J. Nonlinear Var. Anal. 7 (2023), No. 3, pp. 309-343
Available online at http://jnva.biemdas.com
https://doi.org/10.23952/jnva.7.2023.3.01

HEAT KERNELS ON UNIT SPHERES AND APPLICATIONS TO GRAPH KERNELS

DER-CHEN CHANG[1], OPHIR FRIEDER[2], CHI-FENG HUNG[3], HAO-REN YAO[4,*]

[1]*Department of Mathematics & Statistics, Georgetown University, Washington, DC 20057, USA & Graduate Institute of Business Administration, Fu Jen Catholic University, New Taipei City 24205, Taiwan*
[2]*Department of Computer Science, Georgetown University and Department of Biostatistics, Bioinformatics & Biomathematics, Georgetown University Medical Center, Washington, DC 20057, USA*
[3]*School of Medicine, Fu Jen Catholic University, New Taipei City, 24205, Taiwan*
[4]*Clinical Center, National Institute of Health, Bethesda, Maryland, 20892, USA*

Abstract. It is known that many statistical and machine learning approaches heavily rely on pairwise distance between data points. The choice of distance function on the underlying manifold has a fundamental impact on performance of these processes. This is closely related to questions of how to appropriately calculate distances, and hence, fundamental solutions (heat kernels) for heat operators can be obtained. In general, it is not so easy to obtain a closed form for heat kernels. We first survey results of heat kernels on radially symmetric Riemannian manifolds, e.g., Euclidean spaces and unit spheres in $\mathbf{R}^n$. For the cases $n = 1,2,3$, we may construct the heat kernel explicitly. But, the computation is much more complicated when $n > 3$. However, by results of Nagase, we may construct parametrices for the heat kernel by using elementary functions so that the error terms can be under controlled. In the second part of the paper, we discuss some results on subRiemannian manifolds, especially 3-dimensional sphere in $\mathbf{C}^2$ as a CR-manifold. We study geodesics connecting two given points on $\mathbb{S}^3$ respecting the Hopf fibration. This geodesic boundary value problem is completely solved in the case of $\mathbb{S}^3$ and some partial results are obtained in the general case. The Carnot-Carathéodory distance is calculated. We also present some motivations related to quantum mechanics. Then we give a brief discussion of Greiner's methods on the heat kernel for the Cauchy-Riemann subLaplacian on $\mathbb{S}^{2n+1}$. We provide a brief discussion on applications of these heat kernels to graph kernels in the last part of the paper.

Keywords. Graph kernels; Heat kernel; Riemannian manifold; subRiemannian manifold; Unit sphere.

1. Introduction

It is known that one of the most general representations of a data set $\{x_j\}_{j=1}^m$, for $x_j \in \mathscr{M} \subset \mathbf{R}^N$, is via a matrix of pairwise distance between each of the data points. The most standard choice in practice is the Euclidean metric. Basically, this choice implicitly assumes that the data do not have any interesting nontrivial geometric structure in their underlying manifold. In fact, we have seen examples that there exist pairs of points in $\mathscr{M}$ which are close in Euclidean

*Corresponding author.
E-mail addresses: chang@georgetown.edu (D.-C. Chang), frieder@georgetown.edu (O. Frieder), skin@mail.fju.edu.tw (C.-F. Hung), hao-ren.yao@nih.gov (H.-R. Yao)
Received March 3, 2023; Accepted April 12, 2023.

metric but actually far away from each other if one needs to travel along curves with constraints. More precisely, the points along a path that does not cross empty regions across which there is no data but instead follows the "flow" of the data.

The shortest path between two points A and B on a manifold $\mathscr{M}$ is known as the geodesic, with the length of this path corresponding to the geodesic distance, When A and B are very close to each other, then the Euclidean distance provides an accurate approximation to the geodesic distance. However, Euclidean and geodesic distance can be dramatically different when the curvature of $\mathscr{M}$ is large, The accuracy of Euclidean distance in small neighborhoods has exploited to develop algorithms for approximating geodesics distances via graph distances. Such method provides a weighted graph in which edges connect neighbors and weights correspond to the Euclidean distance. Readers can consult the papers by the authors *e.g.,* [36, 37, 38].

Furthermore, using Euclidean metric within local neighborhoods, we need to keep neighborhoods small to control the global approximation error. This creates some shortcomings. For example, when the sample size m is not large enough and when the density d of the data points is not uniform over $\mathscr{M}$ but instead is larger in certain regions than others. The main purpose of this paper is to provide a good strategy for more accurate geodesic distance estimations that improve the local Euclidean approximation while continuing to reply on graph distance algorithms. More precise, we study the spherical model which has the almost unique features of both accounting for positive curvature and having the geodesic distance between any two points in a simple closed form. Here we also recommend a beautiful paper to readers by Li and Dunson [23] for motivation of this subject. Unlike the paper in [23], here we give detailed discussions for elliptic and sub-elliptic heat kernels on the sphere $\mathbb{S}^3 \subset \mathbf{C}^2$.

Let $(\mathscr{M}, g)$ be a Riemannian manifold and let $C^{1,2}(\mathscr{M})$ be the space of functions

$$f : (0,\infty) \times \mathscr{M} \to \mathbf{R},$$

which are continuous on $[0,\infty) \times \mathbf{R}$, C^1-differentiable in the first variable, and C^2-differentiable in the second variable. As usual, define the Laplacian be

$$\Delta = -\mathrm{div}\nabla.$$

Definition 1.1. The operator $Q = \frac{\partial}{\partial t} + \Delta$ defined on the space $C^{1,2}(\mathscr{M})$ is called the *heat operator* on $(\mathscr{M}, g)$.

To invert the heat operator, one needs to study the fundamental solution, *i.e.,* the heat kernel.

Definition 1.2. A fundamental solution K for the heat operator $Q = \frac{\partial}{\partial t} + \Delta_y$ is a function P : $\mathscr{M} \times \mathscr{M} \times (0,\infty) \to \mathbf{R}$ with the following properties:
(1). $K \in C\big(M \times \mathscr{M} \times (0,\infty)\big)$, C^2 in the first variable, and C^1 in the second variable,
(2). $\left(\frac{\partial}{\partial t} + \Delta_y\right) K(\cdot, y, t) = 0$ for all $t > 0$,
(3). $\lim_{t\to 0^+} K(x, \cdot, t) = \delta_x$ for all $x \in \mathscr{M}$.

where δ_x is the Dirac distribution concentrated at x and the limit (3) is considered in the sense of distribution, *i.e.,*

$$\lim_{t\to 0^+} \int_{\mathscr{M}} K(x,y,t)\psi(x)\,dv(x) = \psi(y), \quad \text{for all} \quad \psi \in C_0^\infty(\mathscr{M}), \quad x \in \mathscr{M},$$

where $C_0^\infty(\mathscr{M})$ denotes the collection of smooth functions with compact support, and $dv(x) = \sqrt{|g_{jk}(x)|}dx_1 \wedge \cdots \wedge dx_n$.

1.1. **Compact manifolds.** Let $(\mathscr{M}, g)$ be a compact Riemannian manifold. We define the inner product

$$\langle f, g \rangle_g = \int_{\mathscr{M}} fg\, dv, \quad \text{for all} \quad f, g \in C\infty(\mathscr{M}).$$

Denote $\|f\|_{L^2} = \langle f, f \rangle_g^{\frac{1}{2}}$. The space $L^2(\mathscr{M})$ is obtained from $C^\infty(\mathscr{M})$ by completeness with respect to the norm $\|\cdot\|_{L^2}$.

The real numbers λ for which there is an eigenfunction of f such that $\Delta f = \lambda f$ are called *eigenvalues*. The function f is an eigenfunction of λ. Let

$$V_\lambda(\mathscr{M}, g) = \{ f : \mathscr{M} \to \mathbf{R} : \Delta f = \lambda f \}$$

be the eigenspace with respect to λ. The number $m_\lambda = \dim V_\lambda(\mathscr{M}, g)$ is called the multiplicity of λ.

In the following we shall find the fundamental solution of Q in the case of a compact Riemannian manifold. The spectral theory of the Laplace operator is a consequence of the Riesz-Schauder theory. Hence the following spectrum theorem holds for the Laplace operator on Riemannian manifolds.

Theorem 1.1. *We have the following results.*
(1). *The eigenvalues are nonnegative and form a countable infinite set*

$$0 = \lambda_0 < \lambda_1 < \lambda_2 < \lambda_3 < \cdots,$$

with $\lambda_k \to +\infty$ ad $k \to \infty$ and the series $\sum_{k=1}^{\infty} \frac{1}{\lambda_k^2}$ converges.
(2). *Each eigenvalue λ_k has finite multiplicity m_k. The eigenspaces $V_{\lambda_k}(\mathscr{M}, g)$ and $V_{\lambda_j}(\mathscr{M}, g)$, $k \neq j$ are orthogonal with respect to the inner product $\langle \cdot, \cdot \rangle_g$.*
(3). *From the system of eigenfunctions, using the Gram-Schmidt procedure, one may obtain a complete orthonormal system $\{f_{k_\ell} : k \in \mathbf{N}, \ell = 1, \ldots, m_k\}$ of eigenfunctions, such that*

$$h = \sum_{k=0}^{\infty} \sum_{\ell=1}^{m_k} a_{k\ell} f_{k\ell}, \quad \textit{for all} \quad h \in L^2(\mathscr{M}),$$

with $a_{k\ell} = \langle h, f_{k\ell} \rangle_g$. In particular, the Parseval identity holds

$$\|h\|_{L^2}^2 = \sum_{k=0}^{\infty} \sum_{\ell=1}^{m_k} \langle h, f_{k\ell} \rangle_g^2.$$

By results of Calin-Chang [8] and Mazet–Berger–Gauduchon [29], we may assume that the fundamental solution for the heat operator exists. The following result provides a formula for the fundamental solution on a compact Riemannian manifolds.

Theorem 1.2. *Let $\{f_j : j \in \mathbf{N}\}$ be a complete orthonormal system of eigenfunctions for the Laplace operator on the compact Riemannian manifolds $(\mathscr{M}, g)$, such that*

$$\lambda_0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots.$$

Then the fundamental solution is given by

$$K(x, y, t) = \sum_{j=0}^{\infty} e^{-\lambda_j t} f_j(x) f_j(y). \tag{1.1}$$

Proof. Since the system $\{f_j : j \in \mathbf{N}\}$ is an orthonormal basis of the Hilbert space $L^2(\mathscr{M})$, we assume the existence of a fundamental solution for fixed x and t. Therefore,

$$K(x,\cdot,t) = \sum_{j=0}^{\infty} \omega_j(x,t) f_j,$$

where

$$\omega_j(x,t) = \int_{\mathscr{M}} K(x,y,t) f_j(y)\, dv(y).$$

Differentiating with respect to t yields

$$\begin{aligned}\frac{\partial \omega_j}{\partial t} &= \int_{\mathscr{M}} \frac{\partial K}{\partial t}(x,y,t) f_j(y)\, dv(y) = \left\langle \frac{\partial K}{\partial t}, f_j \right\rangle_g \\ &= -\left\langle \Delta_y K, f_j \right\rangle_g = -\left\langle K, \Delta_y f_j \right\rangle_g = -\lambda_j \left\langle K, f_j \right\rangle_g = -\lambda_j \omega_j.\end{aligned}$$

Hence, $\frac{\partial \omega_j}{\partial t} = -\lambda_j \omega_j$, where $\omega_j(x,t) = \alpha_j(x) e^{-\lambda_j t}$. The function $\alpha_j(x)$ satisfies

$$\begin{aligned}\lim_{t\to 0^+} \omega_j(x,t) &= \lim_{t\to 0^+} \int_{\mathscr{M}} K(x,y,t) f_j(y)\, dv(y) \\ &= \int_{\mathscr{M}} \delta_x(y) f_j(y)\, dv(y) = f_j(x).\end{aligned}$$

On the left hand side, one has

$$\lim_{t\to 0^+} \omega_j(x,t) = \alpha_j(x),$$

and hence $\alpha_j(x) = f_j(x)$. Thus, the equation (1.1) is proved. □

It is extremely interesting to solve the heat operator with initial data. In other words, given a continuous function $h \in C(\mathscr{M})$, find a function $f \in C^{1,2}(\mathscr{M})$ such that

$$\begin{aligned}\Big(\frac{\partial}{\partial t} + \Delta_y\Big) f &= 0 \quad \text{for all} \quad t > 0, \\ \lim_{t\to 0^+} f(x,t) &= h(x) \quad \text{for all} \quad x \in \mathscr{M}.\end{aligned} \tag{1.2}$$

Indeed, we have the following theorem.

Proposition 1.1. *The solution for the above* (1)-(2) *initial value problem is given by the formula*

$$f(x,t) = \int_{\mathscr{M}} K(x,y,t) h(y)\, dv(y), \tag{1.3}$$

where $K(x,y,t)$ is given by (1.1).

Proof. We may prove this proposition by straightforward computation. More precisely,

$$\begin{aligned}\frac{\partial f}{\partial t}(x,t) &= \frac{\partial}{\partial t} \int_{\mathscr{M}} \sum_{j=0}^{\infty} e^{-\lambda_j t} f_j(x) f_j(y) h(y)\, dv(y) \\ &= -\int_{\mathscr{M}} \sum_{j=0}^{\infty} \lambda_j e^{-\lambda_j t} f_j(x) f_j(y) h(y)\, dv(y).\end{aligned}$$

Similarly,

$$\begin{aligned}
\Delta_x f(x,t) &= \Delta_x \int_{\mathscr{M}} \sum_{j=0}^{\infty} e^{-\lambda_j t} f_j(x) f_j(y) h(y)\, dv(y) \\
&= \int_{\mathscr{M}} \sum_{j=0}^{\infty} e^{-\lambda_j t} \Big[\Delta_x f_j(x)\Big] f_j(y) h(y)\, dv(y) \\
&= \int_{\mathscr{M}} \sum_{j=0}^{\infty} \lambda_j e^{-\lambda_j t} f_j(x) f_j(y) h(y)\, dv(y).
\end{aligned}$$

Hence,

$$\Big(\frac{\partial}{\partial t} + \Delta_y\Big) f = 0.$$

We still need to show that $\lim_{t\to 0^+} f(x,t) = h(x)$. Using Property (3) in Definition (1.2) yields

$$\begin{aligned}
\lim_{t\to 0^+} f(x,t) &= \lim_{t\to 0^+} \int_{\mathscr{M}} K(x,y,t) h(y)\, dv(y) \\
&= \int_{\mathscr{M}} \Big[\lim_{t\to 0^+} K(x,y,t)\Big] h(y)\, dv(y) \\
&= \int_{\mathscr{M}} \delta_x(y) h(y)\, dv(y) \\
&= \big\langle \delta_x, h \big\rangle_g = h(x).
\end{aligned}$$

The proof of this proposition is therefore complete. □

2. Heat Kernel on Radially Symmetric Spaces

Definition 2.1. A Riemannian manifold $(\mathscr{M}, g)$ is called *radially symmetric* if for any $x_0 \in \mathscr{M}$, the geodesic sphere $\mathbb{S}(x_0, r)$ centered at x_0 with radius r has constant scalar mean curvature. Here the geodesic sphere centered at x_0 with radius r is defined by

$$\mathbb{S}(x_0, r) = \big\{\gamma(r) :\ \gamma : [0, r(x_0)) \to \mathscr{M},\ \gamma(0) = x_0, \gamma \text{ unit speed geodesic}\big\},$$

with $0 < r < r(x_0)$.

As the geodesics are locally length minimizing curves, the Riemannian distance is measured along the geodesics and it is equal to the arc length parameter s,

$$d(x_0, \gamma(s)) = \text{length}(\gamma) = s.$$

Hence the geodesic sphere can be written as

$$\mathbb{S}(x_0, s) = \big\{x \in \mathscr{M} :\ d(x_0, x) = s\big\}.$$

For example, the Euclidean space $\mathbf{R}^n$ with the standard metric is a radially symmetric space. It is known that the fundamental solution in this case is given by

$$K(x,y,t) = \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x-y|^2}{4t}}, \qquad t > 0. \tag{2.1}$$

This is a product between the volume function $v(t) = t^{-\frac{n}{2}}$ and an exponential with the exponent $-\frac{|x-y|^2}{4t} = -\frac{1}{2}S$ where S is the classical action between the points x and y within time t. We have the following theorem. For detailed discussion, see Calin-Chang [8].

Theorem 2.1. *Let $(\mathscr{M},g)$ be a radially symmetric space about the point $x_0 \in \mathscr{M}$. Then the fundamental solution for the heat operator is given by*

$$K(x_0,x,t) = CV(t)e^{-\frac{1}{2}S} = CV(t)e^{-\frac{d^2(x-x_0)^2}{4t}},$$

where $V(t)$ is the solution of

$$V'(t) = \frac{1}{2}\Delta S \cdot V(t)$$

with the condition $\lim_{t\to 0^+} t^{\frac{n}{2}}V(t) = 1$ and

$$C^{-1} = 2^n \int_0^\infty e^{-y^2}\zeta(x_0,y)\,dy,$$

with ζ define by the following equation:

$$vol\big[\mathbb{S}\big(x_0,2\sqrt{t}y\big)\big] \sim 2^n t^{\frac{n}{2}} \cdot \zeta(x_0,y), \qquad as \quad t\to 0^+.$$

Although we have Theorem 2.1, however, it is quite complicated to calculate an explicit fundamental solution for the heat operator on a n-dimensional Riemannian manifold. The formula (2.1) is a very special case which dealt with a manifold with zero curvature. In this case, the shortest distance between two points x and y is the length of the straight line segment between these two points (which is known as the Euclidean metric). In general, the shortest curve between two points x and y on a manifold $\mathscr{M}$ is known as the *geodesic*. The length of this curve corresponding to the geodesic distance. Now the problem reduces to see whether we may use Euclidean metric to approximate the geodesic distance. Intuitively, if x and y are very close to each other, then Euclidean metric gives an accurate approximation to the geodesic distance. Or, if $\mathscr{M}$ has very very low curvature and is close to flat globally. Otherwise, Euclidean and geodesic distance can be dramatically different. The accuracy of Euclidean metric in small neighborhood has been exploited to develop algorithms for approximating geodesic distances via graph distances. As we have seen before such approached define a weighted graph in which edges connect neighbors and weights correspond to the Euclidean metric. The approximated geodesic distance is the shortest curve on this graph (see [32, 36, 38]).

As we mentioned above, using Euclidean metric within local neighborhoods, we need to keep neighborhoods small to control the global approximation error. This creates problems when the density ρ of the data points is not uniform over $\mathscr{M}$ but instead is larger in certain regions than others. One obvious way is to improve the local Euclidean approximation while continuing to rely on graph distance algorithms. As we can expected, a better local approximation leads to better global approximation error. That is a reason we use another local distance estimator, which has the advantage of providing a simple and transparent modification of Euclidean metric to incorporate curvature. In this paper, we consider the curvature of $\mathscr{M}$ is positive. Then a good candidate to approximate the manifold $\mathscr{M}$ in a local neighborhood using spheres. Using spheres has the almost unique features of both accounting for positive curvature and having the geodesic distance between any two points in a simple closed form. Since we are interested in a smooth compact Riemannian manifold with Riemannian metric g, we need to concentrate on the model case when $\mathscr{M} = \mathbb{S}^n$.

2.1. **Heat kernel on $\mathbb{S}^1$.** We start with the simplest case $n = 1$. Under some convergence and regularity conditions on the function ψ, we have the following famous summation formula, called **Poisson's summation formula** (see Lawden [22]):

$$\sum_{n=-\infty}^{\infty} \psi(x+2n\pi) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} e^{ikx}\int_{-\infty}^{+\infty} e^{-iky}\psi(y)\,dy.$$

As a consequence, we have the next result regarding the theta-function:

$$\theta_3\big(x\big|\tau\big) = 1+2\sum_{n=1}^{\infty} e^{i\pi n^2\tau}\cos(2nx).$$

The result deals with the case of the θ_3-function, but similar transformation formulas work for the other theta-functions.

Lemma 2.1. *(Jacobi's transformation for θ_3)*
If $\tilde{t} = -\frac{1}{t}$, we have

$$\theta_3\big(z\big|\tilde{t}\big) = \sum_{n=-\infty}^{\infty} e^{\frac{itz^2}{t}}\theta_3\big(tz\big|t\big), \tag{2.2}$$

where the third Jacobi theta-function is defined by

$$\theta_3\big(z\big|t\big) = (-it)^{\frac{1}{2}} e^{i\pi n^2 t+2\pi z 2niz}. \tag{2.3}$$

Proof. In fact, we note that we may write the theta-function θ_3 as

$$\theta_3\big(x\big|t\big) = e^{\frac{z^2}{i\pi t}}\sum_{n=-\infty}^{\infty} e^{\frac{it}{4\pi}(u+2n\pi)^2}$$

with $u = \frac{2z}{t}$. Applying Poisson's summation formula with $f(u) = e^{\frac{itu^2}{4\pi}}$ yields

$$\begin{aligned}
\theta_3\big(x\big|t\big) &= \frac{1}{2\pi}e^{\frac{z^2}{i\pi t}}\sum_{k=-\infty}^{\infty} e^{iku}\int_{-\infty}^{+\infty} e^{\frac{itx^2}{4\pi}-iky}\,dy\\
&= \frac{1}{2\pi}e^{\frac{z^2}{i\pi t}}\sum_{k=-\infty}^{\infty} 2\pi\cdot e^{iku}(-it)^{-\frac{1}{2}}e^{-\frac{i\pi k^2}{t}}\\
&= (-it)^{-\frac{1}{2}}e^{\frac{z^2}{i\pi t}}\theta_3\big(\frac{z}{t}\big|-\frac{1}{t}\big).
\end{aligned}$$

Replacing z by tz yields (2.2) and we finish the proof of the lemma. □

Now we shall compute the heat kernel for the Laplacian on the unit circle $\mathbb{S}^1$ using Jacobi's transformation. If ω is the arc length on the unit circle, the eigenfunctions of $\frac{1}{2}\frac{d^2}{d\omega^2}$ on the circle $\mathbb{S}^1$ satisfy

$$\frac{1}{2}\frac{d^2}{d\omega^2}u_k = \lambda_k u_k,$$

where $\lambda_k = -\frac{k^2}{2}$ and $u_k(\omega) = c_k e^{ik\omega}$. The constant c_k can be computed from the orthonormality condition

$$1 = \int_{-\pi}^{+\pi} u_k(\omega)\overline{u_k(\omega)}\,d\omega = 2\pi c_k^2.$$

Hence, $c_k = \frac{1}{\sqrt{2\pi}}$. The functions $u_k(\omega) = \frac{e^{ik\omega}}{\sqrt{2\pi}}$ form a complete orthonormal system on $L^2\big(\mathbb{S}^1, d\omega\big)$. Thus, the heat kernel for $\frac{1}{2}\frac{d^2}{d\omega^2}$ on the circle $\mathbb{S}^1$ is given by

$$
\begin{aligned}
K_1(\omega_0, \omega; \tau) &= \sum_{k=0}^{\infty} e^{-\frac{k^2\tau}{2}} u_k(\omega)\bar{u}_k(\omega_0) = \frac{1}{2\pi}\sum_{k=0}^{\infty} e^{-\frac{k^2\tau}{2}} e^{ik(\omega-\omega_0)} \\
&= \frac{1}{2\pi}\sum_{k=0}^{\infty} e^{-\frac{k^2\tau}{2}+ik(\omega-\omega_0)} = \frac{1}{2\pi}\theta_3\Big(\frac{(\omega-\omega_0}{2}\Big|\frac{i\tau}{2\pi}\Big),
\end{aligned}
$$

by the definition of the theta-function θ_3 given in (2.3). Defining

$$
z = \frac{\omega-\omega_0}{2}, \qquad \tilde{t} = \frac{i\tau}{2\pi} = -\frac{1}{t}, \qquad t = \frac{2\pi i}{\tau}.
$$

Now the Jacobi's transformation formula provides

$$
\begin{aligned}
K_1(\omega_0, \omega; \tau) &= \frac{1}{2\pi}\theta_3\Big(\frac{\omega-\omega_0}{2}\Big|\frac{i\tau}{2\pi}\Big) = \frac{1}{2\pi}\theta_3\big(z\big|\tilde{t}\big) \\
&= \frac{1}{2\pi}(-it)^{\frac{1}{2}} e^{\frac{itz^2}{\pi}}\theta_3\big(tz\big|t\big) \\
&= \frac{1}{\sqrt{2\pi\tau}} e^{-\frac{(\omega-\omega_0)^2}{2\tau}}\theta_3\Big(\frac{i\pi}{\tau}(\omega-\omega_0)\Big|\frac{2\pi i}{\tau}\Big).
\end{aligned}
$$

The above formula is the same as the result that derived by Hamiltonian formalism.

2.2. **Heat kernel on $\mathbb{S}^2$.** Now we turn to the heat kernel for the Laplacian on the sphere $\mathbb{S}^2 = \big\{x \in \mathbf{R}^3 : \ |x| = 1\big\}$. In this case, the Laplacian is given in the spherical coordinates as follows:

$$
\Delta_{\mathbb{S}^2} = \frac{1}{\sin\phi_1}\frac{\partial}{\partial\phi_1}\Big(\sin\phi_1\cdot\frac{\partial}{\partial\phi_1}\Big) + \frac{1}{(\sin\phi_1)^2}\frac{\partial^2}{\partial\phi_2^2}
$$

with $\phi_1 \in (0,\pi)$, $\phi_2 \in [0,2\pi)$, The eigenfunctions are given by the spherical harmonics

$$
Y_m^k(\phi_1,\phi_2) = c_{m,k}P_m^{|k|}(\cos\phi_1)e^{ik\phi_2},
$$

where $m \in \mathbf{Z}_+$ and $k \in \{0, \pm 1, \pm 2, \dots, \pm m\}$. $P_m|k|$ denotes the Legendre function of order m which can be defined by the following generating formula

$$
\sum_{m=0}^{\infty} P_m(x)y^m = \frac{1}{\sqrt{1-2xy+y^2}}.
$$

Since

$$
\Delta_{\mathbb{S}^2}Y_m^k = -m(m+1)Y_m^k,
$$

the eigenvalue corresponding to Y_m^k is $\lambda_m^k = -m^2 - m$; see, for instance [20]. The heat kernel is given by

$$
\begin{aligned}
&K_2(\phi_1^{(0)}, \phi_2^{(0)}; \phi_1, \phi_2; t) \\
&= \sum e^{\lambda_m^k t} Y_m^k(\phi_1,\phi_2)\bar{Y}_m^k(\phi_1^{(0)}, \phi_2^{(0)}) \\
&= \sum_{m,k} c_{m,k}^2 e^{-m(m+1)t} e^{ik(\phi_2-\phi_2^{(0)})} P_m^k(\cos\phi_1)P_m^k(\cos\phi_1^{(0)}) \\
&= \sum_{m=0}^{\infty} e^{-m(m+1)t}\cdot\sum_{k=-m}^{m} c_{m,k}^2 e^{ik(\phi_2-\phi_2^{(0)})} P_m^k(\cos\phi_1)P_m^k(\cos\phi_1^{(0)}).
\end{aligned}
$$

A neater formulas can be obtained if the kernel is represented in terms of the angle ν between the points x_0 and x given by $\nu = \cos^{-1}\left(\langle x_0, x\rangle\right)$. Even if the inverse cosine function is multi-valued, since the heat kernel is an even, 2π-periodic function of ν, we may use any value of ν with $\cos\nu = \langle x_0, x\rangle$. Now the heat kernel can be written as

$$K_2(x_0,x;t) = \frac{e^{\frac{t}{4}}}{2^{\frac{3}{2}}\pi t}\int_\nu^\pi \frac{1}{\sqrt{\cos\nu-\cos\eta}}\cdot\sum_{n=-\infty}^{\infty}(-1)^n(\eta-2\pi n)e^{-\frac{1}{2t}(\nu-2\pi n)^2}\,d\eta.$$

Remark. We may study the heat kernel for Laplacian on the unit sphere $\mathbb{S}^3 = \left\{x\in\mathbf{R}^4:\ |x|=1\right\}$ also. In this case, using spherical coordinates, the Laplace-Beltrami operator can be written as

$$\Delta_{\mathbb{S}^3} = \frac{\partial^2}{\partial\phi_1^2}+\cot\phi_1\cdot\frac{\partial}{\partial\phi_1}+\frac{1}{(\sin\phi_1)^2}\left(\frac{\partial^2}{\partial\phi_1^2}+\frac{\partial^2}{\partial\phi_2^2}-2\cos\phi_1\cdot\frac{\partial^2}{\partial\phi_2\partial\phi_3}\right)$$

We may use a similar method to construct the heat for $\Delta_{\mathbb{S}^3}$. The kernel has the following form:

$$K_3(x_0,x;t) = \frac{1}{(4\pi t)^{\frac{3}{2}}}\frac{d}{\sin d}e^te^{-\frac{d^2}{4t}}, \tag{2.4}$$

where $d = d_{\mathbb{S}^3}(x_0,x)$ is the Riemannian metric on the sphere $\mathbb{S}^3$ which is induced by the Euclidean distance in $\mathbf{R}^4$ to the sphere. The kernel $K(x_0,x;t)$ in (2.4) was first computed by Schulman [31] when $\mathscr{M} = SU(2)$. He also conjectured that the formula (2.4) works in general for Lie groups. However, from results we calculated above for $\mathbb{S}^n$ when $n = 1,2,3$, it seems to us that the heat kernels for odd dimensional sphere have more compact form than even dimensional spheres. That is the goal for the next secion.

2.3. **Recurrence relations for parametrix of heat operator on $\mathbb{S}^n$.** In this case, the Laplace operator $\Delta = d^*d$ acting on functions on the standard n-sphere $(\mathbb{S}^n, g)$ with curvature 1 where d is the exterior differential and d^* is its formal adjoint. We consider the fundamental solution P_n associated to the heat equation with initial condition

$$\left(\frac{\partial}{\partial t}+\Delta\right)f = 0, \qquad f(x,0) = f_0.$$

Since the sphere is invariant under the rotation group, we know that the kernel depends only on t and the "distance" $d(x,y)$ of two points x and y on $\mathbb{S}^n$. Here we are interested when $d(x,y)$ is small. Let us take $0<\theta<\frac{\pi}{2}$ and consider smooth functions on $(0,\infty)\times[0,\theta)$. In the normal coordinates $\mathbf{n}$ at a point, the distance $d(x,y)$ is given by

$$d(x,y) = \cos^{-1}\left(\cos(|\mathbf{n}(x)|)\cdot\cos(|\mathbf{n}(y)|)+\left\langle\mathbf{n}(x),\mathbf{n}(y)\right\rangle_g\frac{\sin(|\mathbf{n}(x)|)}{|\mathbf{n}(x)|}\frac{\sin(|\mathbf{n}(y)|)}{|\mathbf{n}(y)|}\right),$$

where $\left\langle\mathbf{n}(x),\mathbf{n}(y)\right\rangle_g$ is the inner product of the sphere induced by the Riemannian metric g, To simplify our calculation, we assume the g is the metric induced by the Euclidean metric in $\mathbf{R}^n$ to the sphere $\mathbb{S}^n$. Thus the behavior of derivatives of the heat kernel $P_n(t,d(x,y))$ can be explicitly described in terms of this "distance" d by elementary functions up to exponential decay error terms. For more detailed discussion, see Nagase [30]. In other words, take small $\theta>0$. Then for $0 < t \le 1$ and $d = d(x,y) < \theta$, one has

$$\begin{aligned} K_n(t,d) &= p_n(t,d) + \mathscr{O}_\infty\big(e^{-\frac{1}{t}}\big),\\ K_{n+2}(t,d) &= \frac{e^{nt}}{2\pi}\Big(-\frac{1}{\sin d}\frac{\partial}{\partial d}\Big)K_n(t,d) + \mathscr{O}_\infty\big(e^{-\frac{1}{t}}\big),\\ K_{n+1}(t,d) &= \sqrt{2}\int_d^\theta \frac{e^{-(2n+1)t/4}K_{n+2}(t,\eta)\sin\eta}{(\cos d - \cos\eta)^{\frac12}}\,d\eta + \mathscr{O}_\infty\big(e^{-\frac{1}{t}}\big). \end{aligned} \tag{2.5}$$

Here $\mathscr{O}_\infty\big(e^{-\frac{1}{t}}\big)$ is a quantity, for any k, whose k times derivative by d can be estimated as $\mathscr{O}_\infty\big(e^{-\frac{\varepsilon_k}{t}}\big)$ with some positive constant ε. Furthermore, the three heat kernels in (2.5) can be expressed roughly by a finite series of smooth functions. If $n = 2m+1$, then we know from classical result that

$$\begin{aligned} p_{2m+1}(t,d) &= \frac{e^{m^2 t}}{(2\pi)^m(4\pi t)^{\frac12}}\Big(\frac{-1}{\sin d}\frac{\partial}{\partial d}\Big)^m e^{-\frac{d^2}{4t}}\\ &= \frac{e^{-\frac{d^2}{4t}+\frac{(n-1)^2 t}{4}}}{(4\pi t)^{\frac{n}{2}}}\sum_{j=0}^{m-1} t^j p_{2m+1,j}(d^2), \end{aligned} \tag{2.6}$$

with $p_{2m+1,j}(\xi) \in C^\infty(-\theta^2,\theta^2)$, $j = 0,1,2,\dots$. Now let us turn to the case $n = 2m+2$, then we know that (see [30]):

$$\begin{aligned} p_{2m+2}(t,d) &= \frac{e^{\frac{(2m+1)^2 t}{4}}}{2(2\pi)^{m+\frac32}t^{\frac32}}\Big(\frac{-1}{\sin d}\frac{\partial}{\partial d}\Big)^m\int_d^\theta \frac{\eta e^{-\frac{\eta^2}{4t}}}{(\cos d-\cos\eta)^{\frac12}}\,d\eta\\ &= \frac{e^{\frac{(2m+1)^2 t}{4}}}{2(2\pi)^{m+\frac32}t^{\frac32}}\int_d^\theta (\cos d-\cos\eta)^{\frac12}\Big(\frac{\partial}{\partial\eta}\frac{-1}{\sin\eta}\Big)^{m+1}\eta e^{-\frac{\eta^2}{4t}}\,d\eta + \mathscr{O}_\infty\big(e^{-\frac{1}{t}}\big). \end{aligned} \tag{2.7}$$

In fact, equation (2.8) can be written roughly by a finite series of smooth functions plus acceptable error term either:

$$p_{2m+2}(t,d) = \frac{e^{-\frac{d^2}{4t}+\frac{(n-1)^2 t}{4}}}{(4\pi t)^{\frac{n}{2}}}\Big[\sum_{j=0}^N t^j p_{2m+2,j}(d^2) + \mathscr{O}_\infty\big(d^N\big)\Big], \tag{2.8}$$

with $p_{2m+2,j}(\xi) \in C^\infty(-\theta^2,\theta^2)$, $j = 0,1,2,\dots$.

Let us prove that (2.7) implies (2.8). We start with

$$\begin{aligned} &(\cos d-\cos\eta)^{\frac12}\Big(\frac{\partial}{\partial\eta}\frac{-1}{\sin\eta}\Big)^{m+1}\eta e^{-\frac{\eta^2}{4t}}\\ &= \frac{1}{\sqrt2}(\eta^2-d^2)^{\frac12}\Big(\frac{\sin(\frac{d+\eta}{2})}{\frac{d+\eta}{2}}\Big)^{\frac12}\Big(\frac{\sin(\frac{d-\eta}{2})}{\frac{d-\eta}{2})}\Big)^{\frac12}\Big(\frac{\partial}{\partial\eta}\frac{-1}{\sin\eta}\Big)^{m+1}\eta e^{-\frac{\eta^2}{4t}}\\ &= \sum_{j=0}^{m+1} t^{-j}\eta e^{-\frac{\eta^2}{4t}}(\eta^2-d^2)^{\frac12}W_j(\eta^2,d^2), \end{aligned}$$

where $W_j(\mu,\nu) \in C^\infty\big((-\theta^2,\theta^2)\times(-\theta^2,\theta^2)\big)$.

It follows that

$$\begin{aligned}
&\int_d^{\theta} (\cos d - \cos\eta)^{\frac{1}{2}} \Big(\frac{\partial}{\partial\eta}\frac{-1}{\sin\eta}\Big)^{m+1} \eta e^{-\frac{\eta^2}{4t}}\, d\eta \\
=&\, e^{-\frac{d^2}{4t}} \sum_{j=1}^{m+1} t^{-j} \int_d^{\theta} \eta e^{-\frac{\eta^2-d^2}{4t}} (\eta^2-d^2)^{\frac{1}{2}} W_j(\eta^2,d^2)\, d\eta \\
=&\, e^{-\frac{d^2}{4t}} 4t^{\frac{3}{2}} \sum_{j=1}^{m+1} t^{-j} \int_d^{\frac{\theta^2-d^2}{4t}} e^{-\xi}\xi^{\frac{1}{2}} W_j(4t\xi+d^2,d^2)\, d\eta,
\end{aligned} \tag{2.9}$$

Now we have

$$\begin{aligned}
&\frac{(4\pi t)^{m+1} e^{\frac{d^2}{4t}}}{(2\pi)^{m+\frac{3}{2}} t^{\frac{3}{2}}} \int_d^{\theta} (\cos d - \cos\eta)^{\frac{1}{2}} \Big(\frac{\eta}{2t\sin\eta}\Big)^{m+1} \eta e^{-\frac{\eta^2}{4t}}\, d\eta \\
=&\, \frac{1}{\sqrt{\pi}} \int_0^{\frac{\theta^2-d^2}{4t}} e^{-\xi}\xi^{\frac{1}{2}}\, d\xi \cdot \Big(\frac{\sin(\frac{d+\eta}{2})}{\frac{d+\eta}{2}}\Big)^{\frac{1}{2}} \Big(\frac{\sin(\frac{d-\eta}{2})}{\frac{d-\eta}{2})}\Big)^{\frac{1}{2}} \Big(\frac{\eta}{\sin\eta}\Big)^{m+1}\Big|_{\eta=(4t\xi+d^2)^{1/2}} + \mathscr{O}(t) \\
\to&\, \frac{1}{\sqrt{\pi}} \int_0^{\infty} e^{-\xi}\xi^{\frac{1}{2}}\, d\xi \cdot \Big(\frac{\sin(\frac{d+\eta}{2})}{\frac{d+\eta}{2}}\Big)^{\frac{1}{2}} \Big(\frac{\sin(\frac{d-\eta}{2})}{\frac{d-\eta}{2})}\Big)^{\frac{1}{2}} \Big(\frac{\eta}{\sin\eta}\Big)^{m+1}\Big|_{\eta=d} \quad \text{as } t\to 0^+ \\
=&\, \frac{1}{\sqrt{\pi}} \int_0^{\infty} e^{-\xi}\xi^{\frac{1}{2}}\, d\xi \cdot \Big(\frac{\sin d}{d}\Big)^{\frac{1}{2}} \Big(\frac{d}{\sin d}\Big)^{m+1} = \Big(\frac{d}{\sin d}\Big)^{m+\frac{1}{2}}.
\end{aligned}$$

Since $m = \frac{n}{2} - 1$, hence we have,

$$p_{2m+1,0}(d^2) \,=\, \lim_{t\to 0^+} \big(4\pi t\big)^{\frac{n}{2}} \cdot e^{\frac{d^2}{4t} - \frac{(n-1)^2 t}{4}} \cdot p_n(t,d) \,=\, \Big(\frac{d}{\sin d}\Big)^{\frac{n-1}{2}}. \tag{2.10}$$

Now we are in a position to show that $p_n(t,d)$ is a parametrix of the initial value problem for the heat operator (1.2). We have the following theorem.

Theorem 2.2. *Let $\theta > 0$ be a small positive number. Then for $0 < t \le 1$ and $d = d(x,y) < \theta$, one has*
(a). $\Big(\frac{\partial}{\partial d}\Big)^m p_n(t,d) \,=\, \mathscr{O}\big(t^{-\frac{n}{2}-\frac{m}{2}} e^{-\frac{d^2}{5t}}\big)$ *for all* $m \in \mathbf{Z}_+$*,*
(b). $\Big(\frac{\partial}{\partial t} + \Delta\Big) p_n(t,d) \,=\, \mathscr{O}_\infty\big(e^{-\frac{1}{t}}\big)$.
Moreover, we have

$$\lim_{t\to 0^+} \int_{\mathbb{S}^n} p_n(t,d(x,y)) f(y)\, dv(y) \,=\, f(x)$$

for $f \in C^\infty$ *with* $supp(f) \subset \big\{y \in \mathbb{S}^n :\ d(x,y) < \theta\big\}$.

3. SubRiemannian Manifolds and Subelliptic Operators

In the second part of the paper, we turn to subRiemannian cases. SubRiemannian geometry is proved to play an important role in many applications, e.g., in mathematical physics, geometric mechanics, robotics, tomography, neurosystems, and control theory. This geometry enjoys major differences from the Riemannian being a generalization of the latter at the same time, e.g., the notion of geodesic and length minimizer do not coincide even locally, the Hausdorff dimension is larger than the topological dimension of the manifold and the exponential map is never a local diffeomorphism. There exists a large amount of literature developing subRiemannian geometry. Once again, we refer readers to references are [2, 7, 28, 33] for detailed discussions.

The interest to odd-dimensional spheres comes first of all from finite dimensional quantum mechanics modeled over the Hilbert space $\mathbf{C}^n$ where the dimension n is the number of energy levels and the normalized state vectors form the sphere $\mathbb{S}^{2n-1} \subset \mathbf{C}^n$. The problem of controlled quantum systems is basically the problem of controlled spin systems, which is reduced to the left- or right-invariant control problem on the Lie group $SU(n)$. In other words, these are problems of describing the subRiemannian structure of $\mathbb{S}^{2n-1}$ and the subRiemannian geodesics, see e.g., [5, 28]. The special case $n = 2$ is well studied and the subRiemannian structure is related to the classical Hopf fibration, see, e.g., [34, 35]. At the same time, the subRiemannian structure of $\mathbb{S}^3$ comes naturally from the non-commutative group structure of $SU(2)$ in the sense that two vector fields span the smoothly varying distribution of the tangent bundle, and their commutator generates the missing direction. The missing direction coincides with the Hopf vector field corresponding to the Hopf fibration. The subRiemannian geometry on $\mathbb{S}^3$ was studied in [10, 21, 17]. Explicit formulas for geodesics were obtained in [12] by solving the corresponding Hamiltonian system, in [21] from a variational equation, in [6] by applying the structure of the principle $\mathbb{S}^1$-bundle. One of the important helping properties of odd-dimensional spheres is that there always exists at least one globally defined non-vanishing vector field.

Observe that $\mathbb{S}^3$ is compact and many properties and results of subRiemannian geometry differ from the standard nilpotent case, e.g., Heisenberg group or Engel group. In the case $\mathbb{S}^{2n-1}$, $n > 2$, we have no group structure and the main tool is the global action of the group $U(1)$. For example, in our paper we explicitly show that any two points of $\mathbb{S}^3$ can be connected with an infinite number of geodesics.

Because of important applications, we start our paper with the description of n-level quantum systems and motivation given by Berry phases. Further we continue with general formulas for geodesics. Then we concentrate our attention on the *geodesic boundary value problem* finding all subRiemannian geodesics between two given points. In the case of $\mathbb{S}^{2n-1}$ we solve it for the points of the fiber and for $\mathbb{S}^3$ we solve it for arbitrary two points. The Carnot-Carathéodory distance is calculated.

Now let us give more detailed discussion on this subject. Let $\mathscr{M}_n$ be an n-dimensional, smooth connected manifold, together with a smooth distribution $\mathscr{H} \subset T\mathscr{M}_n$ of rank $2 \le k < n$. Such vector bundles are often called *horizontal*. An absolutely continuous curve $\gamma\colon [0,1] \to \mathscr{M}_n$ is called *horizontal* if $\dot\gamma(s) \in \mathscr{H}$ a.e.

Define the following real vector bundles

$$\mathscr{H}^1 = \mathscr{H}, \qquad \mathscr{H}^{k+1} = [\mathscr{H}^k, \mathscr{H}] + \mathscr{H}^k \quad \text{for } k \ge 1,$$

which naturally give rise to the flag

$$\mathscr{H} = \mathscr{H}^1 \subseteq \mathscr{H}^2 \subseteq \mathscr{H}^3 \subseteq \dots.$$

Then we say that a distribution satisfy *bracket generating condition* if $\forall\, x \in \mathscr{M}_n$ $\exists\, k(x) \in \mathbb{Z}_+$ such that

$$\mathscr{H}_x^{k(x)} = T_x\mathscr{M}_n. \tag{3.1}$$

If the dimensions $\dim \mathscr{H}_x^k$ do not depend on x for any $k \geq 1$, we say that $\mathscr{H}$ is a *regular distribution*. The least k such that (3.1) is satisfied is called the *step* of $\mathscr{H}$.

The following classical result shows the precise relation between the notion of path-connectedness by means of horizontal curves and the assumption that $\mathscr{H}$ is a bracket generating distribution, (see Chow [14]).

Theorem 3.1. *(Chow's Theorem) Given any two points $A, B \in \mathscr{M}_n$, there is a piecewise C^1 horizontal curve $\gamma : [0,\tau] \to \mathscr{M}_n$:*

$$\gamma(0) = A, \;\; \gamma(\tau) = B,$$

and

$$\dot{\gamma}(s) = \sum_{k=1}^{m} a_k(s) X_k.$$

Here $\{X_1, \ldots, X_m\}$ is a basis for the regular distribution $\mathscr{H}$.

Definition 3.1. A subRiemannian structure over a manifold $\mathscr{M}_n$ is a pair $(\mathscr{H}, \langle \cdot,\cdot \rangle)$, where $\mathscr{H}$ is a bracket generating distribution and $\langle \cdot,\cdot \rangle$ a fiber inner product defined on $\mathscr{H}$. In this setting, the length of an absolutely continuous horizontal curve $\gamma : [0,1] \to \mathscr{M}_n$ is

$$\ell(\gamma) := \int_0^1 \|\dot{\gamma}(s)\| ds = \int_0^1 \sqrt{a_1^2(s) + \cdots + a_m^2(s)}\, ds,$$

where $\|\dot{\gamma}(s)\|^2 = \langle \dot{\gamma}(s), \dot{\gamma}(s) \rangle$ whenever $\dot{\gamma}(s)$ exists. The triple $(\mathscr{M}_n, \mathscr{H}, \langle \cdot,\cdot \rangle)$ is called subRiemannian manifold.

The shortest length $d_{cc}(A,B)$ is called a *Carnot-Carathéodory distance* between $A, B \in \mathscr{M}_n$ which is given by

$$d_{cc}(A,B) := \inf \ell(\gamma),$$

where the infimum is taken over all absolutely continuous horizontal curves joining A and B.

The following theorem in topology gives a very strong restriction on the problem of finding globally defined subRiemannian structures over spheres.

Theorem 3.2. *(Adams' Theorem [1]) Let $\mathbb{S}^{n-1} = \{x \in \mathbf{R}^n : \|x\|^2 = 1\}$ be the unit sphere in $\mathbf{R}^n$, with respect to the usual Euclidean norm $\|\cdot\|$. Then $\mathbb{S}^{n-1}$ has precisely $\rho(n) - 1$ linearly independent, globally defined and non vanishing vector fields, where $\rho(n)$ is defined in the following way: if*

$$n = (2a+1)2^b, \quad \textit{and} \quad b = c + 4d,$$

where $0 \leq c \leq 3$, then $\rho(n) = 2^c + 8d$.

The condition that a manifold $\mathscr{M}_n$ has maximal number of linearly independent globally defined non-vanishing vector fields is usually rephrased as saying that $\mathscr{M}_n$ is *parallelizable*. It was proved by Bott and Milnor that $\mathbb{S}^n$ is a parallelizable sphere only when $n = 1, 3, 7$. It is impossible to find a globally defined basis for bracket generating distributions, except for $\mathbb{S}^3$ and $\mathbb{S}^7$. The fact that $\mathbb{S}^3$ and $\mathbb{S}^7$ can be considered as the unit spheres in quaternions and octonions.

4. $\mathbb{S}^3$ AS A SUBRIEMANNIAN MANIFOLD

The unit 3-sphere centered at the origin is the set of $\mathbf{R}^4$ defined by

$$\mathbb{S}^3 = \{(x_1,x_2,x_3,x_4) \in \mathbf{R}^4 : \; x_1^2+x_2^2+x_3^2+x_4^2 = 1\}.$$

It is often convenient to regard $\mathbf{R}^4$ as the space with two complex dimensions $\mathbf{C}^2$ or the quaternions $\mathbb{H}$. The unit 3-sphere is then given by

$$\mathbb{S}^3 = \{(z_1,z_2) \in \mathbb{C}^2 : \; |z_1|^2+|z_2|^2 = 1\} \quad \text{or} \quad \mathbb{S}^3 = \{q \in \mathbb{H} : \; |q|^2 = 1\}.$$

The last description represents the sphere $\mathbb{S}^3$ as a set of unit quaternions and (as the set of unit complex number) it can be considered as a group $Sp(1)$, where the group operation is just a multiplication of quaternions. The group $Sp(1)$ is a three-dimensional Lie group, isomorphic to the Lie group $SU(2)$ by the isomorphism $\mathbb{C}^2 \ni (z_1,z_2) \to q \in \mathbb{H}$. The unitary group $SU(2)$ is the group of matrices

$$\begin{pmatrix} z_1 & z_2 \\ -\bar{z}_2 & \bar{z}_1 \end{pmatrix}, \qquad z_1,z_2 \in \mathbb{C}, \quad |z_1|^2+|z_2|^2 = 1,$$

where the group law is given by the multiplication of matrices. Identify $\mathbf{R}^3$ with pure imaginary quaternions. The conjugation $qh\bar{q}$ of a pure imaginary quaternion h by a unit quaternion q defines a rotation in $\mathbf{R}^3$ and since $|qh\bar{q}| = |h|$, the map $h \mapsto qh\bar{q}$ defines a two-to-one homomorphism $Sp(1) \to SO(3)$. The Hopf map $\pi : \mathbb{S}^3 \to \mathbb{S}^2$ can be defined by

$$\mathbb{S}^3 \ni q \mapsto qi\bar{q} = \pi(q) \in \mathbb{S}^2.$$

The Hopf map defines a principle circle bundle also known as the Hopf bundle. Topologically $\mathbb{S}^3$ is a compact, simply-connected, 3-dimensional manifold without boundary.

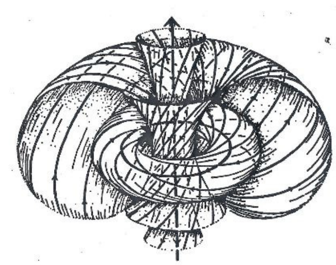

FIGURE 1. The Hopf fibration can be visualized using a stereographic projection of $\mathbb{S}^3$ into $\mathbf{R}^3$.

We mentioned only the small part of properties of the unit 3-sphere that find numerous applications in complex geometry, topology, group theory, mathematical physics and others branches of mathematics. In the present paper we give a new emphasis to the unit 3-sphere, considering it as a subRiemannian manifold. The subRiemannian structure comes naturally from the noncommutative group structure of the sphere in a sense that two vector fields spans the smoothly varying distribution of the tangent bundle and their commutator generates the missing direction. The subRiemannian metric is defined as a restriction of the Euclidean inner product from $\mathbf{R}^4$ to the distribution. The present paper devoted to the description of the subRiemannian geodesics on the sphere. The subRiemannian geodesics are defined as a projection into the manifold of the solution to the corresponding Hamiltonian system. We give the explicit formulae using the different parametrization and discuss the number of geodesics starting from the unity of the group.

4.1. **Left-invariant vector fields and the horizontal distribution.** To calculate the left invariant vector fields we use the definition of $\mathbb{S}^3$ as a set of unit quaternions equipped with the following noncommutative multiplication: let $x=(x_1,x_2,x_3,x_4)$, $y=(y_1,y_2,y_3,y_4)$, then

$$\begin{aligned} x\circ y &= (x_1,x_2,x_3,x_4)\circ(y_1,y_2,y_3,y_4) \\ &= \Big((x_1y_1-x_2y_2-x_3y_3-x_4y_4),(x_2y_1+x_1y_2-x_4y_3+x_3y_4), \\ &\quad (x_3y_1+x_4y_2+x_1y_3-x_2y_4),(x_4y_1-x_3y_2+x_2y_3+x_1y_4)\Big). \end{aligned} \tag{4.1}$$

The law (4.1) gives us the left translation $L_x(y)$ of an element $y=(y_1,y_2,y_3,y_4)$ by an element $x=(x_1,x_2,x_3,x_4)$. The left-invariant basis vector fields are defined as $X(x)=(L_x(y))_*X(0)$, where $X(0)$ are basis vectors at the unity of the group. The matrix corresponding to the tangent map $(L_x(y))_*$ calculated by (4.1) becomes

$$(L_x(y))_*=\begin{pmatrix} x_1 & -x_2 & -x_3 & -x_4 \\ x_2 & x_1 & -x_4 & x_3 \\ x_3 & x_4 & x_1 & -x_2 \\ x_4 & -x_3 & x_2 & x_1 \end{pmatrix}.$$

Calculating the action of $(L_x(y))_*$ on the basis of unit vectors of $\mathbf{R}^4$ we obtan four vector fields

$$\begin{aligned} X_1(x) &= +x_1\partial_{x_1}+x_2\partial_{x_2}+x_3\partial_{x_3}+x_4\partial_{x_4}, \\ X_2(x) &= -x_2\partial_{x_1}+x_1\partial_{x_2}+x_4\partial_{x_3}-x_3\partial_{x_4}, \\ X_3(x) &= -x_3\partial_{x_1}-x_4\partial_{x_2}+x_1\partial_{x_3}+x_2\partial_{x_4}, \\ X_4(x) &= -x_4\partial_{x_1}+x_3\partial_{x_2}-x_2\partial_{x_3}+x_1\partial_{x_4}. \end{aligned} \tag{4.2}$$

It is easy to see that the vector $X_1(x)$ is the unit normal to $\mathbb{S}^3$ at x with respect to the usual inner product $\langle\cdot,\cdot\rangle$ in $\mathbf{R}^4$, hence, we denote $X_1(x)$ by N. Moreover,

$$\langle N,X_2(x)\rangle=\langle N,X_3(x)\rangle=\langle N,X_4(x)\rangle=0, \quad \text{and} \quad |X_k(x)|^2=\langle X_k(x),X_k(x)\rangle=1$$

for $k=2,3,4$, and any $x\in\mathbb{S}^3$. The matrix

$$\begin{pmatrix} -x_2 & x_1 & x_4 & -x_3 \\ -x_3 & -x_4 & x_1 & x_2 \\ -x_4 & x_3 & -x_2 & x_1 \end{pmatrix}$$

has rank 3, and we conclude that the vector fields $X_2(x)$, $X_3(x)$, $X_4(x)$ form an orthonormal basis with respect to $\langle\cdot,\cdot\rangle$ of the tangent space $T_x\mathbb{S}^3$ at any point $x\in\mathbb{S}^3$. Let us denote the vector fields by

$$X_3=X, \quad X_4=Y, \quad X_2=Z.$$

The vector fields possess the following commutation relations

$$[X,Y]=XY-YX=2Z, \quad [Z,X]=2Y, \quad [Y,Z]=2X.$$

Let $\mathscr{D}=\operatorname{span}\{X,Y\}$ be the distribution generated by the vector fields X and Y. Since $[X,Y]=2Z\notin\mathscr{D}$, it follows that $\mathscr{D}$ is not involutive. The distribution $\mathscr{D}$ will be called *horizontal*. Any curve on the sphere with the velocity vector contained in the distribution $\mathscr{D}$ will be called a *horizontal curve*. Since $T_x\mathbb{S}^3 = \text{span}\{X,Y,Z = 1/2[X,Y]\}$, the distribution is bracket generating. We define the metric on the distribution $\mathscr{D}$ as the restriction of the metric $\langle\cdot,\cdot\rangle$ onto $\mathscr{D}$, and the same notation will be used. The manifold $(\mathbb{S}^3,\mathscr{D},\langle\cdot,\cdot\rangle)$ is a step two subRiemannian manifold.

Remark. Notice that the choice of the horizontal distribution is not unique. The relations $[Z,X] = 2Y$ and $[Y,Z] = 2X$ imply possible choices $\mathscr{D} = span\{X,Z\}$ or $\mathscr{D} = span\{Y,Z\}$. The geometries defined by different horizontal distributions are cyclically symmetric, so we restrict our attention to the $\mathscr{D} = span\{X,Y\}$.

We also can define the distribution as a kernel of the following one form

$$\omega = -x_2dx_1 + x_1dx_2 + x_4dx_3 - x_3dx_4$$

on $\mathbb{R}^4$. One can easily check that

$$\omega(X) = 0, \quad \omega(Y) = 0, \quad \omega(Z) = 1 \neq 0, \quad \omega(N) = 0.$$

Hence, $\ker\omega = span\{X,Y,N\}$, and the horizontal distribution can be written as

$$\mathbb{S}^3 \ni x \to \mathscr{D}_x = \ker\omega \cap T_x\mathbb{S}^3.$$

Let $\gamma(s) = (x_1(s),x_2(s),x_3(s),x_4(s))$ be a curve on $\mathscr{S}^3$. Then the velocity vector, written in the left-invariant basis, is

$$\dot\gamma(s) = a(s)X(\gamma(s)) + b(s)Y(\gamma(s)) + c(s)Z(\gamma(s)),$$

where

$$\begin{aligned} a &= \langle\dot\gamma,X\rangle = -x_3\dot x_1 - x_4\dot x_2 + x_1\dot x_3 + x_2\dot x_4, \\ b &= \langle\dot\gamma,Y\rangle = -x_4\dot x_1 + x_3\dot x_2 - x_2\dot x_3 + x_1\dot x_4, \\ c &= \langle\dot\gamma,Y\rangle = -x_2\dot x_1 + x_1\dot x_2 + x_4\dot x_3 - x_3\dot x_4. \end{aligned} \tag{4.3}$$

The following proposition holds.

Proposition 4.1. *Let $\gamma(s) = (x_1(s),x_2(s),y_1(s),y_2(s))$ be a curve on $\mathbb{S}^3$. The curve γ is horizontal, if and only if,*

$$c = \langle\dot\gamma,Z\rangle = \langle\dot\gamma,X\rangle = -x_2\dot x_1 + x_1\dot x_2 + x_4\dot x_3 - x_3\dot x_4 = 0. \tag{4.4}$$

The manifold $\mathbb{S}^3$ is connected which satisfies the bracket generating condition. Once again, by Chow's Theorem [14], there exists a piecewise C^1 horizontal curves connecting two arbitrary points on $\mathbb{S}^3$. In fact, smooth horizontal curves connecting two arbitrary points on $\mathbb{S}^3$ were constructed in [10].

Proposition 4.2. *The horizontality property is invariant under the left translation.*

Proof. It can be shown that (4.3) do not change under the left translation. This implies the conclusion of the proposition. □

5. Hamiltonian System

Ones we have a system of curves, in our case the system of horizontal curves, we can define the length as in the Riemannian geometry. Let $\gamma\colon[0,1]\to\mathbb{S}^3$ be a horizontal curve such that $\gamma(0)=x$, $\gamma(1)=y$, then the length $\ell(\gamma)$ of γ is defined as the following

$$\ell(\gamma)=\int_0^1\langle\dot\gamma,\dot\gamma\rangle^{1/2}\,dt=\int_0^1\big(a^2(t)+b^2(t)\big)^{1/2}\,dt. \tag{5.1}$$

Now we are able to define the distance between the points x and y by minimizing the integral (5.1) or the corresponding energy integral $\int_0^1\big(a^2(t)+b^2(t)\big)\,dt$ under the non-holonomic constraint (4.4). This is Lagrangian approach. The Lagrangian formalism was applied to study the subRiemannian geometry of $\mathbb{S}^3$ in [10]. In the Riemannian geometry the minimizing curve locally coincide with a geodesic, but it is not the case in the subRiemannian manifolds. The interesting examples and discussions can be found, for instance in [24, 26, 27, 28, 33]. Given the subRiemannian metric we can form a Hamiltonian function defined on the cotangent bundle of $\mathbb{S}^3$. The geodesics in the subRiemannian manifolds defines as a projection into the manifold of the solution to the corresponding Hamiltonian system. It is a good generalization of the Riemannian case in the following sense. The Riemannian geodesics (that are defined as curves with vanishing acceleration) lift to solutions of the Hamilton system on the cotangent bundle.

In the present paper we are interested in the construction of subRiemannian geodesics on $(\mathbb{S}^3,\mathscr{D},\langle\cdot,\cdot\rangle)$. Let us write the left invariant vector fields X,Y,Z, using the matrices

$$I_1=\begin{bmatrix}0&0&-1&0\\0&0&0&-1\\1&0&0&0\\0&1&0&0\end{bmatrix},\quad I_2=\begin{bmatrix}0&0&0&-1\\0&0&1&0\\0&-1&0&0\\1&0&0&0\end{bmatrix},\quad I_3=\begin{bmatrix}0&-1&0&0\\1&0&0&0\\0&0&0&1\\0&0&-1&0\end{bmatrix}. \tag{5.2}$$

Then

$$X=\langle I_1x,\nabla x\rangle,\quad Y=\langle I_2x,\nabla x\rangle,\quad Z=\langle I_3x,\nabla x\rangle.$$

The Hamilton function is defined as

$$H=\frac12(X^2+Y^2)=\frac12\Big(\langle I_1x,\xi\rangle^2+\langle I_2x,\xi\rangle^2\Big),$$

where $\xi=\nabla x$. Then the Hamilton system follows as

$$\begin{aligned}\dot x&=\frac{\partial H}{\partial\xi}\quad\Rightarrow\quad \dot x=\langle I_1x,\xi\rangle\cdot(I_1x)+\langle I_2x,\xi\rangle\cdot(I_2x),\\ \dot\xi&=-\frac{\partial H}{\partial x}\quad\Rightarrow\quad \dot\xi=\langle I_1x,\xi\rangle\cdot(I_1\xi)+\langle I_2x,\xi\rangle\cdot(I_2\xi).\end{aligned} \tag{5.3}$$

As was mentioned a geodesic is the projection of a solution of the Hamiltonian system onto the x-space. We obtain the following properties.

1. Since $\langle I_1x,x\rangle=\langle I_2x,x\rangle=\langle I_3x,x\rangle=0$, multiplying the first equation of (5.3) by x, we obtain

   $$\langle\dot x,x\rangle=0\quad\Rightarrow\quad|x|^2=const.$$

   We conclude that *any solution of the Hamiltonian system belongs to the sphere.* Taking the constant equal to 1, we obtain geodesics on $\mathbb{S}^3$.

2. Multiplying the first equation of (5.3) by I_3x, we obtain
$$\langle \dot{x}, I_3x\rangle = 0, \tag{5.4}$$
by the role of multiplication between I_1, I_2, and I_3. The reader easily recognizes the horizontality condition $\langle \dot{x}, Z\rangle = 0$ in (5.4). It means that *any solution of the Hamiltonian system is a horizontal curve*.
3. Multiplying the first equation of (5.3) by I_1x, and then by I_2x, we obtain
$$\langle \xi, I_1x\rangle = \langle \dot{x}, I_1x\rangle, \qquad \langle \xi, xI_2\rangle = \langle \dot{x}, xI_2\rangle.$$
From the other side we know that $\langle \dot{x}, I_1x\rangle = a$ and $\langle \dot{x}, xI_2\rangle = b$. The Hamiltonian function can be written in the form
$$H = \frac{1}{2}\Big(\langle I_1x, \xi\rangle^2 + \langle I_2x, \xi\rangle^2\Big) = \frac{1}{2}\Big(\langle I_1x, \dot{x}\rangle^2 + \langle I_2x, \dot{x}\rangle^2\Big) = \frac{1}{2}\Big(a^2 + b^2\Big).$$
Thus *the Hamiltonian function gives the kinetic energy* $H = \frac{|\dot{q}|^2}{2}$ *and it is a constant along the geodesics.*
4. If we multiply the first equation of (5.3) by $\dot{x}$, then
$$|\dot{x}|^2 = \langle I_1x, \xi\rangle^2 + \langle I_2x, \xi\rangle^2 = \langle I_1x, \dot{x}\rangle^2 + \langle I_2x, \dot{x}\rangle^2 = a^2 + b^2 = 2H.$$
Then
$$|\dot{x}|^2 = a^2 + b^2. \tag{5.5}$$

5.1. **Velocity vector with constant coordinates.** We know that along geodesics the length of the velocity vector is constant. Let us start from the simplest case, when the coordinates of the velocity vector are constant. Suppose that $\dot{a} = \dot{b} = 0$. The first line of system (5.3) can be written as follows
$$\begin{aligned} \dot{x}_1 &= -ax_3 - bx_4 & \qquad \dot{x}_3 &= +ax_1 - bx_2 \\ \dot{x}_2 &= -ax_4 + bx_3 & \qquad \dot{x}_4 &= +ax_2 + bx_1. \end{aligned} \tag{5.6}$$
Differentiation of system (5.6) yields
$$\begin{aligned} \ddot{x}_1 &= -a\dot{x}_3 - b\dot{x}_4 & \qquad \ddot{x}_3 &= +a\dot{x}_1 - b\dot{x}_2 \\ \ddot{x}_2 &= -a\dot{x}_4 + b\dot{x}_3 & \qquad \ddot{x}_4 &= +a\dot{x}_2 + b\dot{x}_1. \end{aligned} \tag{5.7}$$
We substitute the first derivatives from (5.6) in (5.7) to obtain
$$\ddot{x}_k = -r^2x_k, \quad r^2 = a^2 + b^2, \quad k = 1,2,3,4. \tag{5.8}$$

Theorem 5.1. *The set of geodesics with constant velocity coordinates form a unit sphere* $\mathbb{S}^2$ *in* $\mathbf{R}^3$

Proof. We are looking for horizontal geodesics parametrized by the arc length and starting from the point $x(0) = x_0$. So, we set $r = 1$ and $a = \cos\psi$, $b = \sin\psi$, where ψ is a constant from $[0, 2\pi)$. Solving the equation (5.8), we obtain the general solution $x(s) = A\cos s + B\sin s$. We conclude that $A = x_0$ from the initial data. To find B, let us substitute the general solution in equations (5.6) and obtain $B = (aI_1 + bI_2)x_0$. Thus, the horizontal geodesics with constant horizontal coordinates are
$$x(s) = x_0\cos s + (\cos\psi I_1 + \sin\psi I_2)x_0\sin s.$$

Since the geodesics are invariant under the left translation it is sufficient to describe the situation at the unity element, *e.g.*, $x_0 = (1,0,0,0)$ of $\mathbb{S}^3$. In this case, the geodesics are

$$\begin{aligned} x_1 &= \cos s, & x_3 &= \cos\psi \sin s, \\ x_2 &= 0, & x_4 &= \sin\psi \sin s. \end{aligned} \tag{5.9}$$

We see that the set of geodesics with constant velocity coordinates form the unit sphere $\mathbb{S}^2$ in $\mathbb{R}^3 = \{(x_1, 0, x_3, x_4)\}$. The parameter $\psi \in [0, 2\pi)$ corresponds to the initial velocity. □

The sphere (5.9) is a direct analogue of the horizontal plane on the Heisenberg group $\mathbb{H}^1$ at the unity. Let us calculate the analogue of the vertical axis in $\mathbb{H}^1$. We wish to find an integral curve for the vector field Z. In other words, we solve the system

$$\begin{aligned} a &= \langle \dot\gamma, X\rangle = -x_3\dot x_1 - x_4\dot x_2 + x_1\dot x_3 + x_2\dot x_4 = 0, \\ b &= \langle \dot\gamma, Y\rangle = -x_4\dot x_1 + x_3\dot x_2 - x_2\dot x_3 + x_1\dot x_4 = 0, \\ c &= \langle \dot\gamma, Z\rangle = -x_2\dot x_1 + x_1\dot x_2 + x_4\dot x_3 - x_3\dot x_4 = 1, \\ n &= \langle \dot\gamma, N\rangle = +x_1\dot x_1 + x_2\dot x_2 + x_3\dot x_3 + x_4\dot x_4 = 0. \end{aligned} \tag{5.10}$$

The rank of the system is 1 and it is redused to

$$\begin{aligned} \dot x_1 &= -x_2, & \dot x_3 &= +x_4, \\ \dot x_2 &= +x_1, & \dot x_4 &= -x_3. \end{aligned}$$

Differentiating once more, we obtain the equation $\ddot x = -x$. The initial point is $x(0) = x_0$. Then system (5.10) gives the value of the initial velocity $\dot x(0) = I_3 x_0$. Taking into account this initial data we get the equation of the *vertical line*

$$x(s) = x_0 \cos s + I_3 x_0 \sin s.$$

In particular, at the point $(1,0,0,0)$ the equation of vertical line is

$$x_1 = \cos s, \quad x_2 = \sin s, \quad x_3 = 0, \quad x_4 = 0, \qquad s \in [0, 2\pi]. \tag{5.11}$$

5.2. **Velocity vector with non-constant coordinates.** *Cartesian coordinates*

Fix the initial point $x^{(0)} = (1,0,0,0)$. It is convenient to introduce complex coordinates $z = x_1 + ix_2$, $w = x_3 + ix_4$, $\varphi = \xi_1 + i\xi_2$, and $\psi = \xi_3 + i\xi_4$. Hence, the Hamiltonian admits the form $2H = |\bar w\varphi - z\bar\psi|^2$. The corresponding Hamiltonian system becomes

$$\begin{aligned} \dot z &= w(\bar w\varphi - z\bar\psi), & z(0) &= 1, \\ \dot w &= -z(w\bar\varphi - \bar z\psi), & w(0) &= 0, \\ \dot{\bar\varphi} &= \bar\psi(w\bar\varphi - \bar z\psi), & \bar\varphi(0) &= A - iB, \\ \dot{\bar\psi} &= -\bar\varphi(\bar w\varphi - z\bar\psi), & \bar\psi(0) &= C - iD. \end{aligned}$$

Here the constants B, C, and D have the following dynamical meaning: $\dot w(0) = C + iD$, and $B = -i\ddot w(0)/2\dot w(0)$ or if we write in real variables $C = \dot x_3(0)$, $D = \dot x_4(0)$, $B = \frac{1}{2}(\dot x_3(0)\ddot x_4(0) - \dot x_4(0)\ddot x_3(0))$. This complex Hamiltonian system has the first integrals

$$\begin{aligned} z\psi - w\varphi &= C + iD, \\ z\bar\varphi + w\bar\psi &= A - iB, \end{aligned}$$

and we have $|z|^2+|w|^2=1$ and $2H=C^2+D^2=1$ as an additional normalization. Therefore,

$$\begin{aligned}\varphi &= z(A+iB)-\bar{w}(C+iD),\\ \psi &= \bar{z}(C+iD)+w(A+iB).\end{aligned}$$

Let us introduce an auxiliary function $p=\bar{w}/z$. Substituting φ and ψ in the Hamiltonian system, we have the equation for p as

$$\dot{p}=(C+iD)p^2-2iBp+(C-iD), \quad p(0)=0.$$

The solution is

$$p(s)=\frac{(C-iD)\sin(s\sqrt{1+B^2})}{\sqrt{1+B^2}\cos(s\sqrt{1+B^2})+iB\sin(s\sqrt{1+B^2})}.$$

Taking into account that $\dot{z}\bar{z}=-w\dot{\bar{w}}$, we obtain the solution

$$z(s)=\Big(\cos(s\sqrt{1+B^2})+i\frac{B}{\sqrt{1+B^2}}\sin(s\sqrt{1+B^2})\Big)e^{-iBs}, \tag{5.12}$$

and

$$w(s)=\frac{C+iD}{\sqrt{1+B^2}}\sin(s\sqrt{1+B^2})e^{iBs}. \tag{5.13}$$

If $B=0$, we obtain the solutions with constant horizontal velocity coordinates

$$z(s)=\cos s, \qquad w(s)=(\dot{x}_3(0)+i\dot{x}_4(0))\sin s$$

from the previous section.

Theorem 5.2. *Let A be a point on the vertical line, i. e., $A=(\cos k,\sin k,0,0)$, $k\in[0,2\pi)$, then there are countably many geodesics γ_n connecting $O=(1,0,0,0)$ with A. They have the following parametric equations*

$$\begin{aligned} z_n(s) &= \Big(\cos\big(s\frac{\pi n}{\sqrt{\pi^2n^2-k^2}}\big)-i\frac{k}{\pi n}\sin\big(s\frac{\pi n}{\sqrt{\pi^2n^2-k^2}}\big)\Big)e^{\frac{isk}{\sqrt{\pi^2n^2-k^2}}}, \\ w_n(s) &= \big(\dot{x}_3(0)+i\dot{x}_4(0)\big)\frac{\sqrt{\pi^2n^2-k^2}}{\pi n}\sin\big(s\frac{\pi n}{\sqrt{\pi^2n^2-k^2}}\big)e^{\frac{-isk}{\sqrt{\pi^2n^2-k^2}}}, \end{aligned} \tag{5.14}$$

$n\in\mathbb{Z}\setminus\{0,\pm1\}$, $s\in[0,s_n]$, where $s_n=\sqrt{\pi^2n^2-k^2}$ is the length of the geodesic γ_n.

Proof. Since we use the condition $2H=|\dot{z}|^2+|\dot{w}|^2=1$ we conclude that geodesics parametrized by the length arc and the length of a geodesic at the value of parameter $s=l$ is equal to l. If $A=(z(s),w(s))$ belongs to the vertical line starting at $O=(1,0,0,0)$ then $|z(s)|=1$ and $|w(s)|=0$ provided that $-Bs=k$. It implies

$$\cos^2(s\sqrt{1+B^2})+\frac{B^2}{1+B^2}\sin^2(s\sqrt{1+B^2})=1, \quad \sin(s\sqrt{1+B^2})=0, \quad -Bs=k.$$

Equations are satisfied if

$$s_n=\sqrt{\pi^2n^2-k^2}, \quad B_n=-\frac{k}{\sqrt{\pi^2n^2-k^2}}, \quad n\in\mathbb{Z}\setminus\{0,\pm1\}.$$

We conclude, that for any $n\in\mathbb{Z}\setminus\{0,\pm1\}$ there is a constant $B_n=-\frac{k}{\sqrt{\pi^2n^2-k^2}}$ such that the corresponding geodesic $\gamma_n(s)$, $s\in[0,s_n]$, with the equation (5.14) joins the points O with A and the length of geodesic is equal to $s_n=\sqrt{\pi^2n^2-k^2}$. □

So far we have a clear picture of trivial geodesics whose velocity has constant coordinates. They are essentially unique (up to periodicity). The situation with geodesics joining the point $(1,0,0,0)$ with the points of the vertical line A has been described in the preceding theorem. Let us consider the general case of points on $\mathbb{S}^3$.

Theorem 5.3. *Given an arbitrary point $(z_1,w_1)\in\mathbb{S}^3$ which neither belongs to the vertical line A nor to the horizontal sphere $\mathbb{S}^2$, there is a countable number of geometrically different geodesics joining the initial point $(z_0,w_0)\in\mathbb{S}^3$ with (z_1,w_1), $z_0=1$, $w_0=0$.*

Proof. Let us denote

$$w_1=\rho e^{i\varphi},\quad z_1=re^{i\alpha},\quad C+iD=e^{i\theta}.$$

Then from (5.12) and (5.13) we have that

$$r^2=1-\frac{1}{1+B^2}\sin^2(s\sqrt{1+B^2}),\quad \text{and}\ \ \varphi=Bs+\theta. \tag{5.15}$$

We suppose for the moment that the angles $s\sqrt{1+B^2}$ and sB are from the first quadrant. Other cases are treated similarly. Then we have

$$z=(\sqrt{1-(1+B^2\rho^2)}+iB\rho)e^{i(\theta-\varphi)},$$

and

$$\theta=\theta(B)=\alpha+\varphi-\tan^{-1}\frac{B\rho}{\sqrt{1-(1+B^2)\rho^2}}.$$

The first expression in (5.15) leads to the value of the length parameter s

$$s=\frac{1}{\sqrt{1+B^2}}\sin^{-1}(\rho\sqrt{1+B^2}),$$

and the second to

$$\varphi=\theta+\frac{B}{\sqrt{1+B^2}}\sin^{-1}(\rho\sqrt{1+B^2}).$$

Substituting $\theta(B)$ in the latter equation we obtain

$$\sin\left((\varphi-\theta(B))\sqrt{1+\frac{1}{B^2}}\right)=\rho\sqrt{1+B^2}, \tag{5.16}$$

as an equation for the parameter B. Observe that $\varphi-\theta(B)$ is a bounded function and $\lim_{B\to 0}\theta(B)\neq 0$. Indeed, if the latter limit were vanishing, then the value of given φ would be zero and the solution of the problem would be only $B=0$ which is the trivial case excluded from the theorem. So the left-hand side of equation (5.16) is a function of B which is bounded by 1 in absolute value and fast oscillating about the point $B=0$. The right-hand side of (5.16) is an even function increasing for $B>0$. Therefore, there exists a countable number of non-vanishing different solutions $\{B_n\}$ of the equation (5.16) within the interval $|B|\le\sqrt{\frac{1}{\rho^2}-1}$ with a limit point at the origin. The geodesics $(z(s),w(s)$ with the parameters B_n and $\theta(B_n)$ start from the point $(1,0,0,0)$ with the velocities $(\dot z(0),\dot w(0))=(0,e^{i\theta(B_n)})$ in different directions, and therefore, they sweep out different geometric loci about the initial point, which finishes the proof. $\square$

This theorem reveals a clear difference of subRiemannian geodesics on the sphere from those for the Heisenberg group. In the latter case the number of geodesics joining the origin with a point neither from the vertical axis nor from the horizontal plane is finite.

5.3. **Hyperspherical coordinates.** Let us use the hyperspherical coordinates to find geodesics in with conconstant velocity coordinates.

$$\begin{aligned} x_1 + ix_2 &= e^{i\xi_1}\cos\eta, \\ x_3 + ix_4 &= e^{i\xi_2}\sin\eta, \qquad \eta \in [0,\pi/2], \quad \xi_1,\xi_2 \in [0,2\pi). \end{aligned} \tag{5.17}$$

The horizontal coordinates are written as

$$\begin{aligned} a &= \dot\eta\cos(\xi_1-\xi_2) + (\dot\xi_1+\dot\xi_2)\sin(\xi_1-\xi_2)\frac{\sin 2\eta}{2}, \\ b &= -\dot\eta\sin(\xi_1-\xi_2) + (\dot\xi_1+\dot\xi_2)\cos(\xi_1-\xi_2)\frac{\sin 2\eta}{2}, \\ c &= \dot\xi_1\cos^2\eta - \dot\xi_2\sin^2\eta. \end{aligned}$$

The horizontality condition in hyperspherical coordinates is

$$\dot\xi_1\cos^2\eta - \dot\xi_2\sin^2\eta = 0.$$

The horizontal sphere (5.9) is obtained from the parametrization (5.17), if we set $\xi_1 = 0$, $\xi_2 = \psi$, $\eta = s$. We have

$$a^2 + b^2 = 1 = \dot\eta^2 \quad\Longrightarrow\quad a = \cos\psi, \quad b = \sin\psi.$$

The vertical line is obtained from the parametrization (5.17) setting $\eta = 0$, $\xi_1 = s$.

If write the vector fields N, Z, X, Y in the hyperspherical coordinates, we obtain

$$N = -2\cot\left(2\eta\partial_\eta\right), \qquad Z = \partial_{\xi_1} - \partial_{\xi_2},$$

$$X = \sin(\xi_1-\xi_2)\tan\eta\,\partial_{\xi_1} + \sin(\xi_1-\xi_2)\cot\left(\eta\partial_{\xi_2}\right) + 2\cos(\xi_1-\xi_2)\partial_\eta,$$

and

$$Y = \cos(\xi_1-\xi_2)\tan\eta\,\partial_{\xi_1} + \cos(\xi_1-\xi_2)\cot\left(\eta\partial_{\xi_2}\right) - 2\sin(\xi_1-\xi_2)\partial_\eta.$$

In this parametrization the similarity with the Heisenberg group can be shown. Two horizontal vector fields X, Y produce as a commutator the constant vector field Z that orthogonal to the horizontal vector fields at each point of the manifold. In hyperspherical coordinates it is easy to see that the form $\omega = \cos^2\eta\, d\xi_1 - \sin^2\eta\, d\xi_2$ that defines the horizontal distribution is a contact form because of

$$\omega \wedge d\omega = \sin(2\eta)\, d\eta \wedge d\xi_1 \wedge d\xi_2 = 2dV,$$

where dV is the volume form. The sub-Laplacian is the following

$$\frac{1}{2}(X^2+Y^2) = \frac{1}{2}(\tan^2\eta\cdot\partial_{\xi_1}^2 + \cot^2\eta\cdot\partial_{\xi_2}^2 + 4\partial_\eta^2 + 2\partial_{\xi_1}\partial_{\xi_2}).$$

The Hamilton is

$$H(\xi_1,\xi_2,\eta,\psi_1,\psi_2,\theta) = \frac{1}{2}(\tan^2\eta\cdot\psi_1^2 + \cot^2\eta\cdot\psi_2^2 + 4\theta^2 + 2\psi_1\psi_2).$$

The Hamilton system is

$$\begin{aligned}
\dot{\xi}_1 &= \frac{\partial H}{\partial \psi_1} = \psi_1 \tan^2\eta + \psi_2 \\
\dot{\xi}_2 &= \frac{\partial H}{\partial \psi_2} = \psi_2 \cot^2\eta + \psi_1 \\
\dot{\eta} &= \frac{\partial H}{\partial \theta} = 4\theta \\
\dot{\psi}_1 &= -\frac{\partial H}{\partial \xi_1} = 0 \\
\dot{\psi}_2 &= -\frac{\partial H}{\partial \xi_2} = 0 \\
\dot{\theta} &= -\frac{\partial H}{\partial \eta} = -\psi_1^2 \frac{\tan\eta}{\cos^2\eta} + \psi_2^2 \frac{\cot\eta}{\sin^2\eta}.
\end{aligned}$$

We solve the Hamiltonian system for the following initial dates: $\eta(0) = 0$, $\xi(0) = 0$, $\xi_2(0) = \xi_2(0)$, $\psi_1(0) = \psi_1$, $\psi_2(0) = \psi_2$, $\theta(0) = \frac{\dot{\eta}(0)}{4} = \theta_0$.

We see that ψ_1 and ψ_2 are constant. From the third and the last equations, we have

$$\ddot{\eta} = -4\psi_1^2 \frac{\sin\eta}{\cos^3\eta} + 4\psi_2^2 \frac{\cos\eta}{\sin^3\eta}$$

$$\dot{\eta}\, d\dot{\eta} = \Big(-4\psi_1^2 \frac{\sin\eta}{\cos^3\eta} + 4\psi_2^2 \frac{\cos\eta}{\sin^3\eta}\Big)\, d\eta$$

$$\dot{\eta}^2 = C - 4\frac{\psi_1^2 \sin^2\eta + \psi_2^2 \cos^2\eta}{\cos^2\eta \sin^2\eta},$$

and

$$C\cos^2\eta(0)\sin^2\eta(0) = \dot{\eta}^2(0)\cos^2\eta(0)\sin^2\eta(0) + 4\psi_1^2 \sin^2\eta(0) + 4\psi_2^2\cos^2\eta(0).$$

Since $\eta(0) = 0$, we have $\psi_2 = 0$. If $\psi_1 = 0$, then $\dot{\xi}_1 = \dot{\xi}_2 = 0$ and $\dot{\eta} = 4\theta_0$, and we obtain the variety of the trivial geodesics (5.9) up to the parametrization $s \mapsto 4\theta_0 s$. To fined other geodesics we suppose that $\psi_1 \neq 0$. Moreover, we notice that $C = \dot{\eta}(0) + 4\psi_1 > 0$.

We continue to solve the Hamiltonian system finding $\eta(s)$.

$$\frac{\cos\eta\, d\eta}{\sqrt{C\cos^2\eta - 4\psi_1^2}} = ds.$$

Denoting $\sin\eta = p$, we have

$$\frac{dp}{\sqrt{-Cp^2 + C - 4\psi_1^2}} = ds. \tag{5.18}$$

Integrating (5.18) from 0 to s, we have

$$s + A = \frac{1}{\sqrt{|C|}} \sin^{-1}\Big(\sqrt{\frac{C}{C - 4\psi_1^2}} \sin\eta(s)\Big),$$

where A we find setting $s=0$. Then $A=\frac{1}{\sqrt{|C|}}\sin^{-1}0=0$. We express

$$\sin^2\eta(s)=\frac{C-4\psi_1^2}{C}\sin^2(\sqrt{C}s). \tag{5.19}$$

From the Hamiltonian system, we find

$$\xi_2(s)=\psi_1 s+\xi_2(0) \tag{5.20}$$

and

$$\dot{\xi}_1=\psi_1\frac{\sin^2\eta(s)}{1-\sin^2\eta(s)}=\psi_1\frac{\sin^2(\sqrt{C}s)}{a+\sin^2(\sqrt{C}s)},\quad a=\frac{C}{C-4\psi_1^2}.$$

It gives

$$\xi_1(s)=-\psi_1 s+\frac{\psi_1}{2|\psi_1|}\tan^{-1}\Big[\frac{2|\psi_1|}{\sqrt{C}}\tan(\sqrt{C}s)\Big]. \tag{5.21}$$

Let us suppose that the geodesics parametrized on the interval $[0,1]$. If the initial point and the finite point are on the vertical line: $\eta(0)=\eta(1)=0$, then

$$0=\sin^2\eta(1)=\frac{C-4\psi_1^2}{C}\sin^2(\sqrt{C})\quad\Rightarrow\quad C=\pi^2n^2.$$

The value of C and $\xi_1(1)$ gives us the value of $\xi_2(1)$. Setting $C=\pi^2n^2$ into the equation for $\xi_1(1)$ we find $\psi_1=-\xi_1(1)$. Then

$$\xi_2(1)=-\xi_1(1)+\xi_2(0).$$

The finite point on the vertical line corresponds to the value $\xi_1(1),\eta(1)=0$ and $\xi_2=-\xi_1(1)+\xi_2(0)$.

We also note that the square of the velocity

$$|v|^2=\dot{\eta}^2(s)+(\dot{\xi}_1(s)+\dot{\xi}_2(s))^2\frac{\sin^2 2\eta(s)}{2}$$

is constant along geodesics. Taking $\eta(0)=0$, we obtain

$$|v|^2=\dot{\eta}^2(0)=C-4\psi_1^2.$$

In the case when a geodesic terminated at the vertical line at $\xi_1(1)$, we obtain their lengths

$$\ell_n=\sqrt{C-4\psi_1^2}=\sqrt{\pi^2n^2-4\xi_1^2(1)}.$$

We see that this result coincides with the result given by Theorem (5.2).

Using the hyperspherical coordinate we can find the number geodesics joining $\xi_1=\xi_2=\eta=0$ with an arbitrary point ξ_1, ξ_2 and η. We formulate it in the following theorem.

Theorem 5.4. *Let A be a point with the coordinate $A=(\xi_1,\xi_2,\eta)$, ξ_1,ξ_2,η are different from zero, then there are countably many geodesics γ_n connecting $O=(0,0,0)$ with A. The square of length ℓ^2 of a geodesic satisfies to the equation*

$$\ell^2\cos^2(\sqrt{\ell^2+4\xi_2})=\frac{4\xi_2^2\sin^2\eta}{\tan^2\big(2(\xi_1+\xi_2)\big)}. \tag{5.22}$$

The geodesics have the following parametric equations

$$\begin{aligned}
\big(\eta(s)\big)_n &= \sqrt{\frac{l_n^2}{\ell_n^2+4\xi_2^2}}\sin^2\big(s\sqrt{\ell_n^2+4\xi_2^2}\big),\\
\big(\xi_2(s)\big)_n &= \xi_2 s,\\
\big(\xi_1(s)\big)_n &= -\xi_2 s + 2\frac{\xi_2}{|\xi_2|}\tan^{-1}\big(\frac{2|\xi_2|}{\sqrt{\ell_n^2+4\xi_2^2}}\tan\sqrt{\ell_n^2+4\xi_2^2}\big),
\end{aligned}$$

where $n\in\mathbb{N}$ and ℓ_n one of the solutions of the equation (5.22).

Proof. Suppose that the geodesics parametrized on the interval $[0,1]$ and they start from the point $\xi_1(0)=0$, $\eta(0)=0$, $\xi_2(0)=0$ and terminate at a point $\xi_1(1)=\xi_1$, $\eta(1)=\eta$, $\xi_2(1)=\xi_2$. Then $\psi_1=\xi_2$ from (5.20) and

$$\tan^2\big(2(\xi_1+\xi_2)\big) = \frac{4\xi_2^2}{C}\tan^2\sqrt{C}$$

from (5.21), where we assumed that $\xi_2>0$ for the moment. The formula (5.21) gives

$$\sin^2\eta = \frac{C-4\xi_2^2}{C}\sin^2\sqrt{C}.$$

Comparing two last expressions, we obtain

$$(C-4\xi_2^2)\cos^2\sqrt{C} = \frac{4\xi_2^2\sin^2\eta}{\tan^2\big(2(\xi_1+\xi_2)\big)}. \tag{5.23}$$

We observe that the length of a geodesic γ parametrized on the interval $[0,1]$ can be calculated by the following

$$\ell(\gamma) = \int_0^1 |v(0)|\,ds = |\dot\eta^2(0)| = \sqrt{C-4\xi_2^2}.$$

From this and (5.23), we obtain (5.22)

The function $\ell^2\cos^2(\sqrt{l^2+4\xi_2})$ oscillates with increasing amplitude, therefore the graph of the function $\ell^2\cos^2(\sqrt{\ell^2+4\xi_2})$ has infinitely many intersections with the given value $m(\xi_1,\xi_2,\eta) = \frac{4\xi_2^2\sin^2\eta}{\tan^2\big(2(\xi_1+\xi_2)\big)}$, see Figure (2). We conclude, that we have infinitely many geodesics which lengths satisfy the equation (5.22). To find the expressions (5.23) we replace C and ψ_1 by known values. □

Remark. Theorem 5.4 is analogue of Theorem 5.3. The difference is such that in Theorem 5.3 the number of geodesics is parametrized by the angle of the initial velocity and this show that the geodesics have different locus. In Theorem 5.4 we can easily to know the lengths of geodesics, that are given by (5.22).

5.4. **Hopf fibration.** There is a close relation between the subRiemannian sphere S^3 and the Hopf fibration. Let $\mathbb{S}^2$ and $\mathbb{S}^3$ be unit 2-dimensional and 3-dimensional sphere respectively. We remind that the Hopf fibration is a principal circle bundle over two-sphere given by the map $h:\mathbb{S}^3\to\mathbb{S}^2$:

$$h(x_1,x_2,x_3,x_4) = ((x_1^2+x_2^2)-(x_3^2+x_4^2), 2(x_1x_4+x_2x_3), 2(x_2x_4-x_1x_3)).$$

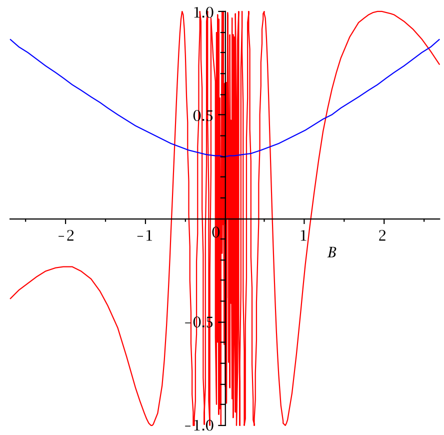

FIGURE 2. Solutions of the equation $\ell^2 \cos^2(\sqrt{\ell^2+4\xi_2}) = m(\xi_1,\xi_2,\eta)$

Another way to define the Hopf fibration is to write

$$h(q) = qiq^* \in \mathbb{S}^2, \quad q \in \mathbb{S}^3, \quad i = (0,1,0,0).$$

The fiber through the unity of the group $(1,0,0,0)$ has equation $(\cos\theta, \sin\theta, 0, 0)$ and as we see coincides with the equation of the vertical line at this point. The sphere $\mathbb{S}^2$ represents the horizontal “plane” filled out by the geodesics with constant horizontal coordinates.

Definition 5.1. Let $Q \to M$ be a principle G-bundle with the horizontal distribution $\mathscr{D}$ on Q. A subRiemannian metric on Q that has distribution $\mathscr{D}$ and it is invariant under the action of G is called a metric of bundle type.

In our situation $\mathbb{S}^3 \to \mathbb{S}^2$ is a principle $\mathbb{S}^1$-bundle given by Hopf map. The subRiemannian metric on the distribution $\mathscr{D} = \text{span}\{X,Y\}$ was defined as the restriction of the euclidean metric $\langle\cdot,\cdot\rangle$ from $\mathbf{R}^4$ and we used the same notation $\langle\cdot,\cdot\rangle$ for subRiemannian metric. We state the following.

Proposition 5.1. *The subRiemannian metric $\langle\cdot,\cdot\rangle$ on $\mathbb{S}^3$ is a metric of bundle type.*

Proof. The action of the group $\mathbb{S}^1$ on $q = (x_1,x_2,x_3,x_4) \in \mathbb{S}^3$ can be written as $q \circ e^{it}$, $e^{it} = (\cos t + i\sin t) \in \mathbb{S}^1$, $t \in [0,2\pi)$, where $\circ$ is the quaternion multiplication. If we write $q = (e^{i\xi_1}\cos\eta, e^{i\xi_2}\sin\eta)$, $\tilde{q} = q \circ e^{it}$ then $\tilde{q} = (e^{i(\xi_1+t)}\cos\eta, e^{i(\xi_2-t)}\sin\eta)$. To show that the metric $\langle\cdot,\cdot\rangle$ is of bundle type we have to show that the metric is invariant under the action of the group $\mathbb{S}^1$. The metric $\langle\cdot,\cdot\rangle$ at any q is given by the matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos^2\eta & 0 \\ 0 & 0 & \sin^2\eta \end{bmatrix}$$

and it is easy to see that it is invariant under the action $\tilde{q} = q \circ e^{it}$. $\square$

6. Heat Kernel of the Elliptic and Subelliptic Laplacian on $\mathbb{S}^3$

Let us write the Laplace and subLaplace operators on $\mathbb{S}^3$ first. As usual, we set

$$Z_k = \sum_{j=1}^{n+1} a_{k_j} \frac{\partial}{\partial z_j}, \qquad k = 1, \dots, n+1$$

be a holomorphic vector field on $\mathbf{C}^{n+1}$. Then it is easy to see that

$$\left\{\sqrt{2}Z_1, \dots, \sqrt{2}Z_{n+1}\right\}$$

is an orthonormal basis of $T^{(1,0)}(\mathbf{C}^{n+1})$ with respect to the Euclidean metric. As usual, set $z_k = x_k + ix_{k+n+1}$, $k = 1, \dots, n+1$. Denote Z_k^* the differential operator adjoint to Z_k with respect to the Euclidean volume form

$$dx = dx_1 \wedge dx_2 \wedge \cdots \wedge dx_{n+1} = \frac{(-1)^{n^2-1}}{2^{n+1}} dz_1 \wedge d\bar{z}_1 \wedge \cdots \wedge dz_{n+1} \wedge d\bar{z}_{n+1}.$$

The Laplacian on $\mathbf{C}^{n+1}$ is given by

$$\Delta = -\sum_{k=1}^{n+1} \left(Z_k^* Z_k + \bar{Z}_k^* Z_k\right) = 2\sum_{k=1}^{n+1} \frac{\partial}{\partial z_k}\frac{\partial}{\partial \bar{z}_k}. \tag{6.1}$$

It is known that the Cauchy-Riemann subLaplacian on the unit sphere $\mathbb{S}^{2n+1}$ is given by

$$\Delta_{cr} = -\sum_{k=1}^{n} \left(W_k^* W_k + \bar{W}_k^* W_k\right)\Big|_{\mathbb{S}^{2n+1}}, \tag{6.2}$$

where $\left\{\sqrt{2}W_1, \dots, \sqrt{2}W_n\right\}$ is a local orthonormal basis of those holomorphic vector fields on $\mathbf{C}^{n+1}$ which are perpendicular to the holomorphic vector field

$$N = \sum_{j=1}^{n+1} \frac{z_j}{|z|}\frac{\partial}{\partial z_j} \tag{6.3}$$

with $|z| = z \cdot \bar{z}$. It is obvious that W_k and $\bar{W}_k$, $k = 1, \dots, n$ are tangential to $\mathbb{S}^{2n+1}$. Using vector N and W_k and $\bar{W}_k$, we may rewrite the Laplace operator Δ as follows (see [13] and [19]):

$$\Delta = -2\mathrm{Re}\Big(N^*N + \sum_{k=1}^{n} W_k^* W_k\Big). \tag{6.4}$$

Let Δ_{cr} be the restriction of Δ to $C^\infty(\mathbb{S}^{2n+1})$. Thus, by (6.2) and (6.4) tells us that

$$\Delta_{cr} = \Delta_{\mathbb{S}} + 2\mathrm{Re}\big(N^*N\big)\Big|_{C^\infty(\mathbb{S}^{2n+1})}.$$

Thus Δ_{cr} is globally defined on $\mathbb{S}^{2n+1}$ and by (6.1) and (6.3), we know that it is fully explicitly defined. Notice that the operator $\Delta_{\mathbb{S}}$ is elliptic. However, Δ_{cr} is only a sub-elliptic.

6.1. **Heat kernel for elliptic Laplacian on $\mathbb{S}^3$.** Recall the heat kernel for the Laplace-Beltrami operator:

$$\Delta = -\frac{1}{2}\sum_{j=1}^{n} X_j^* X_j = \frac{1}{2}\sum_{j=1}^{n} X_j^2 + \cdots .$$

Here $X_1,\ldots,X_n$ represent n linearly independent vector fields on an n-dimensional manifold $\mathscr{M}_n$. When $\mathscr{M}_n$ is compact without boundary or relatively compact domain in $\mathbf{R}^n$, Greiner (see [19]) showed that the heat kernel takes the form

$$P(x,y;t) = \frac{1}{(2\pi t)^{\frac{n}{2}}} e^{-\frac{d^2(x,y)}{2t}} \big(a_0 + a_1 t + a_2 t^2 + \cdots\big).$$

The a_j's are functions of x and y. Note that

$$\frac{\partial}{\partial t}\Big(\frac{d^2(x,y)}{2t}\Big) + \frac{1}{2}\sum_{j=1}^{n}\Big(X_j \frac{d^2(x,y)}{2t}\Big)^2 = 0,$$

i.e., $\frac{d^2(x,y)}{2t}$ is a solution of the *Hamilton-Jacobi equation*.

We are interested in finding the solving kernels for the operators $\frac{\partial}{\partial t} - \Delta_{cr}$. It is reasonable to expect the kernel has the form:

$$P_{cr}(x,y;t) = \frac{c}{t^{\alpha}} e^{-S(x,y;t)} \qquad \text{for some suitable } \alpha.$$

See, for example, Beals-Gaveau-Greiner [3, 4] and Calin-Chang-Greiner [9].

The *modified complex action* function $S(x,y;t)$ plays the role of $\frac{d^2(x,y)}{2t}$ in the Riemannian cases and satisfies the Hamilton-Jacobi equation

$$\frac{\partial S}{\partial t} + H\Big(x, y, \frac{\partial S}{\partial x}, \frac{\partial S}{\partial y}\Big) = 0.$$

In our case, we need to find a solution S of the following equation induced by $\frac{\partial}{-}\Delta_{cr}$:

$$0 = \frac{\partial S}{\partial t} + \frac{1}{2}\Big[\Big(\frac{\partial S}{\partial \theta}\Big)^2 + \frac{1}{\cos^2\theta}\Big(\frac{\partial S}{\partial \phi}\Big)^2\Big]. \tag{6.5}$$

This is a first order partial differential equation whose solution can be found in the form of an action integral S along the bicharacteristics $\eta + H$; thus

$$H\big(\theta,\omega,\tau\big) = \frac{1}{2}\Big\{\omega^2 + \frac{\tau^2}{\cos^2\theta} - \tau^2\Big\}.$$

In fact, $\theta(s)$, $\phi(s)$ and $\omega(s)$ may be obtained from the reduced Hamiltonian

$$H = \frac{1}{2}\Big(\omega^2 + \tau^2\cdot\tan^2\theta\Big),$$

as solutions of

$$\dot{\theta} = \omega = \pm\sqrt{2H - \tau^2\cdot\tan^2\theta}, \quad \text{and} \quad \dot{\phi} = \tau\cdot\tan^2\theta. \tag{6.6}$$

In other words, (6.6) yields

$$\pm\frac{d\theta}{ds} = \sqrt{2H - \frac{2\tau^2}{1+\cos(2\theta)}}$$
$$= \frac{|\tau|k}{\sin(2\theta)}\frac{k^2-1}{k^2}\sqrt{1-\left(\frac{k^2}{k^2-1}\cos(2\theta)-\frac{1}{k^2-1}\right)^2},$$

with $k^2 = \frac{2H}{\tau^2}$. In other words, one has

$$\frac{d\left(\frac{k^2}{k^2-1}\cos(2\theta)-\frac{1}{k^2-1}\right)}{\sqrt{1-\left(\frac{k^2}{k^2-1}\cos(2\theta)-\frac{1}{k^2-1}\right)^2}} = \pm 2|\tau|k\,ds.$$

This yields

$$\frac{k^2}{k^2-1}\cos(2\theta)-\frac{1}{k^2-1} = \sin\Big(\pm 2|\tau|k(s-s_0)\Big), \tag{6.7}$$

with some s_0 still to be determined. When $s=0$, (6.7) becomes

$$1 = \sin\Big(\pm 2|\tau|ks_0\Big).$$

Using an addition formula and expanding the sine function in (6.7), we have

$$\frac{k^2}{k^2-1}\cos(2\theta)-\frac{1}{k^2-1} = \cos\big(2|\tau|ks\big). \tag{6.8}$$

Subtracting 1 from each side leads of (6.8), we obtain

$$\sin^2\theta(s) = \left(1-\frac{\tau^2}{2H}\right)\sin^2\big(\sqrt{2H}s\big). \tag{6.9}$$

On the other hand, the second equation $\dot\phi = \tau\cdot\tan^2\theta$ in (6.6) and

$$\cos^2\big(\theta(s)\big) = \cos^2\big(\sqrt{2H}s\big)+\frac{\tau^2}{2H}\sin^2\big(\sqrt{2H}s\big) = \frac{1+\frac{\tau^2}{2H}\tan^2\big(\sqrt{2H}s\big)}{1+\tan^2\big(\sqrt{2H}s\big)}$$

imply that

$$d\phi = \frac{d\big(\frac{\tau}{\sqrt{2H}}\tan\big(\sqrt{2H}s\big)\big)}{1+\big(\frac{\tau}{\sqrt{2H}}\tan\big(\sqrt{2H}s\big)\big)^2}.$$

Assume that $\phi(0)=0$, then the ϕ-component of the bicharacteristic of $\Delta_{\mathbb{S}}$ is given by

$$\phi(s) = \tan^{-1}\left(\frac{\tau}{\sqrt{2H}}\tan\big(\sqrt{2H}s\big)\right), \tag{6.10}$$

which can be extended to all $s>0$.

Now we may choose parameters H and τ so that $\phi(1)=\phi$ and $\theta(1)=\theta$. Then (6.9) and (6.10) imply that

$$\frac{\sin^2\theta}{\sin^2(\sqrt{2H}t)} = 1-\frac{\tau^2}{2H} = 1-\frac{\tan^2\phi}{\tan^2(\sqrt{2H}t)}.$$

This implies that

$$\sin^2\big(\sqrt{2H}t\big)-\cos^2\big(\sqrt{2H}t\big)\cdot\tan^2\phi = \sin^2\theta.$$

Substracting 1 from each sides of the above equation gives us

$$\cos\left(\sqrt{2Ht}\right) = \cos\theta\cdot\cos\phi.$$

Thus

$$\sqrt{2Ht} = \cos^{-1}\left(\cos\theta\cdot\cos\phi\right)+2k\pi = \zeta+2k\pi, \qquad k\in\mathbf{Z},$$

where $\cos\zeta = \cos\theta\cdot\cos\phi$; the angle ζ subtends the arc from $(1,0,\ldots,0)$ to the point z. Therefore, the equation $\dot S=\omega\cdot\dot\theta+\tau\cdot\dot\phi-H$ implies that $\dot S=H=\frac{E^2}{2}$. Hence we have

$$S = Ht = \frac{\left(\cos^{-1}(\cos\theta\cdot\cos\phi)+2k\pi\right)^2}{2t} = \frac{(\zeta+2k\pi)^2}{2t}, \qquad k\in\mathbf{Z}. \tag{6.11}$$

It is easy to check that S satisfies the Hamilton-Jacobi equation (6.5). In conclusion, we have the following theorem for elliptic Laplacian on $\mathbb{S}^3$.

Theorem 6.1. *Given $z,w\in\mathbb{S}^3$, let ζ be the angle which subtends the arc that joins z and w on a great circle, $0\le\zeta<\pi$. Then the heat kernel $P_{\mathbb{S}}$ of $\Delta_{\mathbb{S}}$ on $\mathbb{S}^3$ is given by*

$$P_{\mathbb{S}}(z,w,;t) = \frac{e^{\frac{t}{2}}}{(2\pi t)^{\frac{3}{2}}}\sum_{k\in\mathbf{Z}} e^{-\frac{(\zeta+2k\pi)^2}{2t}}\varphi\left(\zeta+2k\pi\right),$$

where

$$\varphi(\zeta) = \left(\frac{\zeta}{\sin\zeta}\right).$$

6.2. **Heat kernel for subLaplacian on $\mathbb{S}^3$.** Let $P_{cr}(z,w;t)$ be the heat kernel of Δ_{cr} with pole at z which satisfies

$$\begin{aligned}\frac{\partial P_{cr}}{\partial t}-\Delta_{cr}P_{cr} &=0\\ \lim_{t\to 0^+} P_{cr} &=\delta(z-w).\end{aligned} \tag{6.12}$$

Since Δ_{cr} is invariant with respect to complex (unitary) rotations of $\mathbf{C}^{n+1}$, it follows that $P_{cr}(z,w;t)$ is a function depending only on one complex variable

$$z\cdot\bar w = \sum_{k=1}^{n+1} z_k\cdot\bar w_k = z_1\cdot\bar w_1+\cdots+z_{n+1}\cdot\bar w_{n+1} = \cos\theta\cdot e^{i\phi}$$

and its complex conjugate $\bar z\cdot w$, or of 2 real variables θ and ϕ with $0\le\theta\le\frac{\pi}{2}$ and $0\le\phi\le 2\pi$. Thus, we may write

$$P_{cr}(z,w;t) = P_{cr}\left(\left|z\cdot w\right|,\arg(z\cdot w);t\right) = P_{cr}\left(\cos\theta,\phi;t\right).$$

In particular, we may assume that $z=(1,0,\ldots,0)$, and then $w_1 = z\cdot\bar w = \cos\theta\cdot e^{i\phi}$. One extends θ, ϕ to a complex system of spherical coordinates $(\theta,\phi)=(\theta_1,\ldots,\theta_n,\phi_1,\ldots,\phi_n,\phi_{n+1})$ on $\mathbb{S}^{2n+1}$, and then w can be identified as (θ,ϕ). Hence, Δ_{cr} restricted to functions of θ and ϕ yields the operator

$$\mathscr{L}_{cr} = \frac{1}{2}\frac{\partial^2}{\partial\theta^2}+\left((n-1)\cot\theta+\cot(2\theta)\right)\frac{\partial}{\partial\theta}+\frac{1}{2}\tan^2\theta\cdot\frac{\partial^2}{\partial\phi^2}. \tag{6.13}$$

It is obvious that $\mathscr{L}_{cr}$ is not elliptic since $\frac{1}{2}\tan^2\theta\cdot\frac{\partial^2}{\partial\phi^2}$ vanishes at $\theta=0$. However, we know that

$$\Big[\frac{\partial}{\partial\theta},\tan\theta\frac{\partial}{\partial\phi}\Big] = \frac{1}{\cos^2\theta}\cdot\frac{\partial}{\partial\phi},$$

which does not vanish at $\theta=0$, hence $\mathscr{L}_{cr}$ is a sub-elliptic operator of step 2.

We can follow the idea in Greiner's celebrated paper [19] to look for a solution S_{cr} of the Hamilton-Jacobu equation induced by the sub-elliptic heat equation $\frac{\partial}{\partial t}-\mathscr{L}_{cr}$:

$$0 = \frac{\partial S_{cr}}{\partial t}+\frac{1}{2}\left[\Big(\frac{\partial S_{cr}}{\partial\theta}\Big)^2+\big(1-\frac{1}{\cos^2\theta}\big)\Big(\frac{\partial S_{cr}}{\partial\phi}\Big)^2\right], \tag{6.14}$$

in the form of an action integral along the bicharacteristics $\eta+H$ with

$$H(\theta,\omega,\tau) = \frac{1}{2}\Big(\omega^2+\tan^2\theta\cdot\tau^2\Big),$$

as solutions of

$$\dot{\theta} = -\omega = \sqrt{2H-\tan^2\theta\cdot\tau^2} \qquad \text{and} \qquad \dot{\phi} = \tan^2\theta\cdot\tau.$$

The first equation above yields

$$\begin{aligned}\pm\frac{d\theta}{ds} &= \sqrt{2H+\big(1-\frac{2}{1+\cos(2\theta)}\big)\tau^2}\\ &= \frac{|\tau|}{\sin(2\theta)}\frac{k^2}{\sqrt{k^2+1}}\sqrt{1-\big(\frac{k^2+1}{k^2}\cos(2\theta)-\frac{1}{k^2}\big)^2},\end{aligned}$$

where $k^2=\frac{2H}{\tau^2}$. Now by standard method, one has

$$\frac{k^2+1}{k^2}\cos(2\theta)-\frac{1}{k^2} = \sin\Big(\pm 2|\tau|\sqrt{k^2+1}(s-s_0)\Big), \tag{6.15}$$

with some s_0 that we need to determined later. When $s_0=0$, (6.15) becomes

$$1 = \sin\Big(\pm 2|\tau|\sqrt{k^2+1}s\Big).$$

Using the addition formula of the sine function in (6.15) gives us

$$\frac{k^2+1}{k^2}\cos(2\theta)-\frac{1}{k^2} = \cos\Big(2|\tau|\sqrt{k^2+1}s\Big).$$

Then substract 1 from each side leads to

$$\sin^2\big(\theta(s)\big) = \frac{E^2}{E^2+\tau^2}\sin^2\big((E^2+\tau^2)s\big).$$

Hence,

$$\begin{aligned}&1-\sin^2\big(\theta(s)\big)\\ =&1-\frac{E^2}{E^2+\tau^2}\sin^2\big((E^2+\tau^2)s\big)\\ =&\big(1+\frac{\tau^2}{E^2+\tau^2}\tan^2\big(\sqrt{E^2+\tau^2}s\big)\cdot\big(1+\tan^2\big(\sqrt{E^2+\tau^2}s\big)\big)^{-1}.\end{aligned}$$

Combining with $\dot\phi = \tan^2\theta\cdot\tau$, we conclude that

$$\begin{aligned} d\phi &= -\tau\, ds + \frac{\tau\, ds}{\cos^2\theta(s)} \\ &= -\tau\, ds + d\Big(\frac{\tau}{\sqrt{E^2+\tau^2}}\tan\big(\sqrt{E^2+\tau^2}\,s\big)\Big) \\ &\quad\times\Big(1+\frac{\tau}{\sqrt{E^2+\tau^2}}\tan\big(\sqrt{E^2+\tau^2}\,s\big)\Big)^{-2}, \end{aligned}$$

which implies that

$$\phi(s)-\phi(0) = \tan^{-1}\Big(\frac{\tau}{\sqrt{E^2+\tau^2}}\tan\big(\sqrt{E^2+\tau^2}\,s\big)\Big) - \tau s.$$

In order to find the action function $S = Ht = \frac{E^2}{2}t$, we may assume that $\phi(0)=0$. Then at $s=t$, one has

$$\tan\big(\phi+\tau t\big) = \frac{\tau}{\sqrt{E^2+\tau^2}}\tan\big(\sqrt{E^2+\tau^2}\,t\big). \tag{6.16}$$

Applying standard argument, we have

$$S = Ht = \frac{-(\tau t)^2 + \big(\cos^{-1}(\cos\theta\cdot\cos(\phi+\tau t)) + 2k\pi\big)^2}{2t}.$$

Since τ does not appear in the final form of P_{cr}, so we may use an old trick to sum over $v=\tau t$ where τ is treated as a running parameter, even though $\phi(0)=0$ in (6.16) is supposed to fix it.

Denote

$$\gamma(v) = -v^2 + \big(\cos^{-1}(\cos\theta\cdot\cos(\phi+v))\big)^2,$$

then S will be positive on the imaginary axis so we set $g(v) = \gamma(-iv)$ and obtain

$$\begin{aligned} g(v) &= v^2 + \big(\cos^{-1}(\cos\theta\cdot\cos(\phi-iv))\big)^2 = v^2+\nu^2 \\ &= v^2 - \big(\cosh^{-1}(\cos\theta\cdot\cosh(v+i\phi))\big)^2 = v^2-\rho^2, \end{aligned}$$

where

$$\cos\nu = \cos\theta\cdot\cos(\phi-iv) = \cosh\rho, \qquad \rho = i\nu.$$

In summary, we have the following theorem.

Theorem 6.2. *The heat kernel P_{cr} of Δ_{cr} on $\mathbb{S}^3$ is given by*

$$P_{cr}(z,w,;t) = \frac{e^{\frac{t}{2}}}{(2\pi t)^2}\sum_{k\in\mathbf{Z}}\int_{-\infty}^{+\infty} e^{-\frac{(v^2-(\rho+2ki\pi))^2}{2t}}\Big(\frac{\rho}{\sinh\rho}\Big)\,dv,$$

where $(\theta,\phi)\neq(0,\pi)$.

7. Real-world Applications

We conclude with a represetative real-world application that utilizes the kernel and manifold concept, specifically clinical data modeling. Modeling patient clinical records, such as Electronic Health Records (EHRs), is complicated for multiple reasons. Among complicating factors are those that arise from both the need to compare patient records across patients and the necessity to model the duration between medical events, where the duration and the impact of that duration varies.

In [36, 37], a kernel method to model graph-based medical data was proposed. The disclosed approach provides similarity measurements across patient records. In [11, 38], the authors further discussed how distance affects EHR representation learning. They demonstrated the performance difference between Euclidean and cosine distance-defined kernel functions. In most cases, cosine rather than Euclidean distance performs better due to its spherical properties. The degree of improvement using cosine rather than Euclidean distance roughly correlated to the duration of the affliction treated, the longer the affliction the greater was the improvement. Note, however, that for predominantly one-shot treatment ailments, such as those diseases treated by a single regimen of antibiotics, Euclidean distance often performed slightly, but rarely statistically significantly, better. The aforementioned approach, trained and relying on nationalized health records, is in limited clinical use and is disclosed in [16].

Most recently, in [39], the authors identified the representation difference between Euclidean and cosine distance manifolds and how they benefit self-supervised contrastive learning [25]. Such difference enhances self-supervised learning by discovering robust representations of complex patient medical records across different geometries. The geometry properties are built along with Euclidean and cosine distance manifolds.

Remark. We may consider a manifold with negative curvature also. In that case, we need to approximate the manifold in a local neighborhood using a hyperbolic space:

$$\mathbf{h}^n = \{x \in \mathbf{R}^n : \, |x| < 2\} \quad \text{with} \quad g_{\mathbf{h}} = 4\frac{\sum_{j=1}^n dx_j \otimes dx_j}{(4-|x|^2)^2}.$$

One may construct the heat kernel $K_n^{\mathbf{h}}(t,\rho(x,y))$ on the hyperbolic space as follows:

$$K_n^{\mathbf{h}}(t,\rho(x,y)) = \frac{e^{-\ell^2 t}}{(2\pi)^\ell (4\pi t)^{\frac{1}{2}}}\left(\frac{-1}{\sinh\rho}\frac{\partial}{\partial\rho}\right)^\ell e^{-\frac{\rho^2}{4t}}, \; n = 2\ell+1$$

and

$$K_n^{\mathbf{h}}(t,\rho(x,y)) = \frac{e^{-(2\ell_1)^2 t/4}}{2(2\pi)^\ell (2\pi t)^{\frac{3}{2}}}\left(\frac{-1}{\sinh\rho}\frac{\partial}{\partial\rho}\right)^\ell \int_r^\infty \frac{se^{-\frac{s^2}{4t}}}{(\cosh(s)-\cosh\rho)^{\frac{1}{2}}}, \; n = 2\ell+2$$

The method to obtain the result is inspired by results in [15] and [18]. We omit the details here.

Acknowledgements

Der-Chen Chang was partially supported by NSF grant DMS-1408839 and a McDevitt Endowment Fund at Georgetown University. Chi-Feng Hung was supported by the research grants from the National Science and Technology Council, Far Eastern Memorial Hospital, and Taoyuan Armed Force General Hospital (numbers 110-2320-B-030-004-MY3, 109-FEMH-FJU-02, and 0000) in Taiwan).

References

[1] J.F. Adams, Vector fields on spheres, Ann. Math. Second Series 75 (1962), 603-632.

[2] A. Agrachev, Yu. Sachkov, Control theory from the geometric viewpoint, Encyclopaedia of Math. Sci. 87, Control Theory and Optimization, II. Springer-Verlag, Berlin, 2004.

[3] R. Beals, B. Gaveau, P.C. Greiner, On a geometric formula for the fundamental solution of subelliptic Laplacian, Math. Nachr. 181 (1996), 81-163.

[4] R. Beals, B. Gaveau, P.C. Greiner, The Green function of model two step hypoelliptic operators and the analysis of certain tangential Cauchy-Riemann complexes, Adv. Math. 121 (1996), 288-345.

[5] U. Boscain, G. Charlot, J.P. Gauthier, S. Guérin, H.-R.Jauslin, Optimal Control in laser-induced population transfer for two- and three-level quantum systems, J. Math. Phys. 43 (2002), 2107-2132.

[6] U. Boscain, F. Rossi, Invariant Carnot-Caratheodory metrics on $S^3, \mathrm{SO}(3), \mathrm{SL}(2)$, and lens spaces, SIAM J. Control Optim. 47 (2008), 1851-1878.

[7] O. Calin, D.C. Chang, Sub-Riemannian Geometry, General Theory and Examples, Encyclopedia of Mathematics and Its Applications, Vol. 126, Cambridge University Press, Cambridge, 2009.

[8] O. Calin, D.C. Chang, Geometric Mechanics on Riemannian Manifolds: Applications to Partial Differential Equations, Applied and Numerical Harmonic Analysis 29, Birkhäuser, Boston, Massachusetts, 2004.

[9] O. Calin, D.-C. Chang, P. Greiner, Geometric analysis on the Heisenberg group and its generalizations, AMS/IP Studies in Advanced Mathematics, 40, Amer. Math. Soc., Providence, RI; International Press, Somerville, MA, 2007.

[10] O. Calin, D.C. Chang, I. Markina, Sub-Riemannian geometry of the sphere $\mathbb{S}^3$, Canadian J. Math. 61 (2009), 821-839.

[11] D.C. Chang, O. Frieder, C.F. Hung, H.R. Yao, The analysis from nonlinear distance metric to kernel-based prescription prediction system, J. Nonlinear Var. Anal. 5 (2021), 179-19.

[12] D.C. Chang, I. Markina, A. Vasil'ev, Sub-Riemannian geodesics on the 3-D sphere, Complex Anal. Oper. Theory 3 (2009), 361-377.

[13] D.C. Chang, S.Y. Li, A zeta function associates to subLaplacian on the unit sphere in $\mathbf{C}^n$", J. Anal. Math. 86 (2002), 25-48.

[14] W.L. Chow, Uber Systeme van Linearen partiellen Differentialgleichungen erster Ordnung, Math. Ann. 117 (1939), 98-105.

[15] E.B. Davies, N. Mandouvalos, Heat kernel bounds on hyperbolic space and Kleinian groups, Proc. London Math. Soc. 57 (1988), 182-208.

[16] O. Frieder, D.C. Chang, H.R. Yao, Method and System for Assessing Drug Efficacy Using Multiple Graph Kernel Fusion, US Patent 11,238,966, (2022).

[17] M. Godoy Molina, I. Markina, Sub-Riemannian geometry on parallelizable spheres, Revista Matem. Iberoamericana, arXiv 0901.1406, 2009.

[18] A. Grigor"yan, M. Naguchi, The heat kernel on hyperbolic space, Bull. London Math. Soc. 30 (1998), 643-650.

[19] P.C. Greiner, A Hamiltonian approach to the heat kernel of a subLaplacian on $\mathbb{S}^{2n+1}$, arXiv 1303.0457v1, 2013.

[20] S.J. Gustavson, I.M. Sigal, Mathematical Concepts of Quantum Mechanics, Springer, Universitext, 2003.

[21] A. Hurtado, C. Rosales, Area-stationary surfaces inside the sub-Riemannian three-sphere, Math. Ann. 340 (2008), 675-708.

[22] D.F. Lawden, Elliptic Functions and Applications, 80, Springer, New York, 1989.

[23] D. Li, D. Dunson, Geodesic distance estimation with spherelets, arXiv: 1907.00296v2, [math.ST], 17, Apr 2020.

[24] W. Liu, H.J. Sussman, Shortest paths for sub-Riemannian metrics on rank-two distributions, Mem. Amer. Math. Soc. 118, 1995.

[25] X. Liu, F. Zhang, Z. Hou, L. Mian, Z. Wang, J. Zhang, J. Tang, Self-supervised learning: Generative or contrastive, IEEE Trans. Knowledge and Data Engineering, 35 (2021), 857-876.

[26] R. Montgomery, Abnormal minimizers, SIAM J. Control Optim. 32 (1994), 1605-1620.

[27] R. Montgomery, Survey of singular geodesics, Sub-Riemannian geometry, Progr. Math. vol. 144, pp. 325-339, Birkhäuser, Basel, 1996.

[28] R. Montgomery, A tour of subriemannian geometries, their geodesics and applications, Mathematical Surveys and Monographs, vol. 91. Amer. Math. Soc., Providence, RI, 2002.

[29] E. Mazet, M. Berger, P. Gauduchon, Le spectre d'une variété riemannienne, Lecture Notes in Math., vol. 194, Springer-Verlag, Berlin, 1971.

[30] M. Nagase, Expressions of the heat kernels on spheres by elementary functions and their recurrence relations, Saitama. Math. J. 27 (2010), 25-34.

[31] L.S. Schulman, Techniques and applications of path integration, Dover Publications, Inc., Mincola, New York, 2005.

[32] V.D. Silva, J.B. Tenenbaum, Global versus local methods in nonlinear dimensionality reduction, Advanced in Neural Information Proceedings Systems, 721-728, (2003).

[33] R.S. Strichartz, Sub-Riemannian geometry, J. Differential Geom. 24 (1986), 221-263, 1986; Correction, ibid. 30 (1989), 595-596.

[34] H. Urbantke, Two-level quantum systems: states, phases, and holonomy, Amer. J. Phys. 59 (1991), 503-509.

[35] H. Urbantke, The Hopf fibration–seven times in physics, J. Geom. Phys. 46 (2003), 125-150.

[36] H.R. Yao, D.C. Chang, O. Frieder, W. Huang, T.S. Lee, Graph Kernel prediction of drug prescription, 2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI). IEEE, 2019.

[37] H.R. Yao, D.C. Chang, O. Frieder, W. Huang, T.S. Lee, Multiple Graph Kernel Fusion Prediction of Drug Prescription, Proceedings of the 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics, 2019.

[38] H.R. Yao, D.C. Chang, O. Frieder, W. Huang, I.C. Liang, C.F. Hung, Cross-Global Attention Graph Kernel Network Prediction of Drug Prescription, Proceedings of the 11th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics, 2020.

[39] H.R. Yao, N. Cao, K, Russell, D.C. Chang, O. Frieder, J, Fineman, Self-supervised Representation Learning on Electronic Health Records with Graph Kernel Infomax, arXiv preprint arXiv:2209.00655, 2022.