(6) \quad \mathcal{L} = \frac{1}{2}\langle J, J \rangle + \frac{i}{2}\langle \psi, \dot{\psi} \rangle,$$

where $J = g^{-1}\dot{g} \in L\mathfrak{g}$ and it has a supersymmetry

$$(7) \quad \begin{aligned} \delta g &= ig\psi, \\ \delta \psi &= -J - i\psi\psi. \end{aligned}$$

Exactly as in Proposition 2 in [2], the quantum supercharge \hat{Q} and the Hamiltonian operator \hat{H} satisfy the SUSY algebra and are given by

$$(8) \quad \begin{aligned} \hat{Q} &= \langle \hat{L}, \hat{\psi} \rangle + \frac{i}{6} \langle \hat{\psi}, [\hat{\psi}, \hat{\psi}] \rangle \\ \hat{H} &= \hat{Q}^2 = \frac{1}{2} \Delta + \frac{1}{48} f_{abc} f^{abc} \hat{I} = \frac{1}{2} \Delta + \frac{R}{12} \hat{I}, \end{aligned}$$

where $\hat{L} \equiv \hat{l}_a T^a$ (see [2]) and $R = -n/4$ is a scalar curvature of the pseudo-Riemannian metric on G .

As in [8], we are interested in the propagator between two points connected by the exponential map, the identity element 1_G of G and the element e^h , where $h \in \mathfrak{g}$. We denote $\mathfrak{h} \subset \mathfrak{g}$ to be a Cartan subalgebra containing h .

In compact Lie group case, the propagator was introduced by the formula (4.25) in [2], where both bosonic and fermionic traces were involved, which results in the integration over $\Pi T L G$, where $L G$ is the free loop space of G . However, in the non-compact case the volume of G is infinite, so we need to consider the following representation for the propagator

$$(9) \quad \begin{aligned} &e^{\frac{i\langle \rho, \rho \rangle T}{2}} \langle e^h | e^{-\frac{i\Delta T}{2}} | 1_G \rangle \\ &= \text{Tr}_{\mathcal{H}_F} \left((-1)^F c_n \hat{\chi}_1 \dots \hat{\chi}_n \langle 1_G | e^{-i\hat{H}t + i\langle h, \hat{r} \rangle} | 1_G \rangle \right) \\ &= \int_{\Pi T \Omega G} e^{iS^h} \mathcal{D}g \mathcal{D}\psi, \end{aligned}$$

where $\hat{\chi}^a$ are fermionic zero modes associated with ψ^a , normalized such that $[\hat{\chi}^a, \hat{\chi}^b] = \delta^{ab}$, i.e. $\hat{\chi}^a = \hat{\psi}^a$ for $a = 1, \dots, d_p$ and $\hat{\chi}^a = i\hat{\psi}^a$ for $a = d_p + 1, \dots, n$, ΩG is the space of based loops on G , $c_n = \pm i^{n(n-1)/2}$ is an overall phase to ensure the trace identity (see discussions in [2]), and S^h is the Minkowski action

$$(10) \quad S^h = \frac{1}{2} \int_0^T (\langle J^h, J^h \rangle + i \langle \psi, \dot{\psi} \rangle) dt, \quad J^h = J + \frac{1}{T} \text{Ad}_{g^{-1}} h.$$

Note that the integration in (9) goes over $PT\Omega G$, so the product $\chi_1 \dots \chi_n$ of fermionic zero modes drop out of the path integral. As in [2], in the derivation we used the Freudenthal-de Vries’s ‘strange formula’ $n = 24 \langle \rho, \rho \rangle$,³ where $\rho \in \mathfrak{h}_{\mathbb{C}}^*$ is a Weyl vector (a half-sum of positive roots of $\mathfrak{g}_{\mathbb{C}}$).

³We still denote the naturally induced Cartan-Killing form on the dual space $\mathfrak{h}_{\mathbb{C}}^*$ as $\langle \cdot, \cdot \rangle$.

Remark 1. The bosonic and fermionic ‘integration measures’ are

$$\mathcal{D}g = \prod_{0 \leq t \leq T} dg(t) \quad \text{and} \quad \mathcal{D}\psi = \prod_{0 \leq t \leq T} d\psi(t),$$

where

$$dg = \sqrt{|\det g(x)|} dx^1 \cdots dx^n$$

is the volume form of the pseudo-Riemannian metric on G associated with the Cartan-Killing form (see formula (A.2) in Appendix A) and

$$d\psi = \frac{1}{\sqrt{|\det g(e)|}} d\psi^1 \cdots d\psi^n, \quad \psi = \psi^a T_a \in \Pi\mathfrak{g}.$$

It should be noted that original supersymmetry δ on $\Pi T L G$, given by the formulas (7), does not preserve the domain of integration $\Pi T \Omega G$, so it cannot be used for the localization of the path integral (9). The space of based loops ΩG can be thought of as an infinite-dimensional symplectic manifold, so finding a $\Pi T \Omega G$ preserving supersymmetry transformation is equivalent to finding a Hamiltonian vector field generating the circle action. Intuitively it is not obvious that there is such transformation that preserves the boundary condition $g(0) = 1_G$. However, we found that the action S^h is invariant with respect to following deformed supersymmetry, which is compatible with this constraint,

$$(11) \quad \begin{aligned} \delta_0 g &= ig\psi, \\ \delta_0 \psi &= -J + J(0) - i\psi\psi. \end{aligned}$$

Note that δ_0 is completely ‘local’ in the sense that it treats the pair (g, ψ) as an element of $\Pi T \Omega G$. Thus we may apply the standard equivariant localization principle with the supersymmetry δ_0 and to choose the following deformation⁴

$$(12) \quad \begin{aligned} V &= -\frac{1}{2} \int_0^T \langle j^h, \dot{\psi} \rangle dt, \\ \delta_0 V &= \int_0^T \left(\frac{1}{2} \langle j^h, j^h \rangle + \frac{i}{2} \langle \dot{\psi}, (\partial_t + \text{ad}_{J^h}) \dot{\psi} \rangle \right) dt, \end{aligned}$$

⁴Note that the fixed points of (11) are closed geodesics, so one can use the Duistermaat-Heckman’s exact stationary phase approximation directly. For compact G , this approach was envisioned in [10] and was based on the earlier work [11]; later it was explicitly worked out in [12] to give the Eskin trace formula, up to the DeWitt term (see [2] for the further discussion).

which satisfying $\delta_0^2 V = 0$.

The important difference in this case is that the path integral is oscillating and the domain of integration is the based loop space on a ‘non-compact’ target. Hence we necessarily have to check the potential subtlety arising from the boundary in the paths space, which might break the invariance of the path integral for invariant deformation. To verify that it is not the case, we use the identity

$$(13) \quad \int_{\Pi T\Omega G} e^{i(S^h + \lambda\delta V)} \mathcal{D}\mu = \int_{\Pi T\Omega G} e^{iS^h} \mathcal{D}\mu + \int_0^\lambda d\tilde{\lambda} \int \delta \left(iV e^{i(S + \tilde{\lambda}\delta V)} \right) \mathcal{D}\mu,$$

where $\mathcal{D}\mu = \mathcal{D}g\mathcal{D}\psi$. It follows from the Stokes theorem that the last term is a contribution from the boundary of the space of fields. Since in our case there is no bosonic zero mode, this boundary contribution vanishes due to the highly oscillating nature of the integrand, like in the classical Riemann-Lebesgue lemma. As a result, the path integral is invariant under the transformation $S^h \rightarrow S^h + \lambda\delta V$ with arbitrary ‘real’ λ .

Therefore, as we take the limit $\lambda \rightarrow \infty$, the path integral, because of the highly oscillatory nature of the deformation, localizes onto the stationary points $\dot{j}^h = 0$. The computation is completely analogous to that in [2], and for a regular, semisimple $h \in \mathfrak{g}$ we obtain⁵

$$(14) \quad \begin{aligned} & \langle e^h | e^{-\frac{i\Delta T}{2}} | 1_G \rangle \\ &= \frac{e^{-\frac{i\langle \rho, \rho \rangle T}{2}}}{(2\pi iT)^{d_{\mathfrak{v}}/2} (-2\pi iT)^{d_{\mathfrak{k}}/2}} \sum_{\gamma \in \Gamma^h} \prod_{\alpha \in R_{\mathfrak{g}_{\mathbb{C}}, \mathfrak{h}_{\mathbb{C}}}^+} \frac{\frac{1}{2} \langle \alpha, h + \gamma \rangle}{\sinh \frac{1}{2} \langle \alpha, h + \gamma \rangle} e^{\frac{i\langle h + \gamma, h + \gamma \rangle}{2T}}. \end{aligned}$$

Here Γ^h is the characteristic lattice in the Cartan subalgebra \mathfrak{h} containing h and $R_{\mathfrak{g}_{\mathbb{C}}, \mathfrak{h}_{\mathbb{C}}}^+$ is a set of positive roots in $\mathfrak{g}_{\mathbb{C}}$ with respect to $\mathfrak{h}_{\mathbb{C}}$.⁶ This compactly reproduces results in [8] and reinforces the localization procedure described above.⁷

⁵We remark that this approach, based on using the deformed supersymmetry in $\Pi T\Omega G$, is also equally applicable to the case when G is a compact Lie group. The approach taken in [2] was to employ localization at the level of $\Pi T L G$, which uses the new localization principle.

⁶Throughout the paper, we adopt the standard conventions on \mathfrak{g} and G , used in [2, 9].

⁷Note that the overall prefactor is slightly different from [8], where [8] has $(2\pi iT)^{-n/2}$ in terms of our convention. This is because [8] used the measure $dg = \sqrt{|\det g(x)|} dx^1 \cdots dx^n$ without the absolute value of the determinant.

Remark 2. Note that prefactors of $(2\pi i)^{d_p/2}(-2\pi i)^{d_e/2}$ in the denominator of (14) are Lorentzian (Pseudo-Riemannian) infinite-dimensional analogues of the prefactor $(2\pi)^{\dim M/2}$ in a Gaussian type integral over the bosonic and fermionic modes over a finite-dimensional manifold M (cf. [2, Remark 6] for the case of Euclidean path integral). Moreover, according to $\prod_{n=1}^{\infty} z = 1/\sqrt{z}$, $-\pi < \arg z \leq \pi$ as in [2], so

$$\prod_{n=1}^{\infty} (-1) = e^{-\frac{\pi i}{2}}, \quad \prod_{n=1}^{\infty} i = e^{-\frac{\pi i}{4}} \quad \text{and} \quad \prod_{n=1}^{\infty} (-i) = e^{\frac{\pi i}{4}}.$$

4. Selberg trace formula on compact Riemann surfaces

The Selberg trace formula relates the spectrum of the Laplace operator of constant negative curvature metric on a Riemann surface with the geodesic length spectrum. By the uniformization theorem, such Riemann surface can be realized as a double coset $X = \Gamma \backslash G / K = \Gamma \backslash \mathbb{H}$, where $G = \mathrm{SL}(2, \mathbb{R})$, $K = \mathrm{SO}(2)$ and $\mathbb{H} = G/K$ in the Lobachevsky (hyperbolic) plane. Here Γ is a cofinite discrete subgroup of G , and in this paper we assume that Γ is compact, so X is a compact Riemann surface (an orbifold Riemann surface, if Γ contains elements of finite order).

Here we advocate the point of view that physically it is more advantageous to consider the gauged sigma model on the target $\Gamma \backslash G$ instead of the sigma model on the Riemann surface X . In particular, this is because $\Gamma \backslash G$ is always smooth, while $X = \Gamma \backslash G / K$ may have orbifold singularities.

Remark 3. Note that the group Γ acts on \mathbb{H} by Möbius transformations, so actually $X \simeq \bar{\Gamma} \backslash \mathbb{H}$, where $\bar{\Gamma} = \Gamma / \{\pm I\} \subset \mathrm{PSL}(2, \mathbb{R})$ is a Fuchsian group, a discrete subgroup of $\mathrm{PSL}(2, \mathbb{R})$. This can be understood from the fact that $\{\pm I\}$ is a subgroup of $K = \mathrm{SO}(2)$. Therefore, for any compact orbifold Riemann surface X we can always lift the Fuchsian group $\bar{\Gamma}$ to a discrete subgroup Γ of G with the property that $-I \in \Gamma$, which will be assumed in this section.

4.1. The gauged sigma model

We start by considering the quotient space $\Gamma \backslash G$, which plays a fundamental role in the representation theory. It is quite natural to ask whether there is a supersymmetric system similar to that in the case of G . The answer is remarkably simple, because under the purely bosonic projection $\pi_{\Gamma} : G \rightarrow \Gamma \backslash G$

the Lagrangian (6) remains supersymmetric with the same supersymmetry (7). After the projection, the Hilbert space becomes

$$\mathcal{H}_{\Gamma \backslash G} = L^2(\Gamma \backslash G, dg) \otimes \mathcal{H}_{F, \mathfrak{g}}.$$

Physical realization of the right quotient by $K = \text{SO}(2)$ is achieved by gauging the global symmetry $g \rightarrow gk$, $\psi \rightarrow \text{Ad}_{k^{-1}}\psi$, where $k \in K$, by a connection $A = A^3 T_3$ in the principal K -bundle over S_T^1 . The simplest such extension that preserves the supersymmetry is given by the gauged sigma model on $\Gamma \backslash G$ with the action

$$(15) \quad S_A[g, \psi] = \int_0^T \mathcal{L}_0 dt.$$

Here

$$(16) \quad \mathcal{L}_0 = \frac{1}{2} \langle J_A, J_A \rangle + \frac{i}{2} \langle \psi, \partial_t^A \psi \rangle,$$

where

$$J_A = J - A \quad \text{and} \quad \partial_t^A = \partial_t + \text{ad}_A$$

are gauge covariant objects. The supersymmetry (7) also gets modified

$$(17) \quad \begin{aligned} \delta g &= ig\psi, \\ \delta \psi &= -J_A - i\psi\psi, \\ \delta A &= 0. \end{aligned}$$

However, the physical spectrum of the quantum system with the Lagrangian \mathcal{L}_0 no longer corresponds to the eigenstates of a purely bosonic Laplacian on a Riemann surface X since the fermionic part is being gauged as well. To see this, we first perform canonical quantization on $\Gamma \backslash G$ and impose the gauge constraint later. In the first step, we can effectively set $A = 0$ and hence the quantum supercharge \hat{Q} and the Hamiltonian \hat{H} are given by the same formulas (8).

Now to impose a gauge invariance, we observe that it follows from (16) that Lagrangian \mathcal{L}_0 has no kinetic term for A , so we have classical Gauss law

$$(18) \quad C_0 : J_A^3 + 2i\psi^1\psi^2 = 0.$$

Quantum mechanically it is realized as the following constraint on the Hilbert space⁸

$$(19) \quad \hat{C}_0 : \{ \hat{T}_3 - 2\hat{N}_{F,\mathfrak{p}} + \hat{I} = 0 \},$$

where $\mathfrak{p} = \mathfrak{g}/\mathfrak{k}$ and $\hat{N}_{F,\mathfrak{p}}$ is the fermion number operator in the subspace $\mathcal{H}_{F,\mathfrak{p}}$ of the fermion Hilbert space $\mathcal{H}_{F,\mathfrak{g}}$.

Indeed, $\mathcal{H}_{F,\mathfrak{p}} = \mathbb{C}^2$ and fermion operators $\hat{\psi}^1$ and $\hat{\psi}^2$ are normalized as follows (see [2] and Appendix A)

$$(20) \quad [\hat{\psi}^1, \hat{\psi}^1] = g^{11} = \frac{1}{8} \quad \text{and} \quad [\hat{\psi}^2, \hat{\psi}^2] = g^{22} = \frac{1}{8},$$

so in the standard basis of \mathbb{C}^2 where $\hat{N}_{F,\mathfrak{p}} = -\frac{1}{2}(\sigma_3 - I)$ is diagonal, we have

$$(21) \quad \hat{\psi}^1 = \frac{1}{4}\sigma_1, \quad \hat{\psi}^2 = \frac{1}{4}\sigma_2 \quad \text{and} \quad 2i\hat{\psi}^1\hat{\psi}^2 = -\frac{1}{8}\sigma_3.$$

The current J_A^3 is quantized as in [2], since $g_{33} = \langle T_3, T_3 \rangle = -8$ we have

$$\hat{J}_A^3 = -\frac{1}{8}(\hat{J}_A)_3 = \frac{i}{8}\hat{T}_3,$$

which gives equation (19).

Thus the physical Hilbert space \mathcal{H}_0 consists of the null space of \hat{C}_0 in the unconstrained Hilbert space $\mathcal{H}_{\Gamma \backslash G}$. Representing element in $L^2(\Gamma \backslash G) \otimes \mathcal{H}_{F,\mathfrak{p}} = L^2(\Gamma \backslash G) \otimes \mathbb{C}^2$ as a two-component column vector $f = \begin{pmatrix} f_+ \\ f_- \end{pmatrix}$ and using that $\hat{T}_3 = \frac{\partial}{\partial \theta}$, we rewrite equation $\hat{C}_0 f = 0$ as

$$\frac{\partial f}{\partial \theta} = i\sigma_3 f,$$

so $f_{\pm}(ge^{i\theta T_3}) = e^{\pm i\theta} f_{\pm}(g)$. This gives a remarkable identification

$$(22) \quad \mathcal{H}_0 = (\mathcal{H}_{\Gamma \backslash G})/\hat{C}_0 = L^2_{1/2}(X) \oplus L^2_{-1/2}(X) \oplus \mathcal{H}_{F,\mathfrak{k}}.$$

Here $L^2_m(X)$ for $m \in \frac{1}{2}\mathbb{Z}$ is the Hilbert space of functions on \mathbb{H} satisfying

$$(23) \quad f \left(\frac{az + b}{cz + d} \right) \frac{|cz + d|^{2m}}{(cz + d)^{2m}} = f(z), \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma,$$

⁸Note that we made a specific choice of a quantization scheme when we promote the classical Gauss law (18) to the quantum Gauss law (19).

and square integrable on a fundamental domain of Γ in \mathbb{H} with respect to the hyperbolic area form $y^{-2}dx dy$ (see Appendix A), and $\mathcal{H}_{F,\mathfrak{k}}$, is the fermionic Hilbert space along \mathfrak{k} .⁹ Restriction of Δ to the subspaces $L^2_{\pm 1/2}(X)$ coincides with the Maass Laplacians of weight $\pm 1/2$ (see Appendix A). In other words, the quantum system corresponding to the Lagrangian \mathcal{L}_0 describes the Maass Laplacians $D_{\pm \frac{1}{2}}$ of weights $\pm 1/2$ on X , together with a single free Majorana fermion along \mathfrak{k} direction.

However, the quantum system associated with the Lagrangian \mathcal{L}_0 is actually anomalous! Namely, since $-I \in \Gamma$, we see that for $m = \pm 1/2$ putting $\gamma = -I$ into (23) gives $f(z) = -f(z)$, so $f(z) = 0$ and instead of $L^2_{\pm 1/2}(X)$ we get a zero dimensional space $\{0\}$. In other words, the physical Hilbert space is empty because one can not obey the Gauss law (19).

4.2. Wilson loops and defects

The inconsistency of the Gauss law for a system with fermions coupled to dynamical gauge fields is a typical manifestation of an anomaly in the Hamiltonian approach [13–17]. In our case, it can be understood as a ‘t Hooft anomaly of \mathbb{Z}_2 one-form global symmetry in quantum mechanics.

We recall that a gauge theory in d space-time dimensions with a gauge group G with a center $\mathcal{Z}(G)$ has an (electric) one-form global symmetry $\mathcal{Z}^{(1)} \subset \mathcal{Z}(G)$ whenever there are no matter fields charged under $\mathcal{Z}^{(1)}$, and hence the theory can be consistently put on a $G/\mathcal{Z}^{(1)}$ -bundle [18]. On a Wilson loop in a representation charged under $\mathcal{Z}^{(1)}$, this one-form symmetry acts by a non-trivial phase.

In our case the gauge group is $K = \text{SO}(2)$ and the matter fields $g \in \Gamma \backslash G$ have charge 2, while $\psi \in \mathfrak{g}$ decomposes into charge 2 and 0 under $\text{SO}(2)$, so the one-form symmetry is $\mathcal{Z}^{(1)} = \mathbb{Z}_2$. This \mathbb{Z}_2 symmetry acts on a Wilson line with the charge n representation of $\text{SO}(2)$ by a phase $(-1)^n$.

Gauging the \mathbb{Z}_2 one-form symmetry changes the structure group to $\text{PSO}(2) = \text{SO}(2)/\mathbb{Z}_2$. In our case the local form of the Lagrangian \mathcal{L}_0 and the Gauss law is insensitive to the global structure of the gauge bundle, so there is no gauge-invariant physical state. Moreover, there is no local gauge-invariant counter-terms that makes theory non-anomalous, so the system on the $\text{PSO}(2)$ -bundle suffers from a gauge anomaly and is physically inconsistent.

⁹Although $d_{\mathfrak{k}} = 1$, we can innocuously extend the definition of the fermionic Hilbert space as discussed in [2].

On the other hand, for our original system with the Lagrangian \mathcal{L}_0 and the gauge group $K = \text{SO}(2)$, the inconsistency of the Gauss law is related to the ‘t Hooft anomaly, so we can make the system non-anomalous by adding a gauge-invariant counterterm. It has the form nA^3 , where n is an odd integer. This defines non-anomalous Lagrangians

$$(24) \quad \mathcal{L}_n = \mathcal{L}_0 + nA^3, \quad \text{where } n = 2k + 1, \quad k \in \mathbb{Z}.$$

Note that \mathcal{L}_n is still invariant under the supersymmetry (17). This counterterm can be also viewed as a one-dimensional version of the Chern-Simons term of level n . Equivalently, it can be thought of as addition to the action of a temporal Wilson loop of charge n ,

$$\mathcal{W}_n(A) = e^{in \int_0^T A^3 dt}.$$

Since $\mathcal{W}_n(A)$ with an odd integer n has charge 1 under the \mathbb{Z}_2 one-form symmetry, we see that the path integral with \mathcal{L}_0 suffers from the ‘t Hooft anomaly for the \mathbb{Z}_2 one-form symmetry which makes the total system \mathcal{L}_n non-anomalous. The insertion of such defect twists the Hilbert space as

$$(25) \quad \mathcal{H}_n = L^2_{\frac{n+1}{2}}(X) \oplus L^2_{\frac{n-1}{2}}(X) \oplus \mathcal{H}_{F,\mathfrak{k}},$$

which clearly demonstrates that for n being an odd integer the physical Hilbert space \mathcal{H}_n is well-defined and the system is non-anomalous.

We note that for each n the system has a single fermionic zero mode

$$\psi^3 = \frac{1}{T} \int_0^T \psi^3(t) dt,$$

which we will abbreviate as $\psi_{\mathfrak{k}}$, and denote by $\hat{\psi}_{\mathfrak{k}}$ properly normalized fermion along \mathfrak{k} direction, so that $[\hat{\psi}_{\mathfrak{k}}, \hat{\psi}_{\mathfrak{k}}] = 1$ and $\hat{\psi}_{\mathfrak{k}}^\dagger = \hat{\psi}_{\mathfrak{k}}$. Therefore, the natural non-supersymmetric observable satisfying extended localization principle discussed in Section 2 is given by

$$(26) \quad \begin{aligned} I_n &\equiv \text{Str}_{\mathcal{H}_n} [\hat{\psi}_{\mathfrak{k}} e^{-i\hat{H}_0 T}] = \text{Tr}_{L^2_{\frac{n-1}{2}}(X)} [e^{-iT\Delta/2}] - \text{Tr}_{L^2_{\frac{n+1}{2}}(X)} [e^{-iT\Delta/2}] \\ &= \frac{e^{-\frac{i(\rho,\rho)T}{2}}}{\text{vol}(\mathcal{G})} \int \mathcal{W}_n(A) \psi_{\mathfrak{k}} e^{iS_A[g,\psi]} \mathcal{D}g \mathcal{D}\psi \mathcal{D}A, \end{aligned}$$

where in the last line, we factor out the volume of the gauge group \mathcal{G} , as it is customary in the gauge theory.

However, for the purposes of the basic Selberg trace formula, one needs to consider the trace of the operator $e^{-iT\Delta/2}$ in the Hilbert space $L_0^2(X)$ of square integrable automorphic forms of weight 0. To achieve this goal, we start with the following extremely naive formula, based on the ‘illegal’ use of telescoping¹⁰ and geometric series

$$(27) \quad \begin{aligned} \text{Tr}_{L_0^2(X)}[e^{-iT\Delta/2}] &\stackrel{?}{=} \sum_{n \in 2\mathbb{Z}+1} \frac{1}{2} \text{sgn}(n) I_n \\ &\stackrel{?}{=} \frac{e^{-\frac{i\langle \rho, \rho \rangle T}{2}}}{\text{vol}(\mathcal{G})} \int \frac{1}{\mathcal{W}_{-1}(A) - \mathcal{W}_1(A)} \psi_{\mathfrak{k}} e^{iS_A[g, \psi]} \mathcal{D}g \mathcal{D}\psi \mathcal{D}A. \end{aligned}$$

Note that each of equalities in (27) does not really make sense since the series in the first line is not convergent! More serious mathematical objection is that each term I_n doesn’t contain any information about the non-zero modes since the spectra of operators $D_{\frac{n+1}{2}}$ and $D_{\frac{n-1}{2}}$ coincide, except for possible finitely many eigenvalues related to zero modes of Hodge Laplace operators (see Appendix A). Physically, this is because $\psi_{\mathfrak{k}}$ is just a free fermion, and hence I_n can be thought of as a standard Witten index for a (twisted) supersymmetric non-linear sigma model on X .

However, it is truly amazing that the first and the last term in (27) are equal, and the derivation does not use I_n at all! To demonstrate this, let us start with the last line in (27), where the domain of integration is

$$L(\Gamma \backslash G) \times \text{III}L\mathfrak{g} \times \mathcal{A}.$$

Here $L(\Gamma \backslash G)$ and $\text{III}L\mathfrak{g}$ are free loop spaces of $\Gamma \backslash G$ and \mathfrak{g} (the latter with the reversed parity), and \mathcal{A} is the space of connections on the principle $\text{SO}(2)$ -bundle over S_T^1 .

We fix the gauge in the last line of (27) by setting $A = \mathbf{h}/T$ with $\mathbf{h} = hT_3$, which for the last term in (27) gives the following expression

$$(28) \quad e^{-\frac{i\langle \rho, \rho \rangle T}{2}} \int_{-\pi}^{\pi} \frac{dh}{2\pi} \frac{i}{2 \sin h} \int \psi_{\mathfrak{k}} e^{iS_f[\mathbf{h}](g, \psi)} \mathcal{D}g \mathcal{D}\psi,$$

where

$$(29) \quad S_f[\mathbf{h}](g, \psi) = \frac{1}{2} \int_0^T \left(\langle J_{\mathbf{h}/T}, J_{\mathbf{h}/T} \rangle + i \langle \psi, \partial_t^{\mathbf{h}/T} \psi \rangle \right) dt$$

¹⁰The formula $\sum_{n=1}^{\infty} (a_n - a_{n+1}) = a_1$, which is valid only when $\lim_{n \rightarrow \infty} a_n = 0$.

is the fixed gauge action. Here the 2π in the denominator of (28) comes from the volume of $SO(2)$. Now the effective path integral for fixed \mathbf{h} is nothing but a twisted partition function for the supersymmetric particle on $\Gamma \backslash G$. Since in (28)–(29) bosonic and fermionic degree of freedoms are decoupled, we have

$$\begin{aligned}
 (30) \quad & e^{-\frac{i\langle \rho, \rho \rangle T}{2}} \int \psi_{\mathfrak{k}} e^{iS_f[\mathbf{h}](g, \psi)} \mathcal{D}g \mathcal{D}\psi \\
 &= \int \psi_{\mathfrak{k}} e^{-\frac{1}{2} \int_0^T \langle \psi, \partial_t^{\mathbf{h}/T} \psi \rangle dt} \mathcal{D}\psi \int_{\Gamma \backslash G} \langle g | e^{-iT\Delta_{\Gamma \backslash G}/2} | g e^{i\mathbf{h}} \rangle dg.
 \end{aligned}$$

In the Hamiltonian representation, the fermion path integral becomes the trace

$$(31) \quad \int \psi_{\mathfrak{k}} e^{-\frac{1}{2} \int_0^T \langle \psi, \partial_t^{\mathbf{h}/T} \psi \rangle dt} \mathcal{D}\psi = \text{Tr}_{\mathcal{H}_{F, \mathfrak{g}}} \left[(-1)^F \hat{\psi}_{\mathfrak{k}} e^{-i\hat{H}_F T} \right],$$

where

$$(32) \quad \hat{H}_F = -\frac{i}{2T} \langle \hat{\psi}, \text{ad}_{\mathbf{h}} \hat{\psi} \rangle = \frac{16ih}{T} \hat{\psi}^1 \hat{\psi}^2$$

and fermion operators $\hat{\psi}^1$ and $\hat{\psi}^2$ are normalized as in (20). Though $d_{\mathfrak{k}} = 1$ and Majorana fermions are anomalous, we can use the approach in [2], which effectively reduces the fermion Hilbert space $\mathcal{H}_{F, \mathfrak{g}}$ to \mathbb{C}^2 , and the operator $(-1)^F \hat{\psi}_{\mathfrak{k}}$ to σ_3 . Then using (21) we obtain

$$(33) \quad -i\hat{H}_F T = -ih\sigma_3 = -2ih \left(N_{F, \mathfrak{p}} - \frac{1}{2} \right),$$

so

$$(34) \quad \text{Tr}_{\mathcal{H}_{F, \mathfrak{g}}} \left[(-1)^F \hat{\psi}_{\mathfrak{k}} e^{-i\hat{H}_F T} \right] = -2i \sin h.$$

Thus we see that the fermionic trace (31) remarkably cancels the factor in (28), arising from the insertion of $1/(\mathcal{W}_{-1}(A) - \mathcal{W}_1(A))$ into the path integral (27).

The final ingredient of the proof is the following basic representation identity

$$(35) \quad \text{Tr}_{L_0^2(X)} \left[e^{-iT\Delta/2} \right] = \int_K \frac{dk}{\text{vol}(K)} \int_{\Gamma \backslash G} \langle g | e^{-iT\Delta_{\Gamma \backslash G}/2} | gk \rangle dg,$$

which is valid for any bi-invariant measure dg on G , and we obtain

$$(36) \quad \begin{aligned} Z(iT) &= \text{Tr}_{L_0^2(X)} \left[e^{-iT\Delta/2} \right] \\ &= \frac{e^{-\frac{i(\rho, \rho)T}{2}}}{\text{vol}(\mathcal{G})} \int \frac{1}{\mathcal{W}_{-1}(A) - \mathcal{W}_1(A)} \psi_{\mathfrak{k}} e^{iS_A[g, \psi]} \mathcal{D}g \mathcal{D}\psi \mathcal{D}A. \end{aligned}$$

Thus we have proved that insertion of a special singular defect into the path integral provides the information about the full spectrum, in contrast with the naive path integral without defects.

Note that formula (36) should be understood in the sense of generalized functions, since $Z(iT)$ is not a convergent series for real T . Instead, we use the boundness of the spectrum of Δ on a Riemann surface X , which makes $Z(iT)$ analytic on the lower half-plane $\text{Im } T < 0$. This allows us to consider the equality (36) as a boundary limit of analytic functions at the lower half-plane. Therefore, once we evaluate $Z(iT)$ using the localization, it is necessary to analytically continue to the imaginary time in order to obtain the canonical partition function.

4.3. Localization and the pre-trace formula

Now we are ready to perform the supersymmetric localization. Since supersymmetry (17) doesn't act on the gauge field and commutes with the gauge symmetry, we can first fix the gauge and then apply the supersymmetric localization to the path integral (28). The bosonic domain of integration in the path integral is a free loop space $L(\Gamma \backslash G)$, so we need to describe its connected components.

For this aim, we use the basic fact that for any topological space \mathcal{X} with a based point x_0 , the based loop space $\Omega_{x_0} \mathcal{X}$ decomposes into the disjoint union of connected components, labeled by the elements of the fundamental group $\pi_1(\mathcal{X}, x_0)$,

$$(37) \quad \Omega_{x_0} \mathcal{X} = \bigsqcup_{\xi \in \pi_1(\mathcal{X}, x_0)} \Omega_{x_0, \xi} \mathcal{X},$$

where $\Omega_{\xi, x_0} \mathcal{X}$ is the set of all based loops in the same homotopy class ξ . In our case $\mathcal{X} = \Gamma \backslash G$ and since $p : G \rightarrow \Gamma \backslash G$ is a covering with fibers Γ , we

have an exact sequence of homotopy groups

$$0 \rightarrow \pi_1(G, e) \rightarrow \pi_1(\Gamma \backslash G, \bar{e}) \rightarrow \pi_0(\Gamma) \rightarrow 0,$$

where e is the unit in G and \bar{e} — its projection to $\Gamma \backslash G$. The covering $p : G \rightarrow \Gamma \backslash G$ is normal, so the image N of $\pi_1(G, e)$ is a normal subgroup of $\pi_1(\Gamma \backslash G, \bar{e})$ and

$$\Gamma \simeq \pi_1(\Gamma \backslash G, \bar{e})/N.$$

From here it is easy to derive that

$$(38) \quad L(\Gamma \backslash G) = \bigsqcup_{g_0 \in F} \bigsqcup_{[\gamma]} \left(\bigcup_{\sigma \in \Gamma_\gamma \backslash \Gamma} \mathcal{P}_{\sigma^{-1}\gamma\sigma}(G, g_0) \right),$$

where F is a connected fundamental domain of Γ in G , $[\gamma]$ runs over all conjugacy classes in Γ with representatives γ , the group Γ_γ is a centralizer of γ in Γ , and $\mathcal{P}_\gamma(G, g_0)$ is a space of paths $g(t)$ in G satisfying $g(0) = g_0$ and $g(T) = \gamma g_0$.

Decomposition (38) can be also obtained directly using the basic fact that for any space \mathcal{X} we trivially have

$$L\mathcal{X} = \bigsqcup_{x_0 \in X} \Omega_{x_0} \mathcal{X}.$$

Specifically, in our case $\mathcal{X} = \Gamma \backslash G$, and for every $x_0 \in \mathcal{X}$ we choose $g_0 \in G$ such that $p(g_0) = x_0$. Since $\Gamma = p^{-1}(x_0)$, every other lift of x_0 is σg_0 , where $\sigma \in \Gamma$. We can always assume that $g_0 \in F$ for some choice of a connected fundamental domain of Γ in G . By a lifting path property, each $\xi(t) \in \Omega_{x_0} \mathcal{X}$ can be uniquely lifted to a path $g(t) \in \mathcal{P}_\gamma(G, g_0)$ for some $\gamma \in \Gamma$. By a homotopy lifting property, loops $\xi_1, \xi_2 \in \Omega_{x_0} \mathcal{X}$ are related by a free homotopy with x_0 moving along $\xi(t) \in \Omega_{x_0} \mathcal{X}$, if $\xi_1(t)$ lifts to $g_1(t) \in \mathcal{P}_{\gamma_1}(G, g_0)$, $\xi(t)$ lifts to $g(t) \in \mathcal{P}_\sigma(G, g_0)$ and $\xi_2(t)$ necessarily lifts to $g_2(t) \in \mathcal{P}_{\gamma_2}(G, \sigma g_0)$ for some $\gamma_1, \gamma_2, \sigma \in \Gamma$. Since Γ is discrete, using $\mathcal{P}_{\sigma\gamma\sigma^{-1}}(G, g_0) = \sigma^{-1}\mathcal{P}_\gamma(G, \sigma g_0)$ we get $\sigma^{-1}\gamma_2\sigma = \gamma_1$, and the decomposition (38) follows.

As a result, $Z(iT)$ takes the following form

$$(39) \quad Z(iT) = \sum_{[\gamma]} Z_{[\gamma]}(iT),$$

where

$$Z_{[\gamma]}(iT) = e^{-\frac{i(\rho,\rho)T}{2}} \int_{-\pi}^{\pi} \frac{dh}{2\pi} \frac{i}{2\sin h} \\ \times \sum_{\sigma \in \Gamma_{\gamma} \backslash \Gamma} \int_F dg_0 \int_{\mathcal{P}_{\sigma^{-1}\gamma\sigma}(G, g_0) \times \Pi L\mathfrak{g}} \psi_{\mathfrak{k}} e^{iS_f[\mathbf{h}](g, \psi)} \mathcal{D}g \mathcal{D}\psi$$

and $[\gamma]$ runs over all conjugacy classes in Γ with representatives γ .

The main advantage of (39) is that the supersymmetry is preserved for each $[\gamma]$, so we can apply the supersymmetric localization for each $Z_{[\gamma]}(iT)$ separately. Though the path integral (28) is not supersymmetric, the extended localization principle, formulated in Section 2 is applicable. As in Section 3, for each $[\gamma]$ we use purely oscillating deformation $V_{[\gamma]}$; the vanishing of contribution from the boundary of the space of fields is supported by the compactness of $\Gamma \backslash G$ and by the arguments in Section 3.

Thus we arrive at the simplified expression

$$(40) \quad Z_{[\gamma]}(iT) = e^{-\frac{i(\rho,\rho)T}{2}} \int_{-\pi}^{\pi} \frac{dh}{2\pi} \frac{i}{2\sin h} \sum_{\sigma \in \Gamma_{\gamma} \backslash \Gamma} \int_F \mathcal{A}_{\sigma^{-1}\gamma\sigma}(g_0; \lambda) dg_0,$$

where for arbitrary $\gamma \in \Gamma$,

$$(41) \quad \mathcal{A}_{\gamma}(g_0; \lambda) = \int_{\mathcal{P}_{\gamma}(G; g_0) \times \Pi L\mathfrak{g}} \psi_{\mathfrak{k}} e^{i(S_f[\mathbf{h}](g, \psi) + \lambda \delta V_{\gamma})} \mathcal{D}g \mathcal{D}\psi.$$

Remark 4. We observe that if deformation V_{γ} is such that $V_{-\gamma} = V_{\gamma}$, then

$$Z_{[\gamma]}(iT) = Z_{[-\gamma]}(iT),$$

which is intuitively clear because $-I$ is also an element of K , therefore γ and $-\gamma$ act identically on the homogeneous space G/K , which we have physically realized as a gauged sigma model. Using (40), this can be easily proved as follows. We have

$$\mathcal{P}_{\gamma}(G, g_0) = \mathcal{P}_{-\gamma}(G, g_0) \cdot k,$$

where the path $k(t) = \exp(\pi t T_3/T)$ connects $-I$ and I in G . Changing variables $g(t) \rightarrow g(t)k(t)$ in the path integral (41) shifts $J \mapsto J + \pi T_3/T$, which can be compensated by the shift $h \rightarrow h + \pi$, which is allowed because the integrand is 2π -periodic function of h .

Remark 5. Assuming that the deformation V_I is invariant under the shift $g(t)g$ with constant $g \in G$, we have from (41),

$$\mathcal{A}_I(g_0; \lambda) = \mathcal{A}_I(e, \lambda).$$

Let G_γ be the centralizer of γ in G . ‘Unfolding’ the fundamental domain $F = \Gamma \backslash G$, we get

$$\begin{aligned} \sum_{\sigma \in \Gamma_\gamma \backslash \Gamma} \int_F \mathcal{A}_{\sigma^{-1}\gamma\sigma}(g_0; \lambda) dg_0 &= \int_{\Gamma_\gamma \backslash G} \mathcal{A}_\gamma(g_0; \lambda) dg_0 \\ (42) \qquad \qquad \qquad &= \int_{G_\gamma \backslash G} dg_0 \int_{\Gamma_\gamma \backslash G_\gamma} \mathcal{A}_\gamma(g_1 g_0; \lambda) dg_1 \\ &= \text{vol}(\Gamma_\gamma \backslash G_\gamma) \int_{G_\gamma \backslash G} \mathcal{A}_\gamma(g_0; \lambda) dg_0. \end{aligned}$$

As a result, we obtain the localized version of the pre-trace formula

$$\begin{aligned} Z(iT) &= \sum_{[\gamma]} \text{vol}(\Gamma_\gamma \backslash G_\gamma) \int_{G_\gamma \backslash G} dg_0 e^{-\frac{i(\rho, \rho)T}{2}} \int_{-\pi}^\pi \frac{dh}{2\pi} \frac{i}{2 \sin h} \mathcal{A}_\gamma(g_0; \lambda) \\ (43) \qquad &= \sum_{[\gamma]} \text{vol}(\Gamma_\gamma \backslash G_\gamma) e^{-\frac{i(\rho, \rho)T}{2}} \int_{-\pi}^\pi \frac{dh}{2\pi} \frac{i}{2 \sin h} \int_{G_\gamma \backslash G} \mathcal{A}_\gamma(g_0; \lambda) dg_0, \end{aligned}$$

where summation goes over conjugacy classes $[\gamma]$ of the discrete subgroup Γ of G , and in the 2nd line we interchanged the order of the integrations. According to Remarks 4 and 5, contributions from the conjugacy classes $[\gamma]$ and $[-\gamma]$ to (43) are the same and for $\gamma = I$ the integral over $G_\gamma \backslash G$ is understood as $\mathcal{A}_I(e; \lambda)$. Also note that, as it follows from Remark 1, the volume form dg_0 is given by formula (A.2) in Appendix A.

We can regard (43) as a path integral version of the pre-trace formula with an invariant deformation from the localization parametrized by λ , where $\lambda = 0$ corresponds to the original pre-trace formula. The main role of the localization is that it is separately applied to each orbital integral over $G_\gamma \backslash G$, labeled by the conjugacy class $[\gamma]$, and upon taking the limit $\lambda \rightarrow \infty$ leads to the desired Selberg trace formula.

4.4. The Selberg trace formula

To make a precise connection with the Selberg trace formula, we group the conjugacy classes in Γ according to the conjugacy classes in $\bar{\Gamma} = \Gamma / \{\pm I\} \subset \text{PSL}(2, \mathbb{R})$. Since $\bar{\Gamma}$ is cocompact, the elements of $\bar{\Gamma}$ except the identity are either hyperbolic or elliptic. We have the following contributions to the pre-trace formula (43).

• *Identity element*

For the conjugacy class $[I] = I$ in Γ we use the following invariant deformation

$$(44) \quad \begin{aligned} V_0 &= -\frac{1}{2} \int_0^T \langle \partial_t^{A_0} J, \partial_t^{A_0} \psi \rangle dt, \\ \delta V_0 &= \frac{1}{2} \int_0^T (\langle \partial_t^{A_0} J, \partial_t^{A_0} J \rangle + i \langle \partial_t^{A_0} \psi, \partial_t^J \partial_t^{A_0} \psi \rangle) dt. \end{aligned}$$

Here A_0 is some arbitrary regular element in \mathfrak{k} , introduced in order to have isolated critical points, say $A_0 = 2\pi T_3/T$. Then from the formula for δV_0 , we see that the path integral \mathcal{A}_I localizes onto a set of critical points

$$(45) \quad g_n(t) = g(0)e^{\frac{2\pi i \nu_n t}{T}}, \quad \text{where } \nu_n = inT_3, \quad n \in \mathbb{Z} \quad \text{and} \quad g(0) \in F.$$

Indeed, equation $\partial_t^{A_0} J = 0$ gives $J(t) = e^{-A_0 t} J_0 e^{A_0 t}$, and since $J \in L\mathfrak{g}$ we obtain $J_0 = cA_0$; then $g(t) = g(0)e^{cA_0 t} \in LG$ implies (for our choice of A_0) that c is an integer.

Around each critical point $g = g_n$ the leading quadratic term in δV_0 has additional fermionic zero modes associated with the kernel of the operator $D = \partial_t^{J_n}$, where $J_n = g_n^{-1} \dot{g}_n = -2\pi n T_3/T$, which are given by

$$(46) \quad \text{Re}(E_\alpha e^{\frac{-2\pi i(\alpha, \nu_n)t}{T}}), \quad \text{Im}(E_\alpha e^{\frac{-2\pi i(\alpha, \nu_n)t}{T}}) \quad \alpha \in R^\mathfrak{k}.$$

To compute the contribution of these critical points, we use the method developed in [2] (and refer to [2, Sect. 5] for more details). Namely, denoting the fermion zero modes by χ and using the decomposition $\psi = \chi + \eta$, we see that

$$(47) \quad \lim_{\lambda \rightarrow \infty} \mathcal{A}(e, \lambda) = \sum_{n \in \mathbb{Z}} e^{iS_n} \int e^{iS_n^{\text{loc}}} \mathcal{D}Y \mathcal{D}\eta d\chi,$$

where $S_n = S_f[\mathbf{h}](g_n(t), 0)$ is the classical contribution to the gauge fixed action (29), $Y = g^{-1} \delta g$ and

$$(48) \quad \begin{aligned} S_n^{\text{loc}} &= \frac{1}{2} \int_0^T \left(\langle D\partial_t^{A_0} Y, D\partial_t^{A_0} Y \rangle + i \langle \chi, \partial_t^A \chi \rangle + i \langle \partial_t^{A_0} \eta, D\partial_t^{A_0} \eta \rangle \right. \\ &\quad \left. + i \langle \partial_t^{A_0} \eta, [DY, \partial_t^{A_0} \chi] \rangle + \frac{i}{2} \langle \partial_t^{A_0} \chi, [[DY, Y], \partial_t^{A_0} \chi] \rangle \right) dt. \end{aligned}$$

Since $[Y, \partial_t^{A_0} \chi]$ is orthogonal to χ , the change of variables

$$(49) \quad \partial_t^{A_0} \eta \rightarrow \partial_t^{A_0} \eta - [Y, \partial_t^{A_0} \chi]$$

is legitimate, and we obtain

$$(50) \quad S_n^{\text{loc}} = \frac{1}{2} \int_0^T \left(\langle D\partial_t^{A_0} Y, D\partial_t^{A_0} Y \rangle + i\langle \chi, \partial_t^A \chi \rangle + i\langle \partial_t^{A_0} \eta, D\partial_t^{A_0} \eta \rangle \right) dt.$$

Evaluating Gaussian path integrals, we finally get

$$(51) \quad \begin{aligned} \mathcal{A}_I(g_0; \lambda) &= \frac{1}{(2\pi i)(-2\pi i)^{1/2}} \frac{\text{Pf}(-D(\partial_t^{A_0})^2)}{\sqrt{\det D^2(\partial_t^{A_0})^2}} \int e^{-\frac{1}{2} \int_0^T \langle \chi, \partial_t^A \chi \rangle dt} d\chi \\ &= \frac{1}{(2\pi i)(-2\pi i)^{1/2}} \cdot \frac{1}{iT \cdot T^{1/2}} \cdot 2(h + 2\pi n) \\ &= \frac{2(h + 2\pi n)}{(2\pi iT)^{3/2}}, \end{aligned}$$

where we used Remark 2 and section 4 of [2] for the computations of each factors.

Now we use the important relation, valid whenever $-I \in \Gamma$, which holds for any Riemannian metric on G and corresponding induced Riemannian volumes on $\Gamma \backslash G$ and K ,

$$(52) \quad \text{vol}(\Gamma \backslash G) / \text{vol}(K) = \text{vol}(X) / 2.$$

Indeed, $X = \Gamma \backslash G / K$ and $\Gamma \cap K = \{\pm I\}$, so $\bar{K} = K / \{\pm I\}$ acts freely on $\Gamma \backslash G$. We emphasize that in our case the Riemannian volume form is induced from Cartan-Killing metric $ds^2 = 4 \text{Tr}(g^{-1} dg g^{-1} dg)$ (see Appendix A), so we obtain

$$(53) \quad Z_I(iT) = \frac{\text{vol}(X) e^{-\frac{iT}{16}}}{2(\pi iT)^{3/2}} \int_{-\pi}^{\pi} \sum_{n \in \mathbb{Z}} \frac{i(h + 2\pi n)}{\sin h} e^{-\frac{4i(h+2\pi n)^2}{T}} dh.$$

Remark 6. We emphasize that the infinite series in (53) is well defined, since the apparent poles at $h = m\pi$ cancel each other once we group together terms with n and $-n - m$. This makes sense, since according to the formula (14), the series (53) is a propagator on $\text{SL}(2, \mathbb{R})$.

However, in order to Wick rotate expression (53) to the Euclidean time we need to incorporate the integration over h . Naively changing the order

of integration and summation, we arrive at a remarkably simple formula

$$(54) \quad Z_I(iT) \stackrel{\text{naive}}{=} \frac{\text{vol}(X)e^{-\frac{iT}{16}}}{2(\pi iT)^{3/2}} \int_{\mathbb{R}} \frac{ih}{\sin h} e^{-\frac{4ih^2}{T}} dh.$$

However, this representation is ambiguous because of the poles at $h = n\pi$ with non-zero integer n , and we need to specify a precise choice of the contour of integration that bypasses these poles.

We claim that there is a unique choice of such contour which gives a physically meaningful thermal partition function after the Wick rotation to the Euclidean time. It resembles the Feynman contour, and we write

$$(55) \quad Z_I(iT) \stackrel{\text{physical}}{=} \frac{\text{vol}(X)e^{-\frac{iT}{16}}}{2(\pi iT)^{3/2}} \int_{\mathbb{R}(1-i\varepsilon)} \frac{ih}{\sin h} e^{-\frac{4ih^2}{T}} dh.$$

It is quite remarkable that after the Wick rotation with $T = -i\beta$ for this choice of the contour of integration, twice the contribution of $\gamma = I$ is equal to the contribution of the identity element to the Selberg trace formula,

$$(56) \quad 2Z_I(\beta) = \frac{\text{vol}(X)e^{-\frac{\beta}{16}}}{(\pi\beta)^{3/2}} \int_{-\infty}^{\infty} \frac{p}{\sinh p} e^{-\frac{4p^2}{\beta}} dp.$$

We argue that the any other choice of the contour gives the physically meaningless partition function. Indeed, suppose we choose another contour that bypasses the singularities at $h = \pi n$ differently. Then at each pole where this contour differs from the Feynman like contour, (55) gets an additional contribution of the form $e^{-iC/T}$ with $C > 0$. However, this term analytically continues to $e^{C/\beta}$ for $\text{Im} T = -\beta < 0$, and this exponential high temperature growth contradicts the Weyl's law [19], which states that $Z(\beta) \sim \beta^{-d/2}$ as $\beta \rightarrow 0$.

• *Hyperbolic elements*

Element $\gamma \in \Gamma$ is hyperbolic, if $|\text{Tr } \gamma| > 2$, and since $Z_{[\gamma]} = Z_{[-\gamma]}$, we can always choose γ such that $\text{Tr } \gamma > 2$. Every such γ is conjugated in G to the element e^{rT_1} , where $r = \log \lambda(\gamma)$ and $\lambda(\gamma)$ is the multiplier of γ . We have $\Gamma_\gamma = \Gamma_{\gamma_0}$ and $G_\gamma = G_{\gamma_0}$, where γ_0 is a primitive hyperbolic element, $\gamma = \gamma_0^k$ for $k \geq 1$. Thus we can choose $\gamma = e^{rT_1}$ and $\gamma_0 = e^{r_0T_1}$ so $G_\gamma = \{e^{aT_1} | a \in \mathbb{R}\} \cup \{-e^{aT_1} | a \in \mathbb{R}\}$. Choosing the measure on G_γ to be da , we have

$$\text{vol}(\Gamma_\gamma \backslash G_\gamma) = \text{vol}(\Gamma_{\gamma_0} \backslash G_{\gamma_0}) = r_0, \quad \text{where } r_0 = \log \lambda(\gamma_0).$$

and

$$Z_{[\gamma]} = r_0 e^{-\frac{i\langle \rho, \rho \rangle T}{2}} \int_{-\pi}^{\pi} \frac{dh}{2\pi} \frac{i}{2 \sin h} \int_{G_\gamma \backslash G} \mathcal{A}_\gamma(g_0; \lambda) dg_0,$$

where the measure on $G_\gamma \backslash G$ should reflect our choice of the measure on G_γ above and will be determined later. Note that the final result is independent of this choice since $G = G_\gamma \times (G_\gamma \backslash G)$.

Now we compute $Z_{[\gamma]}$ using the following deformation

$$(57) \quad \begin{aligned} V &= -\frac{1}{2} \int_0^T \langle \partial_t J, \partial_t \psi \rangle dt, \\ \delta V &= \frac{1}{2} \int_0^T (\langle \partial_t J, \partial_t J \rangle + i \langle \partial_t \psi, \partial_t^J \partial_t \psi \rangle) dt. \end{aligned}$$

The stationary points are determined from the equations $\dot{J} = 0$ and $\partial_t D \dot{\psi} = 0$, where $D = \partial_t^J$, so

$$(58) \quad g_\gamma(t) = e^{rt/T} g_0 \quad \text{and} \quad J_\gamma(t) = \frac{1}{T} \text{Ad}_{g_0^{-1}} \mathbf{r},$$

where $\mathbf{r} = rT_1$ and $\psi = \psi_0 = \sum_{a=1,2} \psi_0^a T_a \in \Pi \mathfrak{p}$. As in the previous case $\gamma = I$, we easily compute the finite-dimensional fermionic integral

$$(59) \quad \int e^{-\frac{h}{2} \langle \psi_0, [T_3, \psi_0] \rangle} d\psi_0 = 2h,$$

and the remaining Gaussian path integral from the localization

$$(60) \quad \begin{aligned} \int e^{\frac{i}{2} \int_0^T (\langle D \partial_t Y, D \partial_t Y \rangle + i \langle \partial_t \eta, D \partial_t \eta \rangle)} \mathcal{D}Y \mathcal{D}\eta &= \frac{1}{(2\pi i)(-2\pi i)^{1/2}} \frac{\text{Pf}(-\partial_t^2 D)}{\sqrt{\det \partial_t^2 D^2}} \\ &= \frac{r}{(2\pi iT)^{3/2} \sinh r}. \end{aligned}$$

Thus we obtain

$$\begin{aligned}
 Z_{[\gamma]}(iT) &= \frac{r_0 r e^{-\frac{iT}{16}}}{(2\pi iT)^{3/2} \sinh r} \int_{-\pi}^{\pi} \frac{dh}{2\pi} \frac{ih}{\sin h} \int_{G_\gamma \backslash G} e^{iS[\mathbf{h}](g_\gamma, 0)} dg_0 \\
 (61) \quad &= \frac{r_0 r e^{\frac{4ir^2}{T} - \frac{iT}{16}}}{(2\pi iT)^{3/2} \sinh r} \int_{-\pi}^{\pi} \frac{dh}{2\pi} \frac{ih e^{\frac{4ih^2}{T}}}{\sin h} \int_{G_\gamma \backslash G} e^{-\frac{i\langle \text{Ad}_{g_0} \mathbf{h}, \mathbf{r} \rangle}{T}} dg_0 \\
 &= \frac{2r_0 r e^{\frac{4ir^2}{T} - \frac{iT}{16}}}{(\pi iT)^{3/2} \sinh r} \int_{-\pi}^{\pi} dh \frac{ih e^{\frac{4ih^2}{T}}}{\sin h} \int_{-\infty}^{\infty} e^{\frac{8ihrn}{T}} dn \\
 &= \frac{1}{2} \left(\frac{1}{\pi iT} \right)^{1/2} \frac{r_0}{\sinh r} e^{\frac{4ir^2}{T} - \frac{iT}{16}}.
 \end{aligned}$$

Here the orbital integral over $G_\gamma \backslash G$ in the third equality is computed as following. In addition to the Iwasawa decomposition $G = ANK$ (see Appendix A), we use the following parametrization of a generic element $g \in SL(2, \mathbb{R})$,

$$(62) \quad g = \begin{pmatrix} e^a & 0 \\ 0 & e^{-a} \end{pmatrix} \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}, \quad a, n \in \mathbb{R}, \quad 0 < \theta \leq 2\pi,$$

with the measure $dg = 8\sqrt{2}dadnd\theta$ is induced from the Cartan-Killing metric $ds^2 = 4 \text{Tr}(g^{-1}dg g^{-1}dg)$. Since $G_\gamma = \{e^{aT_1} \mid a \in \mathbb{R}\} \cup \{-e^{aT_1} \mid a \in \mathbb{R}\}$, we have a natural parametrization of $G_\gamma \backslash G$ as

$$\begin{aligned}
 (63) \quad G_\gamma \backslash G &= \mathbb{Z}_2 \backslash NK \\
 &= \left\{ \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}, \quad a, n \in \mathbb{R}, \quad 0 \leq \theta < \pi, \right\}.
 \end{aligned}$$

The measure da on G_γ fixes the measure on $G_\gamma \backslash G$ to be $dg_0 = 8\sqrt{2}dndk$, where $n \in \mathbb{R}$ and $0 \leq \theta < \pi$. Therefore, the orbital integral becomes

$$(64) \quad \int_{G_\gamma \backslash G} e^{-\frac{i\langle \text{Ad}_{g_0} \mathbf{h}, \mathbf{r} \rangle}{T}} dg_0 = 8\pi\sqrt{2} \int_{-\infty}^{\infty} e^{\frac{8ihrn}{T}} dn = \frac{2\sqrt{2}T}{r} \delta(h).$$

Finally, the last expression (61) of can be analytically continued to $T = -i\beta$ and we obtain

$$(65) \quad Z_{[\gamma]}(\beta) = \frac{1}{2\sqrt{\pi\beta}} \frac{r_0}{\sinh r} e^{-\frac{4r^2}{\beta} - \frac{\beta}{16}}.$$

• *Elliptic elements*

An elliptic element γ can be written as $\gamma = e^\theta$ with $\theta = \theta \tilde{T}_3$, where \tilde{T}_3 is conjugate to T_3 in G , and $\theta \in (0, \pi) \cup (\pi, 2\pi)$. Again, since the intervals $(0, \pi)$ and $(\pi, 2\pi)$ are related by $\gamma = -I$, it is sufficient to compute the case $\theta \in (0, \pi)$ and to use the same deformation as in the hyperbolic case,

$$(66) \quad \begin{aligned} V &= - \int dt \frac{1}{2} \langle \partial_t J, \partial_t \psi \rangle \\ \delta V &= \int dt \frac{1}{2} \langle \partial_t J, \partial_t J \rangle + \frac{i}{2} \langle \partial_t \psi, \partial_t^J \partial_t \psi \rangle. \end{aligned}$$

This gives the set of isolated critical points

$$(67) \quad g_n(t) = e^{\theta n t/T} g_0 \quad \text{and} \quad J_n(t) = \text{Ad}_{g_0^{-1}} \theta_n,$$

where

$$\theta_n = \theta_n \tilde{T}_3 \quad \text{and} \quad \theta_n = \theta + 2\pi n, \quad n \in \mathbb{Z}.$$

Their contribution to the partition function is given by

$$(68) \quad \begin{aligned} &Z_{[\gamma]}(iT) \\ &= \frac{\text{vol}(\Gamma_\gamma \backslash G_\gamma) e^{-\frac{iT}{16}}}{(2\pi iT)^{3/2}} \int_{-\pi}^{\pi} \frac{dh}{2\pi} \sum_{n \in \mathbb{Z}} \int_{G_\gamma \backslash G} dg_0 \frac{ih\theta_n}{\sin h \sin \theta_n} e^{-\frac{4i(\theta_n^2 + h^2) + i \langle \text{Ad}_{g_0} h, \theta_n \rangle}{T}}. \end{aligned}$$

Here we've been careful with the order of integration and summation. The orbital integral over $G_\gamma \backslash G$ can be computed explicitly using the Cartan decomposition $G = KAK$,

$$(69) \quad g = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} \begin{pmatrix} \cos w & \sin w \\ -\sin w & \cos w \end{pmatrix},$$

where $t \geq 0$, $0 \leq \alpha < 2\pi$ and $0 \leq w < \pi$. Here the dg measure is induced from the Cartan-Killing metric $ds^2 = 4 \text{Tr}(g^{-1} dg g^{-1} dg)$ and we have

$$dg = 16\sqrt{2} \sinh(2t) dt d\alpha dw.$$

Since $G_\gamma = K$, the natural parametrization of $g_0 \in G_\gamma \backslash G$ is

$$(70) \quad g_0 = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} \begin{pmatrix} \cos w & \sin w \\ -\sin w & \cos w \end{pmatrix}, \quad t \in \mathbb{R}, \quad 0 \leq w < \pi.$$

As in the hyperbolic case, we choose the measure on G_γ to be $d\alpha$, so

$$(71) \quad \text{vol}(\Gamma_\gamma \backslash G_\gamma) = \theta_p, \quad dg_0 = 16\sqrt{2} \sinh(2t) dt dw,$$

where θ_p is such that $e^{\theta_p T_3}$ is the primitive element in Γ_γ . As before, the measure on $G_\gamma \backslash G$ is determined from $G = G_\gamma \times (G_\gamma \backslash G)$.

The relevant orbital integral becomes

$$\begin{aligned}
 \int_{G_\gamma \backslash G} dg_0 e^{-\frac{i \langle \text{Ad}_{g_0} h, \theta_n \rangle}{T}} &= 16\pi\sqrt{2} \int_0^\infty e^{\frac{8ih\theta_n \cosh(2t)}{T}} \sinh(2t) dt \\
 (72) \qquad \qquad \qquad &= 8\pi\sqrt{2} \int_1^\infty e^{\frac{8ih\theta_n y}{T}} dy \\
 &= \frac{i\pi\sqrt{2}T}{h\theta_n} e^{\frac{8ih\theta_n}{T}} + \sqrt{2}\pi^2 T \delta(h\theta_n),
 \end{aligned}$$

which gives

$$\begin{aligned}
 &Z_{[\gamma]}(iT) \\
 &= \frac{\theta_p e^{-\frac{iT}{16}}}{2(2\pi iT)^{3/2}} \int_{-\pi}^\pi dh \sum_{n \in \mathbb{Z}} \frac{ih}{\sin h} \frac{\theta_n}{\sin \theta_n} e^{-\frac{4i(\theta_n^2 + h^2)}{T}} \left(\frac{i\sqrt{2}T}{h\theta_n} e^{\frac{8ih\theta_n}{T}} + \frac{\sqrt{2}\pi T}{\theta_n} \delta(h) \right) \\
 &= \frac{\theta_p e^{-\frac{iT}{16}}}{4(\pi iT)^{1/2} \sin \theta} \int_{-\pi}^\pi dh \sum_{n \in \mathbb{Z}} \frac{ih}{\pi \sin h} e^{-\frac{4i(\theta_n - h)^2}{T}} \frac{1}{h + i0_+}.
 \end{aligned}$$

To simplify this expression, we formally interchange the integration over h and the summation over the set of critical points. As before, we choose the Feynman-like contour which is uniquely determined by the compatibility with the Weyl's law. Such contour for $\theta \in (0, \pi)$ is $\mathcal{C} = \mathbb{R}(1 + i\epsilon)$, and we obtain

$$(73) \qquad Z_{[\gamma]}(iT) = \frac{\theta_p e^{-\frac{iT}{16}}}{4(\pi iT)^{1/2} \sin \theta} \int_{\mathcal{C}} \frac{i}{\pi \sin h} e^{-\frac{4i(\theta - h)^2}{T}} dh.$$

The analytically continued expression can be compactly written as

$$(74) \qquad Z_{[\gamma]}(\beta) = \frac{\theta_p e^{-\frac{\beta}{16}}}{4(\pi\beta)^{1/2} \sin \theta} \left(\text{v.p.} \int_{\mathbb{R}} \frac{i}{\pi \sinh p} e^{-\frac{4(p+i\theta)^2}{\beta}} dp + e^{\frac{4\theta^2}{\beta}} \right).$$

We can simplify further by representing the exponent $e^{-\frac{4(p+i\theta)^2}{\beta}}$ as another Gaussian integral and changing the order of integrations (which is now

legitimate)

$$\begin{aligned}
 & Z_{[\gamma]}(\beta) \\
 = & \frac{\theta_p e^{-\frac{\beta}{16}}}{4(\pi\beta)^{1/2} \sin \theta} \left(\text{v.p.} \frac{i}{\pi} \sqrt{\frac{\beta}{16\pi}} \int_{\mathbb{R}} dp \int_{\mathbb{R}} \frac{e^{-\frac{\beta u^2}{16}} \cos((p+i\theta)u)}{\sinh p} du + e^{4\theta^2/\beta} \right) \\
 = & \frac{\theta_p e^{-\frac{\beta}{16}}}{16\pi \sin \theta} \left(\frac{1}{\pi} \int_{\mathbb{R}} du \int_{\mathbb{R}} e^{-\frac{\beta u^2}{16}} \frac{\sin(pu) \sinh(\theta u)}{\sinh p} dp + \int_{\mathbb{R}} du e^{-\frac{\beta u^2}{16}} \cosh(\theta u) \right) \\
 = & \frac{\theta_p e^{-\frac{\beta}{16}}}{16\pi \sin \theta} \int_{\mathbb{R}} du e^{-\frac{\beta u^2}{16}} \frac{\cosh((\pi/2 - \theta)u)}{\cosh(\pi u/2)}.
 \end{aligned}$$

• *Final Result*

Collecting results in this section we obtain the Selberg trace formula for the compact Riemann surface $X = \Gamma \backslash \mathbb{H}$,

$$\begin{aligned}
 \text{Tr}_{L_0^2(X)}[e^{-\beta\Delta/2}] = & \frac{\text{vol}(X)e^{-\frac{\beta}{16}}}{(\pi\beta)^{3/2}} \int_{\mathbb{R}} dp \frac{p}{\sinh p} e^{-\frac{4p^2}{\beta}} \\
 & + \sum_{\substack{[\gamma] \text{ hyperbolic} \\ \text{Tr}[\gamma] > 2}} \frac{1}{\sqrt{\pi\beta}} \frac{r_0}{\sinh r} e^{-\frac{4r^2}{\beta} - \frac{\beta}{16}} \\
 (75) \quad & + \sum_{\substack{[\gamma] \text{ elliptic} \\ 0 < \theta < \pi}} \frac{\theta_p}{8\pi \sin \theta} \int_{\mathbb{R}} du e^{-\frac{\beta(u^2+1)}{16}} \frac{\cosh((\pi/2 - \theta)u)}{\cosh(\pi u/2)}.
 \end{aligned}$$

After the rescaling $\Delta \rightarrow \Delta/2$ and $\beta \rightarrow 4\beta$, it exactly matches the standard form of the Selberg trace formula [3, 20]

$$\begin{aligned}
 \text{Tr}_{L_0^2(X)}[e^{-\beta\Delta}] = & \frac{\mu(X)e^{-\frac{\beta}{4}}}{(4\pi\beta)^{3/2}} \int_0^\infty dp \frac{p}{\sinh \frac{p}{2}} e^{-\frac{p^2}{4\beta}} \\
 & + \sum_{\substack{[\gamma] \text{ hyperbolic} \\ \text{Tr}[\gamma] > 2}} \frac{1}{\sqrt{4\pi\beta}} \frac{r_0}{\sinh r} e^{-\frac{r^2}{\beta} - \frac{\beta}{4}} \\
 (76) \quad & + \sum_{\substack{[\gamma] \text{ elliptic} \\ 0 < \theta < \pi}} \frac{\theta_p}{4\pi \sin \theta} \int_{\mathbb{R}} \frac{\cosh((\pi - 2\theta)r)}{\cosh(\pi r)} e^{-\beta(r^2 + \frac{1}{4})} dr,
 \end{aligned}$$

where $\mu(X) = \text{vol}(X)/2$ is the hyperbolic area of X .

4.5. Generalization to arbitrary representations

The trace formula on the Hilbert space $L_0^2(X)$ can be generalized to a trace formula on the Hilbert space $L_0^2(X; \pi)$ associated with a unitary finite-dimensional representation π ¹¹ of the Fuchsian group Γ . Elements $\Psi \in L_0^2(X; \pi)$ can be thought as sections of a vector-bundle on X associated with the representation π or, equivalently, as vector-valued functions Ψ on G , satisfying the following constraints

$$(77) \quad \Psi(\gamma x) = \pi(\gamma)\Psi(x), \quad \Psi(xk) = \Psi(x), \quad x \in G, \gamma \in \Gamma, k \in K.$$

To perform a path integral quantization, we lift the representation π of Γ to the representation of $\pi_1(\Gamma \backslash G, \bar{e})$ and observe that for each connected component of the domain of integration (38) we get an additional weight factor $\chi_\pi(\gamma)$, where χ_π is the character of π (cf. [22–24]). Hence the partition function decomposition (39) is modified as

$$(78) \quad Z(iT; \pi) = \sum_{[\gamma]} \chi_\pi(\gamma) Z_{[\gamma]}(iT),$$

where $Z_{[\gamma]}(iT)$ is given by the same path integral that we previously localized in the case of trivial representation. Consequently, we obtain the trace formula on $L^2(X; \pi)$,

$$(79) \quad \begin{aligned} \text{Tr}_{L^2(X; \pi)}[e^{-\beta\Delta/2}] &= \frac{d_\pi \text{vol}(X) e^{-\frac{\beta}{16}}}{(\pi\beta)^{3/2}} \int_{\mathbb{R}} dp \frac{p}{\sinh p} e^{-\frac{4p^2}{\beta}} \\ &+ \sum_{\substack{[\gamma] \text{ hyperbolic} \\ \text{Tr}[\gamma] > 2}} \frac{\chi_\pi(\gamma)}{\sqrt{\pi\beta}} \frac{r_0}{\sinh r} e^{-\frac{4r^2}{\beta} - \frac{\beta}{16}} \\ &+ \sum_{\substack{[\gamma] \text{ elliptic} \\ 0 < \theta < \pi}} \frac{\chi_\pi(\gamma) \theta_p}{8\pi \sin \theta} \int_{\mathbb{R}} du e^{-\frac{\beta(u^2+1)}{16}} \frac{\cosh((\pi/2 - \theta)u)}{\cosh(\pi u/2)}, \end{aligned}$$

which simply reproduces the standard result [3, 20].

¹¹There should be no confusion between the representation π and the area of the circle of radius 1.

4.6. Generalization to arbitrary weights

It is quite amusing to observe that the path integral derivation of the Selberg trace formula provides a unified simple approach for the case of automorphic forms of arbitrary weight $m \in \frac{1}{2}\mathbb{Z}$.

Let us first consider the case of integer weight $m \in \mathbb{Z}$. As in the case of automorphic functions (the forms of zero weight), one can show the following equality

$$(80) \quad \text{Tr}_{L_m^2(X)}[e^{-iT\Delta/2}] = \frac{e^{-\frac{i(\rho,\rho)T}{2}}}{\text{vol}(\mathcal{G})} \int \frac{\mathcal{W}_{2m}(A)}{\mathcal{W}_{-1}(A) - \mathcal{W}_1(A)} \psi_{\mathfrak{t}} e^{iS_A[g,\psi]} \mathcal{D}g \mathcal{D}\psi \mathcal{D}A.$$

The only difference with the case $m = 0$ is the appearance of the additional Wilson line in the path integral, which depends only on the gauge field. Therefore, the localization procedure is almost identical to the weight zero case with the same prescription for the physical contour! Thus after the Wick rotation we obtain

$$(81) \quad \begin{aligned} \text{Tr}_{L_m^2(X)}[e^{-\beta\Delta/2}] &= \frac{\text{vol}(X)e^{-\frac{\beta}{16}}}{(\pi\beta)^{3/2}} \int_{\mathbb{R}} dp \frac{p \cosh(2mp)}{\sinh p} e^{-\frac{4p^2}{\beta}} \\ &+ \sum_{\substack{[\gamma] \text{ hyperbolic} \\ \text{Tr}[\gamma] > 2}} \frac{1}{\sqrt{\pi\beta}} \frac{r_0}{\sinh r} e^{-\frac{4r^2}{\beta} - \frac{\beta}{16}} \\ &+ \sum_{\substack{[\gamma] \text{ elliptic} \\ 0 < \theta < \pi}} \frac{\theta_p e^{-\frac{\beta}{16}}}{2(\pi\beta)^{1/2} \sin \theta} \left(\int_{\mathbb{R}} \frac{ie^{2mp}}{\pi \sinh p} e^{-\frac{4(p+i\theta)^2}{\beta}} dp + e^{\frac{4\theta^2}{\beta}} \right), \end{aligned}$$

As in Section 4.5, this formula admits a straightforward generalization to the Hilbert space $L_m^2(X; \pi)$ associated with a general unitary finite-dimensional representation π of Γ satisfying $\pi(-I) = \pi(I)$.

However, in case of $m \in \mathbb{Z} + \frac{1}{2}$ we have seen in Section that the Hilbert space $L_m^2(X)$ is empty because $-1 \in \Gamma$. Therefore, it is necessary to consider a non-trivial one-dimensional representation π which obeys the consistency condition $\pi(-I) = -1$. It is striking to see how simply the path integral

approach reproduces the correct result! Thus performing the same manipulation as in $m = 0$ case, in basic case $m = \frac{1}{2}$ we obtain the following formula

$$\begin{aligned}
 \text{Tr}_{L^2_{\frac{1}{2}}(X;\pi)}[e^{-\beta\Delta/2}] &= \frac{\text{vol}(X)e^{-\frac{\beta}{16}}}{8\pi^2} \int_{\mathbb{R}} dp p \coth p e^{-\frac{\beta u^2}{4}} \\
 &+ \sum_{\substack{[\gamma] \text{ hyperbolic} \\ \text{Tr}[\gamma] > 2}} \frac{\chi_{\pi}(\gamma)}{\sqrt{\pi\beta}} \frac{r_0}{\sinh r} e^{-\frac{4r^2}{\beta} - \frac{\beta}{16}} \\
 (82) \quad &+ \sum_{\substack{[\gamma] \text{ elliptic} \\ 0 < \theta < \pi}} \frac{\chi_{\pi}(\gamma)\theta_p}{8\pi \sin \theta} \left(\int_{\mathbb{R}} e^{-\frac{\beta(u^2+1)}{16}} \frac{\sinh((\pi/2 - \theta)u)}{\sinh(\pi u/2)} du + 2i \right).
 \end{aligned}$$

Needless to say that it exactly matches the known result [3, 20].

Remark 7. The path integral approach we developed here is also applicable to compact homogeneous spaces G/K , where G is a compact Lie group and K is its subgroup. In this case we can directly use the Euclidean formalism for formulas like (27), which is a major simplification. This approach provides a compact and simple derivation of known results related to the Eskin trace formula, including the basic case $SU(2)/U(1) \simeq S^2$ (see [25] and references therein).

5. Selberg trace formula on general compact locally symmetric space

Here we generalize further our path integral localization approach to include the Selberg trace formula for a general compact locally symmetric space $X = \Gamma \backslash G/K$, where G is a non-compact real semi-simple Lie group of rank r_G , K is its maximal compact subgroup and Γ is a discrete subgroup of G such that $\Gamma \backslash G$ is compact. For simplicity, we consider only the trace formula for Laplace operator acting on functions on X , which is enough to convey the main novelty.

5.1. General set-up for $\Gamma \backslash G/K$

We start with the same gauged sigma model (16) on G/K with the same supersymmetry (17), where now A is a connection in a principal K -bundle over S^1_T . There is a main technical difficulty if one tries to follow the logic developed for the Riemann surfaces due to the fact that the structure group

K is non-abelian and hence the Gauss law (with or without additional Wilson lines) would be more complicated. Namely, it is difficult to anticipate the analogous insertion of Wilson lines to the path integral that gives the Hilbert space $L_0^2(X)$.

It is remarkable that there is a short-cut, which bypasses the above obstacles and produces the desired path integral identity for the trace formula!

Namely, let P be the principal G -bundle over S_T^1 and let $\nabla_A = d + \text{ad}_A$ be a connection on the associated adjoint bundle $P \times_{\text{Ad}|_K} \mathfrak{g}$, where $\text{Ad}|_K : K \hookrightarrow G \rightarrow \text{Ad}(\mathfrak{g}) \subset \text{GL}(\mathfrak{g})$ is a natural adjoint action of K on \mathfrak{g} . We claim that the generalization of (36) is the following formula

$$(83) \quad \begin{aligned} & \text{Tr}[e^{-iT\Delta/2}] \\ &= \frac{e^{-i\langle \rho, \rho \rangle T/2}}{\text{vol}(\mathcal{G})} \int \frac{\chi(A) e^{i \int_0^T (\frac{1}{2} \langle J_A, J_A \rangle + \frac{i}{2} \langle \psi, \partial_t^A \psi \rangle) dt}}{\text{Pf}' \left(i(\text{Hol}_{S_T^1}^{-1/2}(\nabla_A) - \text{Hol}_{S_T^1}^{1/2}(\nabla_A)) \right)} \mathcal{D}g \mathcal{D}\psi \mathcal{D}A. \end{aligned}$$

Here $\text{Hol}_{S_T^1}(\nabla_A) \in \text{GL}(\mathfrak{g})$ is the the holonomy along the temporal circle and Pf' is a Pfaffian on the non-zero eigenvalue subspace in \mathfrak{g} , which will become clear in a moment. Finally, $\chi(A)$ denotes the insertion of r_G fermionic zero modes, associated with the kernel of the operator ∇_A acting on \mathfrak{g} -valued functions on S_T^1 .

Though the explicit form of $\chi(A)$ for general time-dependent A is complicated, it can be obtained by first fixing the gauge with $A = \text{const}$ where the fermionic zero modes are simple, and then gauge transforming it back to get fermionic zero modes for general A .

Namely, since two connections with the same holonomy $u \in K$ are gauge equivalent and the exponential map $\exp : \mathfrak{k} \rightarrow K$ is surjective, for a connection ∇_A with the holonomy $u = e^{\mathbf{h}}$ where $\mathbf{h} \in \mathfrak{k}$, we can choose the gauge $\dot{A} = 0$, so $A = \mathbf{h}/T$. Since the conjugation by K is a residual gauge symmetry, we can let $\mathbf{h} \in \mathfrak{t}$, where \mathfrak{t} is the Cartan subalgebra of \mathfrak{k} . Since generic A is regular in \mathfrak{k} , so that its centralizer in \mathfrak{k} is \mathfrak{t} , i.e. $Z_{\mathfrak{k}}(A) = \mathfrak{t}$ and so $\ker(\nabla_A) \simeq Z_{\mathfrak{g}}(\mathfrak{t})$. Now it follows from [26, Proposition 6.60] that

$$(84) \quad Z_{\mathfrak{g}}(\mathfrak{t}) = \mathfrak{h}^c,$$

where $\mathfrak{h}^c \subset \mathfrak{g}$ is a maximally compact Cartan subalgebra in \mathfrak{g} which is stable under the Cartan involution θ on \mathfrak{g} , which satisfies $\theta^2 = 1$ with \mathfrak{k} and \mathfrak{p} being $+1$ and -1 eigenspaces. This leads to a naturally decomposition $\mathfrak{h}^c = \mathfrak{t} \oplus \mathfrak{a}$, where $\mathfrak{t} \subset \mathfrak{k}$ and $\mathfrak{a} \subset \mathfrak{p}$.

This clearly shows that for general A there are $r_G = \dim \mathfrak{h}^c$ linear independent fermionic zero modes. Representing them by the elements $\chi^1, \dots, \chi^{r_G}$

associated with the orthonormal basis of \mathfrak{h}^c , we have insertion of zero modes

$$(85) \quad \chi(A) = c_{r_G} \chi^1 \cdots \chi^{r_G}.$$

As in Section 4.2, to prove the main formula (83) it is sufficient to show that, with some appropriate normalization c_G , the fermionic path integral cancels the denominator in (83), which is the following identity

$$(86) \quad \int \chi(A) e^{i \int_0^T \frac{i}{2} \langle \psi, \partial_t^A \psi \rangle dt} \mathcal{D}\psi = \text{Pf}' \left(i(\text{Hol}_{S_T^1}^{-1/2}(\nabla_A) - \text{Hol}_{S_T^1}^{1/2}(\nabla_A)) \right).$$

To prove (86), we again fix the gauge to $A = \mathbf{h}/T$ with $\mathbf{h} \in \mathfrak{t}$ and as in Section 4.2, use the Hamiltonian formalism which generalizes (34).

Let $\mathfrak{h}_{\mathbb{C}}^c$ the complexification of \mathfrak{h}^c which is a Cartan subalgebra of $\mathfrak{g}_{\mathbb{C}}$, which leads to a decomposition $\mathfrak{g}_{\mathbb{C}} = \mathfrak{h}_{\mathbb{C}}^c \oplus \bigoplus_{\alpha \in R_{\mathfrak{g}_{\mathbb{C}}, \mathfrak{h}_{\mathbb{C}}^c}} \mathfrak{g}_{\alpha}$ where $R_{\mathfrak{g}_{\mathbb{C}}, \mathfrak{h}_{\mathbb{C}}^c}$ is a root system of $\mathfrak{g}_{\mathbb{C}}$ w.r.t. $\mathfrak{h}_{\mathbb{C}}^c$. As in the Appendix A of [2], we introduce a Cartan-Weyl basis of $\mathfrak{g}_{\mathbb{C}}$ such that $\{H_i\}$ and $\{E_{\alpha}\}$ with a normalization $\langle H_i, H_j \rangle = \delta_{ij}$ and $\langle E_{\alpha}, E_{\beta} \rangle = \delta_{\alpha, -\beta}$. We can it to expand ψ as

$$\psi = \sum_{i=1}^{r_G} \psi^i H_i + \sum_{\alpha \in R_{\mathfrak{g}_{\mathbb{C}}, \mathfrak{h}_{\mathbb{C}}^c}} \psi^{\alpha} E_{\alpha}.$$

Therefore, canonical quantization gives the commutation relations

$$[\hat{\psi}^i, \hat{\psi}^j] = \delta^{ij} \quad \text{and} \quad [\hat{\psi}^{\alpha}, \hat{\psi}^{-\alpha}] = 1,$$

To make sense of the quantization, we need to impose the following conditions as in Section 3

$$(\hat{\psi}^i)^{\dagger} = \hat{\psi}^i, \quad (\hat{\psi}^{\alpha})^{\dagger} = \hat{\psi}^{-\alpha}.$$

On the path integral side, this means that ψ belongs to an appropriate middle dimensional contour of $\Pi L\mathfrak{g}_{\mathbb{C}}$.

As in Section 4.2, in the Hamiltonian representation the fermion path integral becomes the trace

$$(87) \quad \int \chi(A) e^{-\frac{1}{2} \int_0^T \langle \psi, \partial_t^{\mathbf{h}/T} \psi \rangle dt} \mathcal{D}\psi = \text{Tr}_{\mathcal{H}_{F, \mathfrak{g}}} \left[(-1)^F \hat{\chi}(A) e^{-i\hat{H}_F T} \right],$$

and

$$\begin{aligned}
 \hat{H}_F &= \sum_{\alpha \in R_{\mathfrak{g}_\mathbb{C}; \mathfrak{h}_\mathbb{C}}^+} \frac{i \langle \alpha, \mathbf{h} \rangle}{T} \frac{\hat{\psi}^\alpha \hat{\psi}^{-\alpha} - \hat{\psi}^{-\alpha} \hat{\psi}^\alpha}{2} \\
 (88) \quad &= \sum_{\alpha \in R_{\mathfrak{g}_\mathbb{C}; \mathfrak{h}_\mathbb{C}}^+} \frac{i \langle \alpha, \mathbf{h} \rangle}{T} \left(\hat{\psi}^\alpha \hat{\psi}^{-\alpha} - \frac{1}{2} \right).
 \end{aligned}$$

Now in this gauge, we have an identification $\hat{\chi}^i = \hat{\psi}^i$ with $i = 1, \dots, r_G$ and $c_G = \pm i^{r_G(r_G-1)/2}$, and this makes the product of the inserted operator $\hat{\psi}(A)$ with $(-1)^F$ effectively as a projection operator to the reduced fermionic Hilbert space

$$\mathcal{H}_{\mathfrak{g}/\mathfrak{h}^c} = \bigotimes_{\alpha \in R_{\mathfrak{g}_\mathbb{C}; \mathfrak{h}_\mathbb{C}}^+} \mathbb{C}^2,$$

where each \mathbb{C}^2 is an irreducible module for the pair $\hat{\psi}_\alpha, \hat{\psi}_{-\alpha}$.

For each factor \mathbb{C}^2 associated with a positive root α , we define the fermion number operator by

$$(89) \quad \hat{N}_{F,\alpha} = \hat{\psi}^\alpha \hat{\psi}^{-\alpha}, \quad \alpha > 0.$$

Together with $(\hat{\psi}^\alpha)^\dagger = \hat{\psi}^{-\alpha}$, this means that $\hat{\psi}^\alpha$ is a creation operator and $\hat{\psi}^{-\alpha}$ as an annihilation operator. As a consequence,

$$-i \hat{H}_F T = \sum_{\alpha \in R_{\mathfrak{g}_\mathbb{C}; \mathfrak{h}_\mathbb{C}}^+} \langle \alpha, \mathbf{h} \rangle \left(\hat{N}_{F,\alpha} - \frac{1}{2} \right)$$

and

$$\text{Tr}_{\mathcal{H}_{F,\mathfrak{g}/\mathfrak{h}^c}} \left[e^{-i \hat{H}_F T} \right] = \prod_{\alpha \in R_{\mathfrak{g}_\mathbb{C}; \mathfrak{h}_\mathbb{C}}^+} \left(e^{-\frac{\langle \alpha, \mathbf{h} \rangle}{2}} - e^{\frac{\langle \alpha, \mathbf{h} \rangle}{2}} \right).$$

Therefore, the fermionic path integral (86) for $A = \mathbf{h}/T$ in the Hamiltonian formulation is unambiguously given by¹²

$$(90) \quad \tau(\mathbf{h}) \stackrel{\text{def}}{=} \prod_{\alpha \in R_{\mathfrak{g}_\mathbb{C}; \mathfrak{h}_\mathbb{C}}^+} \left(e^{-\frac{\langle \alpha, \mathbf{h} \rangle}{2}} - e^{\frac{\langle \alpha, \mathbf{h} \rangle}{2}} \right).$$

¹²One can also compute the fermionic path integral directly using the zeta function regularization analogous to [2]. We note that for the consistent result, it is important to determine the path integral measure carefully, especially such that it takes into account our choice of the fermion number in (89).

The product can be written as a Pfaffian of the following skew-symmetric linear operator acting on $\mathfrak{g}/\mathfrak{h}^c$

$$(91) \quad \tau(\mathbf{h}) = \text{Pf} \left(i(e^{-\frac{1}{2}\text{ad}_{\mathbf{h}}} - e^{\frac{1}{2}\text{ad}_{\mathbf{h}}}) \right) \Big|_{\mathfrak{g}/\mathfrak{h}^c},$$

where the sign convention of the Pfaffian is chosen to be consistent with the definition (89) of the fermion number and is given by the right hand side of (90). Since $e^{\text{ad}_x} = \text{Ad}_{e^x}$, we can rewrite $\tau(\mathbf{h})$ as

$$(92) \quad \tau(\mathbf{h}) = \text{Pf} \left(i \left(\text{Ad}_{e^{-\frac{1}{2}\mathbf{h}}} - \text{Ad}_{e^{\frac{1}{2}\mathbf{h}}} \right) \right) \Big|_{\mathfrak{g}/\mathfrak{h}^c}.$$

Finally, the holonomies of gauge equivalent connections are related by the adjoint action of K , so for a connection ∇_A in the adjoint bundle $P \times_{\text{Ad}|_K} \mathfrak{g}$ which is gauge equivalent to \mathbf{h}/T we have

$$(93) \quad \tau(\mathbf{h}) = \text{Pf}' \left(i \left(\text{Hol}_{S_T^1}^{-1/2}(\nabla_A) - \text{Hol}_{S_T^1}^{1/2}(\nabla_A) \right) \right).$$

This proves (86) and together with the DeWitt term originated from the supersymmetry algebra, completes the proof of the main identity (83).

Remark 8. The definition (89) is consistent with the choice of fermion number operator in case $G = \text{SL}(2, \mathbb{R})$ in Sections 4.1–4.2. Namely, we have (see Appendix A)

$$H = \frac{1}{2\sqrt{2}i}T_3, \quad E_\alpha = \frac{1}{4}(T_1 + iT_2), \quad E_{-\alpha} = \frac{1}{4}(T_1 - iT_2),$$

so

$$\langle H, H \rangle = \langle E_\alpha, E_{-\alpha} \rangle = 1 \quad \text{and} \quad \alpha(H) = \frac{1}{\sqrt{2}}, \quad \langle \alpha, \mathbf{h} \rangle = 2ih.$$

Therefore, formula (88) gives Hamiltonian (32).

Now we are ready to obtain the path-integral version of the pre-trace formula from (83) by carefully taking care of the path integral measure. We gauge fix the right hand side of (83) by imposing $\dot{A} = 0$ so that $A = \mathbf{h}/T$ with $\mathbf{h} \in \mathfrak{k}$. Then the gauge fixed path integral in terms of holonomy $u = e^{\mathbf{h}} \in K$ can be written as

$$(94) \quad \begin{aligned} Z(iT) &= \text{Tr}[e^{-iT\Delta/2}] \\ &= \frac{e^{-\frac{i\langle \rho, \rho \rangle T}{2}}}{\text{vol}(K)} \int_K du \int \frac{1}{\tau(\mathbf{h})} \psi_{\mathfrak{h}^c} e^{iS_f[\mathbf{h}](g, \psi)} \mathcal{D}g \mathcal{D}\psi, \end{aligned}$$

where we denote $\psi_{\mathfrak{h}^c} = \psi(\mathbf{h}/T)$. Note that since two connections with same holonomy are gauge equivalent, the integral does not depend on the choice of \mathbf{h} in $u = e^{\mathbf{h}}$.

The second residual gauge symmetry, which is present for non-abelian K , is given global gauge transformations $A \rightarrow kAk^{-1}$ with $k \in K$. Therefore we can gauge fix such that $\mathbf{h} \in \mathfrak{t}$ and $\mathbf{t} = e^{\mathbf{h}} \in T_K$, where T_K is the maximal torus of K . Using the Weyl integration formula we obtain after the gauge fixing

$$(95) \quad \begin{aligned} Z(iT) &= \text{Tr}[e^{-iT\Delta/2}] \\ &= \frac{e^{-\frac{i\langle \rho, \rho \rangle T}{2}}}{|W_K| \text{vol}(T_K)} \int_{T_K} dt |\delta(\mathbf{t})|^2 \int \frac{1}{\tau(\mathbf{h})} \psi_{\mathfrak{h}^e} e^{iS_f[\mathbf{h}](g, \psi)} \mathcal{D}g \mathcal{D}\psi. \end{aligned}$$

Here W_K is the Weyl group of K and the correct radial measure can be read off from

$$(96) \quad \delta(\mathbf{t}) = \det(1 - \text{Ad}_{\mathbf{t}})|_{\mathfrak{t}/\mathfrak{t}} \quad \text{and} \quad |\delta(\mathbf{t})|^2 = \prod_{\alpha \in R_{\mathfrak{t}_{\mathbb{C}}, \mathfrak{t}_{\mathbb{C}}}} \left(e^{\frac{\langle \alpha, \mathbf{h} \rangle}{2}} - e^{-\frac{\langle \alpha, \mathbf{h} \rangle}{2}} \right)^2.$$

Finally, as in Section 4.3, we obtain the path integral pre-trace formula

$$(97) \quad \begin{aligned} Z(iT) &= \sum_{[\gamma]} Z_{[\gamma]}(iT) \\ &= \sum_{[\gamma]} \text{vol}(\Gamma_{\gamma} \backslash G_{\gamma}) \int_{G_{\gamma} \backslash G} \mathcal{F}_{\gamma}(g_0; \lambda) dg_0, \end{aligned}$$

where

$$(98) \quad \begin{aligned} &\mathcal{F}_{\gamma}(g_0; \lambda) \\ &= \frac{e^{-\frac{i\langle \rho, \rho \rangle T}{2}}}{|W_K| \text{vol}(T_K)} \int_{T_K} dt |\delta(\mathbf{t})|^2 \int_{\mathcal{P}_{\gamma}(G; g_0) \times \Pi L_{\mathfrak{g}}} \frac{1}{\tau(\mathbf{h})} \psi_{\mathfrak{h}^e} e^{iS_f[\mathbf{h}]} \mathcal{D}g \mathcal{D}\psi. \end{aligned}$$

As in case of $G = \text{SL}(2, \mathbb{R})$, we compute each contribution to the pre-trace formula (97) separately by using a supersymmetric localization with a γ -dependent deformation $S_f[\mathbf{h}] \rightarrow S_f[\mathbf{h}] + \lambda \delta V_{\gamma}$. As before, we see that conjugacy classes $[\gamma]$ and $[\gamma']$ give the same path integral contribution if $\gamma' = z\gamma$ where $z \in \mathcal{Z}(G)$, the center of G . This is clear geometrically since K contains $\mathcal{Z}(G)$. Therefore, we divide the conjugacy classes in Γ by their action on G/K .

- *Conjugacy classes with trivial action*

They are of the form $[\gamma] = z$, where $z \in \mathcal{Z}_\Gamma(G) = \mathcal{Z}(G) \cap \Gamma$, so we can choose $[\gamma] = I$. In this case the deformation of the same form as in Section 4.4,

$$(99) \quad \begin{aligned} V_0 &= - \int dt \frac{1}{2} \langle \partial_t^{A_0} J, \partial_t^{A_0} \psi \rangle, \\ \delta V_0 &= \int dt \frac{1}{2} \langle \partial_t^{A_0} J, \partial_t^{A_0} J \rangle + \frac{i}{2} \langle \partial_t^{A_0} \psi, \partial_t^J \partial_t^{A_0} \psi \rangle, \end{aligned}$$

where A_0 is regular element in \mathfrak{k} . The isolated critical points are given by

$$(100) \quad g_{\{n_j\}}(t) = \exp\left(\frac{n_j \mathfrak{z}_j t}{T}\right) g_0,$$

where $\{\mathfrak{z}_j\}$ is a basis of characteristic lattice Γ_K of the maximal torus, $T_K = \mathfrak{k}/\Gamma_K$. The computation is analogous to $G = \text{SL}(2, \mathbb{R})$ case in Section 4.4, and we get

$$(101) \quad \begin{aligned} Z_I(iT) &= \frac{\text{vol}(\Gamma \backslash G) e^{-\frac{i \langle \rho, \rho \rangle T}{2}}}{(2\pi iT)^{d_p/2} (-2\pi iT)^{d_\mathfrak{k}/2} i^{\dim(R_{\mathfrak{g}_\mathbb{C}^+; \mathfrak{h}_\mathbb{C}^+})} |W_K| \text{vol}(T_K)} \\ &\times \int_{T_K} \frac{|\delta(\mathfrak{t})|^2}{\tau(\mathfrak{h})} \sum_{\mathfrak{z} \in \Gamma_K} \left(\prod_{\alpha \in R_{\mathfrak{g}_\mathbb{C}^+; \mathfrak{h}_\mathbb{C}^+}} -i \langle \mathfrak{h} - \mathfrak{z}, \alpha \rangle \right) e^{\frac{i \langle \mathfrak{h} - \mathfrak{z}, \mathfrak{h} - \mathfrak{z} \rangle}{2T}} dt, \end{aligned}$$

where $\dim(R_{\mathfrak{g}_\mathbb{C}^+; \mathfrak{h}_\mathbb{C}^+})$ is a number of positive roots in $R_{\mathfrak{g}_\mathbb{C}^+; \mathfrak{h}_\mathbb{C}^+}$.

Thus the total contribution of the conjugacy classes with trivial action on G/K is $|\mathcal{Z}_\Gamma(G)| Z_I(iT)$.

• *Conjugacy classes with non-trivial action*

For a representative $\gamma = e^{\mathfrak{r}}$ of an element in $\Gamma \backslash \mathcal{Z}_\Gamma(G)$ we choose the deformation

$$(102) \quad \begin{aligned} V &= - \int dt \frac{1}{2} \langle \partial_t^{A_\gamma} J, \partial_t^{A_\gamma} \psi \rangle, \\ \delta V &= \int dt \frac{1}{2} \langle \partial_t^{A_\gamma} J, \partial_t^{A_\gamma} J \rangle + \frac{i}{2} \langle \partial_t^{A_\gamma} \psi, \partial_t^J \partial_t^{A_\gamma} \psi \rangle, \end{aligned}$$

where A_γ is a regular element of $\mathfrak{t}^\gamma \equiv \mathfrak{h}_\gamma^c \cap \mathfrak{k}$ where \mathfrak{h}_γ^c is the maximally compact Cartan subalgebra of the centralizer group $Z_\mathfrak{k}(\mathfrak{r}) \subset G$. In this case, we choose \mathfrak{k} and \mathfrak{h}^c such that $\mathfrak{h}_\gamma^c \subset \mathfrak{h}^c$.

Note that we additionally introduced A_γ in the deformation to ensure that the path integral localizes onto isolated critical points. Explicitly, the points in the localization locus are given by

$$(103) \quad g_{\gamma, \{n_j\}}(t) = \exp\left(\frac{(r + n_j \mathfrak{z}_j^\gamma)t}{T}\right) g_0,$$

where \mathfrak{z}_j^γ is a basis of $\Gamma_K^\gamma \equiv \Gamma_K \cap \mathfrak{t}^\gamma$. Thus for each non-trivial conjugacy class $[\gamma]$ we have the following contribution to the pre-trace formula (97),

$$(104) \quad \begin{aligned} Z_{[\gamma]}(iT) = & \frac{\text{vol}(\Gamma_\gamma \backslash G_\gamma)}{(2\pi iT)^{d_p/2} (-2\pi iT)^{d_t/2} |W_K| \text{vol}(T_K)} \int_{T_K} \left(\prod_{\alpha \in R_{\mathfrak{g}_\mathbb{C}; \mathfrak{h}_\mathbb{C}}^+} -i \langle \mathbf{h}, \alpha \rangle \right) \frac{|\delta(\mathbf{t})|^2}{\tau(\mathbf{h})} d\mathbf{t} \\ & \times \sum_{\mathfrak{z}^\gamma \in \Gamma_K^\gamma} e^{\frac{i(\langle \mathbf{h}, \mathbf{h} \rangle + \langle \mathbf{r} + \mathfrak{z}^\gamma, \mathbf{r} + \mathfrak{z}^\gamma \rangle)}{2T}} \left(\prod_{\alpha \in R_{\mathfrak{g}_\mathbb{C}; \mathfrak{h}_\mathbb{C}}^+} \frac{\langle \mathbf{r} + \mathfrak{z}^\gamma, \alpha \rangle}{2i \sinh \frac{\langle \mathbf{r} + \mathfrak{z}^\gamma, \alpha \rangle}{2}} \right) \int_{G_\gamma \backslash G} e^{-\frac{i \langle \text{Ad}_g \mathbf{h}, \mathbf{r} + \mathfrak{z}^\gamma \rangle}{T}} dg. \end{aligned}$$

As in Section 4.4, we formally interchange the integration over h and the summation over the set of critical points, and bypass the singularities in the integration over h by choosing Feynman like r_G -dimensional contour, which is consistent with the Weyl’s law.

5.2. Selberg trace formula for compact hyperbolic 3-manifold

Here we consider another rank 1 example when $X = \Gamma \backslash G/K$ is a compact hyperbolic 3-manifold. It corresponds to the case $G = \text{SL}(2, \mathbb{C})$ and $K = \text{SU}(2)$, so $G/K = \mathbb{H}^3$ is the Lobachevsky space, and Γ is a cocompact purely loxodromic Kleinian group. To simplify notations, here we consider the case $-I \in \Gamma$ and hence $\mathcal{Z}_\Gamma(G) = \{I, -I\}$. Each element in Γ that does not belong to $\mathcal{Z}_\Gamma(G)$ is loxodromic, and by definition it is conjugate to a unique element in $\text{SL}(2, \mathbb{C})$ of the form

$$(105) \quad \begin{pmatrix} a_\gamma & 0 \\ 0 & a_\gamma^{-1} \end{pmatrix}, \quad |a_\gamma| > 1,$$

and we denote $a_\gamma = e^{r+ib}$ where $r > 0$ and $b \in [0, 2\pi)$. Since $\mathcal{Z}_\Gamma(G) = \{I, -I\}$, the elements with (r, b) and $(r, b + \pi)$ act identically on \mathbb{H}^3 .

Firstly, we apply (101) to compute the contribution from the trivial elements which in our case $[\gamma] = I$ or $[\gamma] = -I$. We choose to parametrize

$T_K \simeq U(1)$ by $\mathbf{t} = e^{\mathbf{h}}$ with $\mathbf{h} = h \cdot i\sigma_3$ where $h \in [0, 2\pi)$. Regarding $\mathfrak{g} \simeq \mathfrak{sl}(2, \mathbb{C}) \simeq \mathfrak{so}(3, 1)$ as a real semisimple Lie algebra, we have $\mathfrak{g}_{\mathbb{C}} \simeq \mathfrak{so}(3, 1)_{\mathbb{C}} \simeq \mathfrak{so}(4, \mathbb{C})$, and an elementary computation gives

$$(106) \quad |\delta(\mathbf{t})|^2 = 4 \sin^2 h \quad \text{and} \quad \tau(\mathbf{h}) = -4 \sin^2 h.$$

As in Section 4.4, we can combine the sum over Γ_K and the integral over T_K to a single integral over $\mathbb{R}^{r_K} = \mathbb{R}$, and we obtain

$$(107) \quad Z_I(iT) = \frac{\text{vol}(\Gamma \backslash G) e^{-\frac{iT}{8}}}{(2\pi T)^3 \pi} \int_{\mathbb{R}} dh h^2 e^{-\frac{4ih^2}{T}}.$$

Analytic continuation and evaluation of the resulting integral gives

$$(108) \quad Z_I(\beta) = \text{vol}(\Gamma \backslash G) e^{-\frac{\beta}{8}} \frac{1}{128\pi^{7/2} \beta^{3/2}},$$

where we emphasize that the volume is in terms of the Riemannian measure of the Cartan-Killing metric on G defined in (3). Since $\Gamma \cap K = \mathcal{Z}_{\Gamma}(G) = \{I, -I\}$ the relation between volumes (52) still holds in this case, and using $\text{vol}(K) = 32\sqrt{2}\pi^2$, we have

$$(109) \quad Z_I(\beta) = \frac{\text{vol}(X) e^{-\frac{\beta}{8}}}{2(2\pi\beta)^{3/2}}$$

Secondly, we apply (104) to compute the contribution from the loxodromic elements. Each $\gamma \in \Gamma$ in a form (105) is conjugated to $e^{\mathbf{r}}$, where $\mathbf{r} = (r + ib)\sigma_3$, and the critical points are given by

$$(110) \quad g_{\gamma, n} = \begin{pmatrix} e^{\frac{r+i(b+2\pi n)}{T}} & 0 \\ 0 & e^{-\frac{r+i(b+2\pi n)}{T}} \end{pmatrix}, \quad n \in \mathbb{Z}.$$

In other words, $\Gamma_K^{\gamma} \simeq \mathbb{Z}$ and $\mathbf{r} = 2\pi n i\sigma_3$, where $n \in \mathbb{Z}$. Therefore, formula (104) becomes

$$(111) \quad \begin{aligned} & Z_{[\gamma]}(iT) \\ &= \frac{\text{vol}(\Gamma_{\gamma} \backslash G_{\gamma}) e^{-\frac{iT}{8}}}{8\pi^4 T^3} \int_{-\pi}^{\pi} dh \sum_{n \in \mathbb{Z}} e^{\frac{8i(r^2 - (b+2\pi n)^2 - h^2)}{T}} h^2 \frac{(r^2 + (b + 2\pi n))^2}{|\sinh^2(r + ib)|^2} \\ & \quad \times \int_{G_{\gamma} \backslash G} e^{-\frac{i(\text{Ad}_{\mathbf{g}} \mathbf{h}, \mathbf{r} + \mathbf{r}^{\gamma})}{T}} dg. \end{aligned}$$

To compute the orbital integral, we use the Iwasawa decomposition $G = ANK$ in the form

$$(112) \quad g = e^{aT_1} \exp \begin{pmatrix} 1 & (n_r + in_i) \\ 0 & 1 \end{pmatrix} e^{i\phi\sigma_3} e^{i\theta\sigma_2} e^{i\psi\sigma_3}$$

where $\phi \in [0, \pi)$, $\theta \in [0, \pi/2)$, $\psi \in [0, 2\pi)$ are the standard Euler angles for $SU(2)$. The corresponding Haar measure on G with respect to the Cartan-Killing metric (see Remark 1) is

$$(113) \quad dg = c_G da dn_r dn_i d\phi d\theta d\psi, \quad c_G = 2^{10} \sin(2\theta).$$

To determine the measure of $G_\gamma \backslash G$ where $G_\gamma = \{e^{(a+ib)\sigma_3} \mid a \in \mathbb{R}, b \in [0, 2\pi)\}$, we use the property of the integrand in the orbital integral is invariant under the transformation $g \rightarrow fg$ for any $f \in G_\gamma$. As in Section 4.4, we have a freedom to choose an overall normalization of the Haar measure on G_γ which will not affect the final result. Here we choose $df = \frac{1}{2\pi} da db$, and therefore the measure of the orbital integral over $G_\gamma \backslash G$ effectively becomes $c_G dn_r dn_i d\phi d\theta d\psi$.

The integration over n_r and n_i gives delta functions and the orbital integral becomes

$$(114) \quad \int_{G_\gamma \backslash G} e^{-\frac{i(\text{Ad}_g \mathbf{h}, \mathbf{r} + \mathbf{r} \gamma)}{T}} dg = \frac{(2\pi)^3 T^2}{h^2(r^2 + (b + 2\pi n)^2)} (e^{i16(b+2\pi n)h} + e^{-i16(b+2\pi n)h}).$$

Interchanging the sum over n and the integral over h in (111), we get

$$(115) \quad \begin{aligned} & Z_{[\gamma]}(iT) \\ &= \frac{\text{vol}(\Gamma_\gamma \backslash G_\gamma) e^{-\frac{iT}{8}}}{\pi T} \int_{\mathbb{R}} dh \left(e^{\frac{8i(r^2 - (b+h)^2)}{T}} + e^{\frac{8i(r^2 - (b-h)^2)}{T}} \right) \frac{1}{|\sinh(r+ib)|^2} \\ &= r_0 e^{-\frac{iT}{8}} \frac{1}{2\sqrt{2\pi iT}} e^{\frac{8ir^2}{T}} \frac{1}{|\sinh(r+ib)|^2}, \end{aligned}$$

where we used $\text{vol}(\Gamma_\gamma \backslash G_\gamma) = r_0/2$ where r_0 corresponds to a primitive element γ_0 of γ with $a_{\gamma_0} = e^{r_0 + ib_0}$.

Finally, upon the analytic continuation we get

$$(116) \quad Z_{[\gamma]}(\beta) = \frac{1}{2\sqrt{2\pi\beta}} \frac{r_0}{|\sinh(r+ib)|^2} e^{-\frac{8r^2}{\beta} - \frac{\beta}{8}}$$

Therefore, we obtain the Selberg trace formula on a compact hyperbolic 3-manifold,

$$(117) \quad \text{Tr}[e^{-\frac{\beta\Delta}{2}}] = \frac{\text{vol}(X)e^{-\frac{\beta}{8}}}{(2\pi\beta)^{3/2}} + \sum_{\substack{[\gamma] \in \text{loxodromic} \\ 0 \leq \arg(a_\gamma) < \pi}} \frac{1}{\sqrt{2\pi\beta}} \frac{r_0}{|\sinh(r+ib)|^2} e^{-\frac{8r^2}{\beta} - \frac{\beta}{8}},$$

where for loxodromic elements we simplified the domain to be $0 \leq \arg(a_\gamma) < \pi$ using the property $Z_{[\gamma]} = Z_{[-\gamma]}$. After rescaling the Riemannian metric on \mathbb{H}^3 to have the scalar curvature -1 , we obtain the Selberg trace formula in the form used in the literature [27] (see also [28]).

Appendix A. Laplace operators on $\text{SL}(2, \mathbb{R})$ and $\text{SL}(2, \mathbb{R})/\text{SO}(2)$

Let $T_1 = \sigma_3, T_2 = \sigma_1, T_3 = i\sigma_2$ be the basis of the Lie algebra $\mathfrak{g} = \text{sl}(2, \mathbb{R})$, so that \mathfrak{p} is generated by T_1, T_2 and \mathfrak{k} — by T_3 . In this basis the Killing metric

$$g_{ab} = \langle T_a, T_b \rangle = 4 \text{Tr } T_a T_b$$

is diagonal and

$$(A.1) \quad g_{11} = g_{22} = 8, \quad g_{33} = -8.$$

The Killing form on \mathfrak{g} naturally induces the Cartan-Killing metric on G as $ds^2 = 4 \text{Tr}(g^{-1}dg g^{-1}dg)$. One natural coordinates (x, y, θ) on $G = \text{SL}(2, \mathbb{R}) = \mathbb{R} \times \mathbb{R}_{>0} \times S^1$ are given by the Iwasawa decomposition $G = ANK$,

$$g = \begin{pmatrix} \sqrt{y} & 0 \\ 0 & \frac{1}{\sqrt{y}} \end{pmatrix} \begin{pmatrix} 1 & \frac{x}{y} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}.$$

In the path integral computations in the main text, we need to use a volume form dg for the pseudo-Riemannian metric on G , defined by the Cartan-Killing form. Using the Iwasawa decomposition and (A.1), we obtain

$$(A.2) \quad dg = 4\sqrt{2} \frac{dx dy d\theta}{y^2}.$$

Correspondingly, the Riemannian volume form on $K < G$ is $dk = 2\sqrt{2}d\theta$, so that the induced volume form on $\mathbb{H} = G/K$ is twice the hyperbolic area

form

$$d\mu = \frac{dx dy}{y^2}.$$

In particular, for $X = \Gamma \backslash G / K$ we have $\mu(X) = \text{vol}(X) / 2$.

Defining left-invariant vector fields on G by

$$(\hat{X}f)(g) = \left. \frac{d}{dt} \right|_{t=0} f(ge^{tX}), \quad X \in \mathfrak{g},$$

we have

$$\begin{aligned} \hat{T}_1 &= -2y \sin 2\theta \frac{\partial}{\partial x} + 2y \cos 2\theta \frac{\partial}{\partial y} + \sin 2\theta \frac{\partial}{\partial \theta}, \\ \hat{T}_2 &= 2y \cos 2\theta \frac{\partial}{\partial x} + 2y \sin 2\theta \frac{\partial}{\partial y} - \cos 2\theta \frac{\partial}{\partial \theta}, \\ \hat{T}_3 &= \frac{\partial}{\partial \theta}, \end{aligned}$$

so

$$\Delta = -\frac{1}{8}(\hat{T}_1^2 + \hat{T}_2^2 - \hat{T}_3^2) = -\frac{y^2}{2} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) + \frac{y}{2} \frac{\partial^2}{\partial x \partial \theta}.$$

Note that operator Δ is not elliptic.

Restriction of 2Δ to the subspace of functions f on G satisfying

$$f(ge^{\theta T_3}) = e^{in\theta} f(g), \quad n = 2k \in \mathbb{Z},$$

is the so-called Maass Laplacian of weight k^{13} , elliptic operator $D_k/2$, where

$$D_k f = -y^2 \left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \right) + 2iky \frac{\partial f}{\partial x}.$$

Operator D_k acts the Hilbert space $L_k^2(X)$ of functions $f(z)$ on \mathbb{H} , satisfying

$$f \left(\frac{az + b}{cz + d} \right) \frac{|cz + d|^{2k}}{(cz + d)^{2k}} = f(z), \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$$

and square integrable over the fundamental domain for Γ with respect to hyperbolic area element $y^{-2} dx dy$. When k is half-integer ($n = 2k$ is odd),

¹³We parameterized Laplacians by $k \in \frac{1}{2}\mathbb{Z}$ instead of $n = 2k \in \mathbb{Z}$ since they effectively act on k -differentials on X .

then we assume Γ does not contain $-I^{14}$, or twist with a one-dimensional representation π of Γ such that $\pi(-I) = -1$.

One also considers Hodge Laplacian, the $\bar{\partial}$ -Laplace operator

$$\Delta_k = -4y^{2-2k} \frac{\partial}{\partial z} y^{2k} \frac{\partial}{\partial \bar{z}},$$

acting on the Hilbert space $\mathcal{H}_{2k}(\Gamma)$ of automorphic forms of weight $2k$ for group Γ , functions $f(z)$ satisfying

$$(A.3) \quad f\left(\frac{az+b}{cz+d}\right) = (cz+d)^{2k} f(z), \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$$

and square-integrable with respect to the measure $y^{2k-2} dx dy$. When n is odd, the same conditions on Γ are imposed. Functions $f \in \mathcal{H}_{2k}(\Gamma)$ corresponds to k -differentials $f(z) dz^k$ on a Riemann surface $X = \Gamma \backslash \mathbb{H}$, spinors for $k = 1/2$.

The operators Δ_k are non-negative and the isometry

$$f(z) \mapsto y^k f(z)$$

between Hilbert spaces $\mathcal{H}_{2k}(\Gamma)$ and $L_k^2(X)$ conjugates Δ_k and $D_k + k(k-1)$, so they have the same spectrum.

Introducing Maass operators

$$K_k = (z - \bar{z}) \frac{\partial}{\partial z} + k, \quad L_k = -(z - \bar{z}) \frac{\partial}{\partial \bar{z}} - k,$$

we have

$$D_{k+1} K_k = K_k D_k, \quad D_{k-1} L_k = L_k D_k.$$

From these equations it follows that the spectra of operators D_k and D_{k+1} coincide, except for possible finitely many eigenvalues related to zero modes of Hodge Laplacians (see [29]).

¹⁴Note that in $X = \Gamma \backslash \mathbb{H}$ the Fuchsian group Γ is a subgroup of $\mathrm{PSL}(2, \mathbb{R})$, and there 2^{2g} its lifts to $\mathrm{SL}(2, \mathbb{R})$ that do not contain $-I$, provided all elliptic generators are of odd order.

References

- [1] M. C. Gutzwiller. *The Geometry of Quantum Chaos*. Physica Scripta **T9** (1985), 184–192.
- [2] C. Choi and L. A. Takhtajan. *Supersymmetry and trace formulas. Part I. Compact Lie groups*. JHEP **06** (2024) 026. Erratum: JHEP **02** (2025) 154. [arXiv:hep-th/2112.07942](https://arxiv.org/abs/hep-th/2112.07942).
- [3] A. Selberg. *Harmonic analysis and discontinuous groups in weakly symmetric riemannian spaces with applications to dirichlet series*. The Journal of the Indian Mathematical Society **20** (1956), 47–87.
- [4] J. J. Duistermaat and G. J. Heckman. *On the variation in the cohomology of the symplectic form of the reduced phase space*. Inventiones mathematicae **69** (1982), 259–268.
- [5] M. F. Atiyah and R. Bott. *The moment map and equivariant cohomology*. Topology **23** (1984), 1–28.
- [6] N. Berline and M. Vergne. *Zéros d'un champ de vecteurs et classes caractéristiques équivariantes*. Duke Mathematical Journal **50** (1983), 539–549.
- [7] E. Witten. *Analytic Continuation Of Chern-Simons Theory*. AMS/IP Stud. Adv. Math. **50** (2011), 347–446. [arXiv:hep-th/1001.2933](https://arxiv.org/abs/hep-th/1001.2933).
- [8] N. Krausz and M. S. Marinov. *Exact evolution operator on noncompact group manifolds*. J. Math. Phys. **41** (2000), 5180–5208. [arXiv:quant-ph/9709050](https://arxiv.org/abs/quant-ph/9709050).
- [9] A. Kirillov. Jr. *An introduction to Lie groups and Lie algebras*, Cambridge Studies in Advanced Mathematics, vol. 113, Cambridge University Press, Cambridge (2008).
- [10] M. F. Atiyah. *Circular symmetry and stationary-phase approximation*. Astérisque **131** (1985), 43–59.
- [11] M. F. Atiyah and A. N. Pressley. *Convexity and loop groups*. Arithmetic and geometry, Springer, (1983), 33–63.
- [12] R. F. Picken. *The Propagator for Quantum Mechanics on a Group Manifold From an Infinite Dimensional Analog of the Duistermaat-heckman Integration Formula*. J. Phys. A **22** (1989) 2285.

- [13] L. D. Faddeev. *Operator Anomaly for the Gauss Law*. Phys. Lett. B **145** (1984), 81–84.
- [14] L. D. Faddeev and Samson L. Shatashvili. *Algebraic and Hamiltonian Methods in the Theory of Nonabelian Anomalies*. Teor. Mat. Fiz. **60** (1984), 206–217.
- [15] Philip C. Nelson and Luis Alvarez-Gaumé. *Hamiltonian interpretation of anomalies*. Comm. Math. Phys. **99** (1985), 103–114.
- [16] W. A. Bardeen, S. Elitzur, Y. Frishman, and E. Rabinovici. *Fractional charges: global and local aspects*. Nucl. Phys. B **218** (1983), 445–458.
- [17] D. Delmastro, J. Gomis, P.-S. Hsin, and Z. Komargodski. *Anomalies and Symmetry Fractionalization*. (2022). [arXiv:hep-th/2206.15118](https://arxiv.org/abs/hep-th/2206.15118).
- [18] Davide Gaiotto, Anton Kapustin, Nathan Seiberg, and Brian Willett. *Generalized global symmetries*. JHEP **02** (2015) 172.
- [19] H. Weyl. *Über die asymptotische verteilung der eigenwerte*. Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse **1911** (1911), 110–117.
- [20] D. A. Hejhal. The Selberg trace formula for $\mathrm{PSL}(2, R)$. Vol. I, Lecture Notes in Mathematics, Vol. 548, Springer-Verlag, Berlin-New York, (1976).
- [21] N. P. Landsman. *Quantization and Superselection Sectors. 1. Transformation Group C^* Algebras*. Rev. Math. Phys. **2** (1990), 45–72.
- [22] M. G. G. Laidlaw and C. M. DeWitt. *Feynman functional integrals for systems of indistinguishable particles*. Phys. Rev. D **3** (1971), 1375–1378.
- [23] L. Schulman. *A Path integral for spin*. Phys. Rev. **176** (1968), 1558–1569.
- [24] J. S. Dowker. *Quantum mechanics and field theory on multiply connected and on homogeneous spaces*. J. Phys. A **5** (1972), 936–943.
- [25] R. Camporesi. *Harmonic analysis and propagators on homogeneous spaces*. Phys. Rept. **196** (1990), 1–134.
- [26] A. W. Knap. Lie Groups beyond an Introduction. Progress in Mathematics, vol. 140, Birkhäuser Boston, Inc., Boston, MA, (1996).

- [27] Jü. Elstrodt, F. Grunewald, and J. Mennicke. *Groups acting on hyperbolic space: Harmonic analysis and number theory*. Springer Science & Business Media, (2013).
- [28] J. S Friedman. *The selberg trace formula and selberg zeta-function for cofinite kleinian groups with finite-dimensional unitary representations*. *Mathematische Zeitschrift* **250** (2005), 939–965. [arXiv:math/0410067](https://arxiv.org/abs/math/0410067).
- [29] J. D. Fay. *Fourier coefficients of the resolvent for a fuchsian group*. (1977).

PERIMETER INSTITUTE FOR THEORETICAL PHYSICS
WATERLOO, N2L 2Y5, CANADA
E-mail address: changhachoi@gmail.com

DEPARTMENT OF MATHEMATICS, STONY BROOK UNIVERSITY
STONY BROOK, NY 11794, USA
EULER INTERNATIONAL MATHEMATICAL INSTITUTE, PESOCHNAYA NAB. 10
SAINT PETERSBURG, 197022, RUSSIA
E-mail address: leontak@math.stonybrook.edu

RECEIVED JANUARY 15, 2024

ACCEPTED MAY 13, 2025