High-jet Relations of the Heat Kernel: Embedding Map and Applications

Ke Zhu

Abstract For any compact Riemannian manifold (M, g) and its heat kernel embedding map $\psi_t : M \to l^2$ constructed in [BBG], we study the higher derivatives of ψ_t with respect to an orthonormal basis at x on M. As the heat flow time $t \to 0_+$, it turns out the limiting angles between these derivative vectors are universal constants independent on g, x and the choice of orthonormal basis. Geometric applications to the mean curvature and the Riemannian curvature are given. Some algebraic structures of the ∞-jet space of ψ_t are explored.

1 Introduction

Let (M, g) be a n-dimensional compact Riemannian manifold with smooth metric g. In [BBG], Bérard, Besson and Gallot considered the *heat kernel embedding*

$$\Phi_t : x \in M \to \left\{ e^{-\lambda_j t/2} \phi_j (x) \right\}_{j \geq 1} \in l^2 \tag{1}$$

where λ_j is the j-th eigenvalue of the Laplacian Δ_g of (M, g), $\{\phi_j\}_{j\geq 0}$ is the L^2 orthonormal eigenbasis of Δ_g, and l^2 is the Hilbert space of real-valued, square summable series. Let $\langle , \rangle$ be the standard inner product of l^2, then $\langle \Phi_t (x), \Phi_t (y) \rangle$ is the heat kernel

$$H (t, x, y) = \sum_{j=1}^{\infty} e^{-\lambda_j t} \phi_j (x) \phi_j (y)$$

Ke Zhu
Department of Mathematics and Statistics, Minnesota State University
Mankato, MN 56001, USA
e-mail: ke.zhu@mnsu.edu

on M. It was proved in [BBG] that the *normalized heat kernel embedding*

$$\psi_t : x \in M \to \sqrt{2}\,(4\pi)^{n/4}\, t^{\frac{n+2}{4}} \cdot \left\{e^{-\lambda_j t/2}\phi_j\,(x)\right\}_{j\geq 1} \in l^2 \tag{2}$$

is *asymptotically isometric*, i.e. the induced metric of the embedded image $\psi_t\,(M)$ in l^2 gets closer and closer to the original metric g as the heat flow time $t \to 0_+$. More precisely, as $t \to 0_+$

$$\psi_t^* g_0 = g + \frac{t}{3}\left(\frac{1}{2}S_g \cdot g - \mathrm{Ric}_g\right) + O\left(t^2\right), \tag{3}$$

where g_0 is the standard metric in l^2, S_g is the scalar curvature, and Ric_g is the Ricci curvature of (M, g). The embedding ψ_t is *canonical*, in the sense that it is constructed by the eigenfunctions of the Laplacian of (M, g).

Given any $x \in M$, we choose the normal coordinates $\left(x^1, \cdots x^n\right)$ near x such that $\left\{\frac{\partial}{\partial x^j}\right\}_{1\leq j\leq n}$ is *orthonormal* at x. Let $\overrightarrow{\alpha}$ be the multi-index of the mixed derivative operator $D^{\overrightarrow{\alpha}}$ on this chart. As $t \to 0_+$, it turns out these (l^2-valued) coefficients $\left\{D^{\overrightarrow{\alpha}}\Phi_t\,(x)\right\}$ have a special property: the angles between any two $D^{\overrightarrow{\alpha}}\Phi_t\,(x)$ and $D^{\overrightarrow{\beta}}\Phi_t\,(x)$ converge to universal constants *independent* on the metric g, the point x on M, and the orthonormal basis $\{V_i\}_{1\leq i\leq n} = \left\{\frac{\partial}{\partial x^j}\right\}_{1\leq j\leq n}$ at x, but only on $\overrightarrow{\alpha}$ and $\overrightarrow{\beta}$. The asymptote of the length of $D^{\overrightarrow{\alpha}}\Phi_t\,(x)$ only depends on $\overrightarrow{\alpha}$ and n too (Theorem 2). These *asymptotic relations* in the k-jet space of Φ_t (or ψ_t) have interesting applications:

1. When $k = 1$, [BBG] used them to construct *asymptotically isometric embeddings* $\psi_t = \sqrt{2}\,(4\pi)^{n/4}\, t^{\frac{n+2}{4}}\Phi_t : M \to l^2$ as in (3). They also used the embedding ψ_t to define the *spectral distance* between different metrics on M. When the metrics have a uniform upper bound of the diameter and a uniform lower bound of the Ricci curvature, they established a *precompactness* theorem on the space of such metrics;
2. When $k = 2$, by perturbing the almost isometric embedding ψ_t , [WZ] constructed *isometric embeddings* of compact Riemannian manifolds M into $\mathbb{R}^{q(t)}$ with controlled second fundamental form and $q\,(t) \sim t^{-n/2}$, where the 2-jet bundle of ψ_t appeared in the linearized operator of the isometric embedding problem. It was also showed that for any compact Riemannian manifolds, the mean curvature vectors of the image $\frac{1}{\sqrt{t}}\psi_t\,(M)$ converge to constant length $\sqrt{\frac{n+2}{2n}}$ as $t \to 0_+$;
3. When $k = 3$, we use them to show the embedded image $\frac{1}{\sqrt{t}}\psi_t\,(M) \subset l^2$ is *asymptotically umbilical* in the mean curvature vector direction as $t \to 0_+$ (Proposition 4). We also propose to construct submanifolds in $\mathbb{R}^q$ with constant mean curvature length by truncating and perturbing $\frac{1}{\sqrt{t}}\psi_t\,(M) \subset l^2$;

4. When $k = 2$, but also considering the secondary leading terms in the asymptote of the limiting angles, we prove the "asymptotic" Gauss formula (Proposition 5) to express the *Riemannian curvature* tensor of (M, g) as

$$R(X,Y,Z,W) = \lim_{t\to 0+} \left[\langle \nabla_X \nabla_W \psi_t, \nabla_Y \nabla_Z \psi_t \rangle - \langle \nabla_X \nabla_Z \psi_t, \nabla_Y \nabla_W \psi_t \rangle \right].$$

From this formula we can easily see the symmetry of the Riemannian curvature tensor, including the second Bianchi identity (Lemma 3). We can also express the *Levi-Civita connection* in terms of ψ_t (Lemma 2).

For example, the following 2-jet relations were proved in [WZ] and played a crucial role to construct isometric embeddings of M into $\mathbb{R}^q$. Let $\langle , \rangle$ be the standard inner product in l^2 and $|\cdot|$ be the standard metric in l^2. We have

Theorem 1. *([WZ] Theorem 2) For any $x \in M$, let $\left(x^1, \cdots x^n\right)$ be the normal coordinates near x such that $\left\{\frac{\partial}{\partial x^j}\right\}_{1\le j\le n}$ is* orthonormal *at x. Then for the n first derivatives vectors $\nabla_i \Phi_t$ $(1 \le i \le n)$ and the $n(n+1)/2$ second derivatives vectors $\nabla_i \nabla_j \Phi_t$ $(1 \le i \le j \le n)$ of Φ_t with respect to $\left\{\frac{\partial}{\partial x^j}\right\}_{1\le j\le n}$, as $t \to 0_+$ we have*

$$\frac{\langle \nabla_i \Phi_t(x), \nabla_j \Phi_t(x) \rangle}{|\nabla_i \Phi_t(x)| \, |\nabla_j \Phi_t(x)|} \to \delta_{ij}, \ \ \textit{and} \ \frac{\langle \nabla_i \nabla_j \Phi_t(x), \nabla_k \Phi_t(x) \rangle}{|\nabla_i \nabla_j \Phi_t(x)| \, |\nabla_k \Phi_t(x)|} \to 0,$$

and for $i \ne j$ or $k \ne l$,

$$\frac{\langle \nabla_i \nabla_j \Phi_t(x), \nabla_k \nabla_l \Phi_t(x) \rangle}{|\nabla_i \nabla_j \Phi_t(x)| \, |\nabla_k \nabla_l \Phi_t(x)|} \to 0,$$

except when $\{i, j\} = \{k, l\}$ as sets. Furthermore, for $i \ne j$,

$$\frac{\langle \nabla_i \nabla_i \Phi_t(x), \nabla_j \nabla_j \Phi_t(x) \rangle}{|\nabla_i \nabla_i \Phi_t(x)| \, |\nabla_j \nabla_j \Phi_t(x)|} \to \frac{1}{3}. \tag{4}$$

The convergence is uniform for all x on M.

In this paper, we generalize Theorem 1 to obtain the limiting angles between any two derivative vectors $D^{\overrightarrow{\alpha}} \Phi_t(x)$ and $D^{\overrightarrow{\beta}} \Phi_t(x)$ as $t \to 0_+$. It is a refinement of Proposition 15 in [WZ]; The new observation is that the coefficients there can be translated into certain enumeration problems of graphs and can be solve explicitly (Proposition 2).

In the following we denote the disjoint union of two sets by $\amalg$, and let $|\overrightarrow{\alpha}|$ be the total degree of the derivative operator $D^{\overrightarrow{\alpha}}$.

Definition 1. For any two derivative vectors $D^{\overrightarrow{\alpha}} \Phi_t(x)$ and $D^{\overrightarrow{\beta}} \Phi_t(x)$, suppose there are s distinct indices $\{j_1, j_2, \cdots, j_s\}$ in $\{\overrightarrow{\alpha}\} \amalg \left\{\overrightarrow{\beta}\right\}$, such that each j_r has multiplicity a_r in $\overrightarrow{\alpha}$, multiplicity b_r in $\overrightarrow{\beta}$, and let the average multiplicity $\sigma_r := \frac{a_r+b_r}{2}$ for $r = 1, 2, \cdots, s$. If all $a_r + b_r$ are even, we define the constants

$$A\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) = (-1)^{\frac{|\overrightarrow{\alpha}|-\left|\overrightarrow{\beta}\right|}{2}} \prod_{r=1}^{s} \left[\frac{(2\sigma_r)!}{2^{\sigma_r}(\sigma_r)!}\right], \tag{5}$$

$$\begin{aligned} B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) &= A\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) \Big/ \left[A\left(\overrightarrow{\alpha}, \overrightarrow{\alpha}\right) A\left(\overrightarrow{\beta}, \overrightarrow{\beta}\right)\right]^{1/2} \\ &= (-1)^{\frac{|\overrightarrow{\alpha}|-\left|\overrightarrow{\beta}\right|}{2}} \prod_{r=1}^{s} \frac{(2\sigma_r)!}{\sigma_r!} \Big/ \left[\frac{(2a_r)!}{a_r!}\frac{(2b_r)!}{b_r!}\right]^{1/2}. \end{aligned} \tag{6}$$

If some $a_r + b_r$ is odd, we simply let

$$A\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) = 0 = B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right). \tag{7}$$

We have our main theorem:

Theorem 2. *(Asymptotic high-jet relations) For any two derivative vectors* $D^{\overrightarrow{\alpha}}\Phi_t(x)$ *and* $D^{\overrightarrow{\beta}}\Phi_t(x)$*, as* $t \to 0_+$*,*

1. *If there is an index appeared with odd total multiplicity in* $\{\overrightarrow{\alpha}\} \amalg \left\{\overrightarrow{\beta}\right\}$*, then*
$$\frac{\left\langle D^{\overrightarrow{\alpha}}\Phi_t(x), D^{\overrightarrow{\beta}}\Phi_t(x)\right\rangle}{\left|D^{\overrightarrow{\alpha}}\Phi_t(x)\right| \left|D^{\overrightarrow{\beta}}\Phi_t(x)\right|} \to 0. \tag{8}$$
Especially this is the case when $|\overrightarrow{\alpha}| + \left|\overrightarrow{\beta}\right|$ *is odd;*
2. *If each index appears with even total multiplicity in* $\{\overrightarrow{\alpha}\} \amalg \left\{\overrightarrow{\beta}\right\}$*, then*
$$\frac{\left\langle D^{\overrightarrow{\alpha}}\Phi_t(x), D^{\overrightarrow{\beta}}\Phi_t(x)\right\rangle}{\left|D^{\overrightarrow{\alpha}}\Phi_t(x)\right| \left|D^{\overrightarrow{\beta}}\Phi_t(x)\right|} \to B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) \neq 0; \tag{9}$$
3. *We have*
$$\left|D^{\overrightarrow{\alpha}}\Phi_t(x)\right|^2 \to \frac{1}{(4\pi t)^{n/2}} \left(\frac{1}{2t}\right)^{|\overrightarrow{\alpha}|} A\left(\overrightarrow{\alpha}, \overrightarrow{\alpha}\right). \tag{10}$$

The above convergences are uniform for all x *on* M*.*

For example, for $n \geq 2$, $i \neq j$, $\overrightarrow{\alpha} = (i, i)$ and $\overrightarrow{\beta} = (j, j)$, we recover Theorem 2 in [WZ], because

$$
\begin{aligned}
A\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) &= \prod_{r=1}^{2}\left[(-1)^0 \frac{2!}{2^1 \cdot 1!}\right] = 1, \\
A\left(\overrightarrow{\alpha}, \overrightarrow{\alpha}\right) &= \prod_{r=1}^{1}\left[(-1)^0 \frac{4!}{2^2 \cdot 2!}\right] = 3 = A\left(\overrightarrow{\beta}, \overrightarrow{\beta}\right), \\
B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) &= A\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) \Big/ \left[A\left(\overrightarrow{\alpha}, \overrightarrow{\alpha}\right) A\left(\overrightarrow{\beta}, \overrightarrow{\beta}\right)\right]^{1/2} = \frac{1}{3}.
\end{aligned}
$$

Theorem 2 is a consequence of the more general result in Section 2 (Proposition 3), stated as follows:

Proposition 1. *As $t \to 0_+$, we have*

$$
(4\pi t)^{n/2} (2t)^{\left[\frac{|\overrightarrow{\alpha}|+|\overrightarrow{\beta}|}{2}\right]} \cdot D_y^{\overrightarrow{\beta}} D_x^{\overrightarrow{\alpha}} H(t,x,y)\Big|_{x=y} = A\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) + O(t), \quad (11)
$$

where $[c]$ is the largest integer less or equal to the given real number c. The convergence is uniform for all x on M.

After the completion of this paper, the author received the paper [Ni], which is related to our results. The paper [Ni], motivated by probabilistic questions, describes the small ε-asymptotics of the jets along the diagonal of the integral kernel of the smoothing operator $w\left(\varepsilon\sqrt{\Delta_g}\right)$ for an arbitrary nonnegative even Schwartz function w. For $w(x) = e^{-x^2}$, these reduce to the short time asymptotics of the jets along the diagonal of the heat kernel.

The paper is organized as follows: In Section 2 we first review the heat kernel and its near-diagonal expansion, then introduce certain graphs to aid the computation of the leading terms in (11), and then prove our main Theorem 2. In Section 3 we give applications of the high-jet relations of ψ_t to the mean curvature of $\psi_t(M) \subset l^2$ and the Riemannian curvature of M. In Section 4 we explore more algebraic structures of the high-jet relations, and reformulate them by the lattice geometry on $\mathbb{Z}_+^n$, with the motivation for future applications.

Acknowledgement. The author thanks Xiaowei Wang for many enlightening discussions during the collaboration of [WZ], where some ideas of the current work emerged. The author thanks Liviu I. Nicolaescu for pointing out the probabilistic point of view on our results. He also thanks Clifford Taubes for interest and support. The author thanks the IBS-CGP & MATRIX workshop on Symplectic Topology in 2022, where he learned some related asymptotic limits in tropical geometry.

2 High-jet relations

2.1 Heat kernel and derivatives of the distance function

Let $H(t,x,y)$ be the heat kernel of the Laplacian Δ_g on (M,g),

$$H(t,x,y) := \sum_{j=1}^{\infty} e^{-\lambda_j t}\phi_j(x)\phi_j(y). \tag{1}$$

It is well known that $H(t,x,y)$ has the *Minakshisundaram-Pleijel expansion*

$$H(t,x,y) = \frac{1}{(4\pi t)^{n/2}} e^{-\frac{r^2}{4t}} U(t,x,y)$$

(e.g. [BeGaM], p.213, or [Ch], p.154), where $r = r(x,y)$ is the distance function for points x and y on M,

$$U(t,x,y) = u_0(x,y) + t u_1(x,y) + \cdots + t^p u_p(x,y) + O\left(t^{p+1}\right) \tag{2}$$

as $t \to 0_+$, where the convergence is in C^l for any $l \geq 0$, and

$$\begin{aligned} u_0(x,y) &= [\theta(x,y)]^{-1/2} \text{ (for } x \text{ and } y \text{ close enough)}, \\ u_1(x,x) &= \frac{S_g(x)}{6}, \end{aligned} \tag{3}$$

where

$$\begin{aligned} \theta(x,y) &= \frac{\text{volume density at } y \text{ read in the normal coordinates around } x}{r^{n-1}} \\ &= \left|\det\left(d_y \exp_x\right)\right| = \left|\det\left(g_{ij}(y)\right)\right|^{1/2}, \end{aligned} \tag{4}$$

and in the second line of (4), y is regarded as a point on T_xM, and $d_y \exp_x$ is the differential of the exponential map $\exp_x : T_xM \to M$ ([BeGaM], p.208, [BGV], p.36). Given any x on M, let $\left(x^1, \cdots, x^n\right)$ be the normal coordinates around x and $\{V_j\}_{j=1}^n = \left\{\frac{\partial}{\partial x^j}\right\}_{j=1}^n$ be an orthonormal basis at x. For any unit vector $V \in T_xM$, it was derived in [BBG] that for $x_s = \exp_x(sV)$ and $x_\tau = \exp_x(\tau V)$ with small s and τ,

$$\theta(x_s, x_\tau) = 1 - \mathrm{Ric}_g(\dot{x}(s), \dot{x}(s)) \frac{(s-\tau)^2}{3!} + O\left(|s-\tau|^3\right),$$

so

$$u_0(x_s, x_\tau) = 1 + \frac{1}{12}\mathrm{Ric}_g(\dot{x}(s), \dot{x}(s))(s-\tau)^2 + O\left(|s-\tau|^3\right).$$

Hence

$$
\begin{aligned}
u_0(x,x) &= 1, &(5)\\
\nabla_i u_0(x,y)|_{x=y} &= 0, &(6)\\
\nabla_i U(t,x,y)|_{x=y} &= O(t), &(7)\\
\nabla_i^x \nabla_j^y u_0(x,y)|_{x=y} &= -\frac{1}{6}\mathrm{Ric}_g(V_i,V_j), &(8)
\end{aligned}
$$

For later computations, we will need the following facts on the derivatives of the squared distance function $r^2(x,y)$ on the diagonal $x=y$.

Lemma 1. *For $r=r(x,y)$ on $M\times M$, let $\nabla_i=\nabla^x_{V_j}$ be the partial derivative with respect to V_j in the x variables, $\nabla_{\bar{j}}=\nabla^y_{V_j}$ be the partial derivative with respect to V_j in the y variables, $\nabla_{ij}=\nabla^x_{V_i}\nabla^x_{V_j}$, $\nabla_{i\bar{j}}=\nabla^x_{V_i}\nabla^y_{V_j}$, $\nabla_{ij\bar{k}}=\nabla^x_{V_i}\nabla^x_{V_j}\nabla^y_{V_k}$ and so on. Then we have the following identities*

$$
\begin{aligned}
\nabla_i r^2(x,y)|_{x=y} &= 0 = \nabla_{\bar{\imath}} r^2(x,y)|_{x=y}, &(9)\\
\nabla_{ij} r^2(x,y)|_{x=y} &= 2\delta_{ij} = -\nabla_{i\bar{j}} r^2(x,y)|_{x=y}, &(10)\\
\nabla_{ijk} r^2(x,y)|_{x=y} &= \nabla_{ij\bar{k}} r^2|_{x=y} = \nabla_{i\bar{j}\bar{k}} r^2|_{x=y} = \nabla_{\bar{i}\bar{j}\bar{k}} r^2|_{x=y} = 0, &(11)\\
\nabla_{ijkl} r^2(x,y)|_{x=y} &= -\nabla_{ij\overline{kl}} r^2|_{x=y}(x,y) = -\frac{2}{3}\left(R_{ikjl}(x)+R_{iljk}(x)\right), &(12)
\end{aligned}
$$

where $R_{ijkl}(x)=R(V_i,V_j,V_k,V_l)(x)$ is the Riemannian curvature.

The identities (9) and (10) are well-known (e.g. [BBG]). The identities (11) and (12) can be found in Chapter 16 (p.282) of [De], where $\frac{1}{2}r^2(x,y)$ is called the *world function*, and the limit at $x=y$ is called the coincidence limit.

2.2 Admissible graphs

To compute $\left\langle D^{\overrightarrow{\alpha}}\Phi_t(x), D^{\overrightarrow{\beta}}\Phi_t(x)\right\rangle$, it is useful to introduce certain graphs.

Definition 2. (Vertex set) Let $\mathcal{S}=\{\overrightarrow{\alpha}\}\amalg\left\{\overrightarrow{\beta}\right\}$ be the set with $|\overrightarrow{\alpha}|+\left|\overrightarrow{\beta}\right|$ vertices, where each vertex is an index $i\in\mathcal{S}$, with the sign "+" if it is in $\{\overrightarrow{\alpha}\}$ and sign "−" if it is in $\left\{\overrightarrow{\beta}\right\}$. We say the vertex has color i if we do not distinguish its sign.

Definition 3. (Admissible graph) A graph G on the vertex set $\mathcal{S}=\{\overrightarrow{\alpha}\}\amalg\left\{\overrightarrow{\beta}\right\}$ is called *admissible* if (i) Each vertex is connected to only one another vertex; (ii) Each edge only connects two vertices with the same color; (iii) Each edge has a sign $\pm$ being the negative of the product of the signs of its two end-points. Let Sign(G) be the product of the signs of all edges of G.

Remark 1. The necessary and sufficient condition for $\mathcal{S}=\{\overrightarrow{\alpha}\}\amalg\left\{\overrightarrow{\beta}\right\}$ to have admissible graph(s) is that each index appears in $\mathcal{S}$ with even multiplicity. Especially $|\overrightarrow{\alpha}|+\left|\overrightarrow{\beta}\right|$ must be even. If G is not admissible, we simply let $\mathrm{Sign}(G)=0$.

Proposition 2. *Suppose there are s distinct indices $\{j_1,j_2,\cdots,j_s\}\subset\{1,2,\cdots,n\}$ in $\{\overrightarrow{\alpha}\}\amalg\left\{\overrightarrow{\beta}\right\}$, such that each j_r has multiplicity a_r in $\overrightarrow{\alpha}$, and multiplicity b_r in $\overrightarrow{\beta}$. Let $\sigma_r=\frac{a_r+b_r}{2}$. Then we have*

1. *For each admissible graph G of $\mathcal{S}=\{\overrightarrow{\alpha}\}\amalg\left\{\overrightarrow{\beta}\right\}$, $sign(G)=(-1)^{\frac{|\overrightarrow{\alpha}|-\left|\overrightarrow{\beta}\right|}{2}}$;*
2. *$\#\{admissible\ graphs\ of\ \mathcal{S}\}=\prod_{r=1}^{s}\left[\frac{(2\sigma_r)!}{2^{\sigma_r}(\sigma_r)!}\right]=\left|A\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)\right|$.*

Proof. By the definition of $\mathrm{sign}(G)$, it is the product of the signs of its edges. There are $\frac{|\overrightarrow{\alpha}|+\left|\overrightarrow{\beta}\right|}{2}$ edges, each edge connecting vertices i and i' has the sign $(-1)\cdot\mathrm{sign}(i)\mathrm{sign}(i')$, and the disjoint union of edges covers all vertices of G, so

$$\begin{aligned}\mathrm{sign}\left(G\right)&=\prod_{\text{edge } e \text{ of } G}\mathrm{sign}(e)=(-1)^{\frac{|\overrightarrow{\alpha}|+\left|\overrightarrow{\beta}\right|}{2}}\cdot\prod_{\text{vertex } i \text{ of } G}\mathrm{sign}(i)\\&=(-1)^{\frac{|\overrightarrow{\alpha}|+\left|\overrightarrow{\beta}\right|}{2}}\cdot(-1)^{\left|\overrightarrow{\beta}\right|}\ =(-1)^{\frac{|\overrightarrow{\alpha}|-\left|\overrightarrow{\beta}\right|}{2}},\end{aligned}$$

where the second line is because only vertices in $\overrightarrow{\beta}$ have "$-$" sign. For each index j_r, it has total multiplicity $2\sigma_r$ in $\mathcal{S}=\{\overrightarrow{\alpha}\}\amalg\left\{\overrightarrow{\beta}\right\}$, i.e. there are $2\sigma_r$ vertices with color j_r in $\mathcal{S}$, so the ways to draw disjoint edges on this subset of vertices in $\mathcal{S}$ is

$$\begin{aligned}&\binom{2\sigma_r}{2\sigma_r-2}\binom{2\sigma_r-2}{2\sigma_r-4}\cdots\binom{2}{2}\Big/\sigma_r!\\=&\ \frac{(2\sigma_r)!}{(2\sigma_r-2)!2!}\frac{(2\sigma_r-2)!}{(2\sigma_r-4)!2!}\cdots\frac{2!}{2!}\Big/\sigma_r!\\=&\ \frac{(2\sigma_r)!}{2^{\sigma_r}\left(\sigma_r\right)!}\end{aligned}$$

Considering all distinct indices $j_1,j_2,\cdots$ and j_s in $\mathcal{S}$, we see the number of admissible graphs on $\mathcal{S}$ is $\prod_{r=1}^{s}\left[\frac{(2\sigma_r)!}{2^{\sigma_r}(\sigma_r)!}\right]$. $\square$

Remark 2. The number $A\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$ is related to the moments of the Gaussian $e^{-|x|^2}$ by

$$A\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)=2^{\left[\frac{|\overrightarrow{\alpha}|+\left|\overrightarrow{\beta}\right|}{2}\right]}\pi^{-n/2}\int_{\mathbb{R}^n}e^{-|x|^2}x^{\overrightarrow{\alpha}+\overrightarrow{\beta}}dx.$$

Actually, the above combinatorics of the numbers $A\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$ is a manifestation of a classical *Wick formula* (or *Isserlis' theorem*) of the Gauss integrals, see e.g. [AT], Section 11.6. The formula is frequently used in the Feynman integral.

2.3 High-jet relations from the heat kernel expansion

Given any x on M, let $\{V_i\}_{i=1}^n$ be an orthonormal basis at x, $\nabla_i=\nabla_{V_i}^x$ be the partial derivative with respect to V_i in the x variables, and $\nabla_{\bar{j}}=\nabla_{V_j}^y$ be the partial derivative with respect to V_j in the y variables, where "$-$" indicates the derivative is with respect to the y-variables. We also denote $\nabla_{ij}=\nabla_{V_i}^x\nabla_{V_j}^x$, $\nabla_{i\bar{j}}=\nabla_{V_i}^x\nabla_{V_j}^y$ and so on. Since

$$\langle\Phi_t(x),\Phi_t(y)\rangle=\sum_{j=1}^{\infty}e^{-\lambda_j t}\phi_j(x)\,\phi_j(y)=H(t,x,y)\,,$$

applying the derivative operators $D_x^{\overrightarrow{\alpha}}$ and $D_y^{\overrightarrow{\beta}}$ to the above identity and letting $x=y$, we get

$$\left\langle D^{\overrightarrow{\alpha}}\Phi_t(x)\,,D^{\overrightarrow{\beta}}\Phi_t(x)\right\rangle=D_y^{\overrightarrow{\beta}}D_x^{\overrightarrow{\alpha}}H(t,x,y)\,|_{x=y}. \tag{13}$$

Proposition 3. *Let $\overrightarrow{\alpha}$ and $\overrightarrow{\beta}$ be two multi-indices in the set $\{1,2,\cdots n\}$. Then as $t\to0_+$, we have*

$$(4\pi t)^{n/2}(2t)^{\left[\frac{|\overrightarrow{\alpha}|+\left|\overrightarrow{\beta}\right|}{2}\right]}\cdot D_y^{\overrightarrow{\beta}}D_x^{\overrightarrow{\alpha}}H(t,x,y)\Big|_{x=y}\to A\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right), \tag{14}$$

where $[c]$ is the largest integer less or equal to the given real number c. If $A\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)=0$, we further have

$$(4\pi t)^{n/2}(2t)^{\left[\frac{|\overrightarrow{\alpha}|+\left|\overrightarrow{\beta}\right|}{2}\right]}\cdot D_y^{\overrightarrow{\beta}}D_x^{\overrightarrow{\alpha}}H(t,x,y)\Big|_{x=y}=O(t)\,. \tag{15}$$

The convergence is uniform for all x on M.

Proof. For any multi-index $\vec{\gamma}$, applying the Leibniz rule to

$$H\left(t,x,y\right)=\frac{1}{\left(4\pi t\right)^{n/2}}e^{-\frac{r^2}{4t}}U\left(t,x,y\right)$$

we can write

$$D^{\overrightarrow{\gamma}}H\left(t,x,y\right)=\frac{1}{\left(4\pi t\right)^{n/2}}e^{-\frac{r^2}{4t}}P_{\overrightarrow{\gamma}}\left(t,x,y\right),$$

where $P_{\overrightarrow{\gamma}}\left(t,x,y\right)$ is a polynomial in $D^{\overrightarrow{\mu_j}}\left(-\frac{r^2(x,y)}{4t}\right)$ and $D^{\overrightarrow{\eta}}U\left(t,x,y\right)\Big|_{x=y}$, i.e. each summand is of the form

$$\left(\prod_{j=1}^{l}D^{\overrightarrow{\mu_j}}\left(-\frac{r^2\left(x,y\right)}{4t}\right)\right)\Bigg|_{x=y}\cdot D^{\overrightarrow{\eta}}U\left(t,x,y\right)\Big|_{x=y} \tag{16}$$

for some multi-indices $\overrightarrow{\mu_j}$ and $\overrightarrow{\eta}$ with

$$\sum_{j=1}^{l}|\overrightarrow{\mu_j}|+|\overrightarrow{\eta}|=|\overrightarrow{\gamma}|. \tag{17}$$

For example, when $\overrightarrow{\gamma}=\partial_{x_i}$,

$$P_{\overrightarrow{i}}\left(t,x,y\right)=\partial_i\left(-\frac{r^2\left(x,y\right)}{4t}\right)U\left(t,x,y\right)+\partial_iU\left(t,x,y\right).$$

(more examples of $P_{\overrightarrow{\gamma}}\left(t,x,y\right)$ with $|\overrightarrow{\gamma}|\leq 4$ are in [WZ]). Now we take

$$\vec{\gamma}=\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right),|\overrightarrow{\gamma}|=|\overrightarrow{\alpha}|+\left|\overrightarrow{\beta}\right|,D^{\vec{\gamma}}=D_y^{\overrightarrow{\beta}}D_x^{\overrightarrow{\alpha}},$$

where $\overrightarrow{\beta}$ indicates the derivative $D_y^{\overrightarrow{\beta}}$ is with respect to the y-variables. We have the following claims:

1. As $t\rightarrow 0_+$, the nonzero summands of $P_{\vec{\gamma}}\left(t,x,y\right)$ involving the highest power of $\frac{1}{t}$ must have $\overrightarrow{\mu_j}'s$ with $|\overrightarrow{\mu_j}|=2$ as many as possible, such that each $\overrightarrow{\mu_j}$ is of the form (i,i), $\left(i,\bar{i}\right)$ or $\left(\bar{i},\bar{i}\right)$ for some $i\in\{1,2,\cdots,n\}$, and $|\overrightarrow{\eta}|=0$ or 1. We also have

$$\left|D^{\overrightarrow{\gamma}}H\left(t,x,y\right)|_{x=y}\right|\leq C\frac{1}{\left(4\pi t\right)^{n/2}}\cdot t^{-\left[\frac{|\overrightarrow{\gamma}|}{2}\right]},$$

   where the constant C depends on n and $|\overrightarrow{\gamma}|$.
2. If the total multiplicity of each index in $\{\overrightarrow{\alpha}\}\amalg\left\{\overrightarrow{\beta}\right\}$ is even, then we further have $l=\frac{|\overrightarrow{\gamma}|}{2}$, and $|\overrightarrow{\eta}|=0$ for the the nonzero summands of

$P_{\vec{\gamma}}(t,x,y)$ involving the highest power of $\frac{1}{t}$. These terms are of the same sign $(-1)^{\frac{|\vec{\alpha}|-|\vec{\beta}|}{2}}$, and the total number of such terms is $\left|A\left(\vec{\alpha},\vec{\beta}\right)\right|$.

For Claim 1, from Lemma 1 we have

$$\begin{aligned}
\partial_i\left(-\frac{r^2(x,y)}{4t}\right)|_{x=y} &= 0,\\
\partial_{ij}\left(-\frac{r^2(x,y)}{4t}\right)|_{x=y} &= -\frac{\delta_{ij}}{2t} = -\partial_{i\bar{j}}\left(-\frac{r^2(x,y)}{4t}\right)|_{x=y},\\
\partial_{ijk}\left(-\frac{r^2(x,y)}{4t}\right)|_{x=y} &= 0 = \partial_{ij\bar{k}}\left(-\frac{r^2(x,y)}{4t}\right)|_{x=y},\\
\left|D^{\vec{\mu_j}}\left(-\frac{r^2(x,y)}{4t}\right)|_{x=y}\right| &\le C\cdot\frac{1}{t}, \text{ for any } \vec{\mu}_j. \qquad (18)
\end{aligned}$$

So from the equations in (18) and the total degree condition (17), the summand

$$\left(\prod_{j=1}^{l} D^{\vec{\mu_j}}\left(-\frac{r^2(x,y)}{4t}\right)\right)\Bigg|_{x=y}\cdot D^{\vec{\eta}}U$$

can not exceed $\left[\frac{|\vec{\gamma}|}{2}\right]$ copies of factor $\frac{1}{t}$. The $\left[\frac{|\vec{\gamma}|}{2}\right]$ copies of factor $\frac{1}{t}$ can be achieved if $|\vec{\mu_j}|=2$ for all j, and $|\vec{\eta}|=0$ or 1 depending on $|\vec{\gamma}|$ is even or odd, so

$$\left|D^{\vec{\eta}}U\right| \le |u_0(x,x)| + |\nabla_i u_0(x,y)|_{x=y}| \le 2$$

by (5) and (6) as $t\to 0_+$. Such summand is nonzero if and only if each $\vec{\mu_j}$ is of the form (i,i), $\left(i,\bar{i}\right)$ or $\left(\bar{i},\bar{i}\right)$ for some $i\in\{1,2,\cdots,n\}$ by Lemma 1, so there are at most $3n$ choices of each $\vec{\mu_j}$. Therefore as $t\to 0_+$,

$$\begin{aligned}
\left|D^{\vec{\gamma}}H(t,x,y)|_{x=y}\right| &= \left|\frac{1}{(4\pi t)^{n/2}}e^{-\frac{r^2}{4t}}P_{\vec{\gamma}}(t,x,y)\Big|_{x=y}\right|\\
&\le \frac{1}{(4\pi t)^{n/2}}\cdot\left(\frac{1}{2t}\right)^{\left[\frac{|\vec{\gamma}|}{2}\right]}\cdot(3n)^{\left[\frac{|\vec{\gamma}|}{2}\right]}\cdot 2.
\end{aligned}$$

For Claim 2, since each index appears with even multiplicity in $\{\vec{\alpha}\}\amalg\left\{\vec{\beta}\right\}$, it is easy to see the summands (16) with $l=\frac{|\vec{\gamma}|}{2}$ and $|\vec{\eta}|=0$ can have maximal number of $\vec{\mu_j}$'s satisfying $|\vec{\mu_j}|=2$, with each $\vec{\mu_j}$ is of the form (i,i), $\left(i,\bar{i}\right)$ or $\left(\bar{i},\bar{i}\right)$ for some $i\in\{1,2,\cdots,n\}$. It is of the form

$$\left.\left(\prod_{j=1}^{l} D^{\overrightarrow{\mu_j}}\left(-\frac{r^2(x,y)}{4t}\right)\right)\right|_{x=y} \cdot U(t,x,x).$$

Since

$$\partial_{ij}\left(-\frac{r^2(x,y)}{4t}\right)|_{x=y} = -\frac{\delta_{ij}}{2t} = -\partial_{i\bar{j}}\left(-\frac{r^2(x,y)}{4t}\right)|_{x=y}, \tag{19}$$

the product

$$\left.\left(\prod_{j=1}^{l} D^{\overrightarrow{\mu_j}}\left(-\frac{r^2(x,y)}{4t}\right)\right)\right|_{x=y} = \prod_{j=1}^{l} \operatorname{sign}(\vec{\mu}_j)\cdot\left(\frac{1}{2t}\right)^{l}$$

gives the highest power of $\frac{1}{t}$, where we let $\operatorname{sign}(\vec{\mu}_j) = 1$ if $\overrightarrow{\mu_j} = (i,i)$ or $\left(\bar{i},\bar{i}\right)$, and $\operatorname{sign}(\vec{\mu}_j) = -1$ if $\overrightarrow{\mu_j} = \left(i,\bar{i}\right)$. If these summands do not cancel each other, they give the leading term in $D^{\overrightarrow{\gamma}}H(t,x,y)|_{x=y}$ as $t \to 0_+$, of the form

$$\frac{1}{(4\pi t)^{n/2}}\cdot\left(\frac{1}{2t}\right)^{\frac{|\overrightarrow{\gamma}|}{2}}\cdot(-1)^a A,$$

where A is the number of such summands, and $(-1)^a$ is their common sign.

To determine the number A, we build a one-one correspondence between such summands and admissible graphs in Definition 3. For each $\overrightarrow{\mu_j} = (i,i)$, $\left(i,\bar{i}\right)$ or $\left(\bar{i},\bar{i}\right)$ of the summand, we draw an edge between the corresponding vertices in $\mathcal{S} = \{\overrightarrow{\alpha}\} \amalg \left\{\overrightarrow{\beta}\right\}$. Drawing l edges on $\mathcal{S}$ with no intersecting vertex, we obtain an admissible graph G. The sign on each edge of G reflects the sign in (19). Taking product of them we see $\operatorname{sign}(G)$ is exactly the sign of the summand. Conversely, given any admissible graph G, we can read $\overrightarrow{\mu_j}$ from its edges, and then can write down the corresponding summand $\left.\left(\prod_{j=1}^{l} D^{\overrightarrow{\mu_j}}\left(-\frac{r^2(x,y)}{4t}\right)\right)\right|_{x=y} \cdot U(t,x,x)$ in $\left.D_y^{\overrightarrow{\beta}} D_x^{\overrightarrow{\alpha}} H(t,x,y)\right|_{x=y}$. From Proposition 2 we know $\operatorname{sign}(G) = (-1)^{\frac{|\overrightarrow{\alpha}|-\left|\overrightarrow{\beta}\right|}{2}}$, so the summands are of the same sign $(-1)^{\frac{|\overrightarrow{\alpha}|-\left|\overrightarrow{\beta}\right|}{2}}$. Proposition 2 also gives the signed count of admissible graphs, which is $A\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$. Claim 2 is proved.

If each index appears with even multiplicity in $\{\overrightarrow{\alpha}\} \amalg \left\{\overrightarrow{\beta}\right\}$, then $|\overrightarrow{\gamma}|$ is even, $\left[\frac{|\overrightarrow{\gamma}|}{2}\right] = \frac{|\overrightarrow{\gamma}|}{2}$. From the above argument we have as $t \to 0_+$,

$$
\begin{aligned}
&(4\pi t)^{n/2}(2t)^{\left[\frac{|\vec{\alpha}|+|\vec{\beta}|}{2}\right]}\cdot D_y^{\vec{\beta}} D_x^{\vec{\alpha}} H(t,x,y)\Big|_{x=y} \\
\to\ &(4\pi t)^{n/2}(2t)^{\frac{|\vec{\gamma}|}{2}}\cdot\frac{1}{(4\pi t)^{n/2}}\cdot\left(\frac{1}{2t}\right)^{\frac{|\vec{\gamma}|}{2}} A\left(\vec{\alpha},\vec{\beta}\right)=A\left(\vec{\alpha},\vec{\beta}\right).
\end{aligned}
$$

If some index i appears with odd multiplicity in $\{\vec{\alpha}\}\amalg\left\{\vec{\beta}\right\}$, then by our definition $A\left(\vec{\alpha},\vec{\beta}\right)=0$. For a nonzero summand of the form

$$
\left(\prod_{j=1}^{l} D^{\vec{\mu_j}}\left(-\frac{r^2(x,y)}{4t}\right)\right)\Bigg|_{x=y}\cdot D^{\vec{\eta}}U,
$$

by the equations in (18) and the total degree condition (17), it can have the $\frac{1}{t}$ power at most $\left[\frac{|\vec{\gamma}|-1}{2}\right]$. If the power $\left[\frac{|\vec{\gamma}|-1}{2}\right]$ is achieved, then $|\vec{\eta}|\le 2$ by the total degree condition (17), and $|\vec{\eta}|=2$ is only possible when $|\vec{\gamma}|$ is even. If $|\vec{\eta}|=0$, there will be a term $D^{\vec{\sigma}}\left(r^2\right)|_{x=y}$ in the summand, such that $|\vec{\sigma}|\le 3$ and i appears with odd multiplicity in $\vec{\sigma}$, making the summand zero by Lemma 1. So we must have $|\vec{\eta}|=1$ when $|\vec{\gamma}|$ is odd.

Similar to the argument in Claim 2, we have

$$
\left|D^{\vec{\gamma}} H(t,x,y)\Big|_{x=y}\right|\le C\frac{1}{(4\pi t)^{n/2}}\cdot\left(\frac{1}{t}\right)^{\left[\frac{|\vec{\gamma}|-1}{2}\right]}\left|D^{\vec{\eta}}U\right|. \qquad (20)
$$

When $|\vec{\eta}|=1$, by (7) we have $\left|D^{\vec{\eta}}U\right|\le Ct$. Therefore

$$
\begin{aligned}
&(4\pi t)^{n/2}(2t)^{\left[\frac{|\vec{\alpha}|+|\vec{\beta}|}{2}\right]}\cdot D_y^{\vec{\beta}} D_x^{\vec{\alpha}} H(t,x,y)\Big|_{x=y} \\
\to &\left\{\begin{array}{l}
t^{\left[\frac{|\vec{\gamma}|}{2}\right]}\cdot O\left(t^{-\left[\frac{|\vec{\gamma}|-1}{2}\right]}\right)\cdot O(t)=O(t)\to 0\ (\text{if } |\vec{\gamma}| \text{ is odd}),\\
t^{\left[\frac{|\vec{\gamma}|}{2}\right]}\cdot O\left(t^{-\left[\frac{|\vec{\gamma}|-1}{2}\right]}\right)\cdot O(1)=O(t)\to 0\ (\text{if } |\vec{\gamma}| \text{ is even}).
\end{array}\right\}=A\left(\vec{\alpha},\vec{\beta}\right).
\end{aligned}
$$

Proposition 3 is proved. The convergence is uniform for all x on M, because in *Minakshisundaram-Pleijel expansion*

$$
H(t,x,y)=\frac{1}{(4\pi t)^{n/2}}e^{-\frac{r^2}{4t}}U(t,x,y)
$$

the term

$$U(t,x,y) = u_0(x,y) + tu_1(x,y) + \cdots + t^p u_p(x,y) + O\left(t^{p+1}\right) \tag{21}$$

as $t \to 0_+$, where the convergence is in C^l for any $l \geq 0$ (Theorem 2.30 in [BGV]). $\square$

Now we are ready to give the proof of Theorem 2.

Proof. Letting $\overrightarrow{\alpha} = \overrightarrow{\beta}$ in Proposition 3, and noticing

$$\left\langle D^{\overrightarrow{\alpha}} \Phi_t(x), D^{\overrightarrow{\beta}} \Phi_t(x) \right\rangle = D_y^{\overrightarrow{\beta}} D_x^{\overrightarrow{\alpha}} H(t,x,y)\,|_{x=y},$$

we have

$$\begin{aligned}
&(4\pi t)^{n/2} (2t)^{|\overrightarrow{\alpha}|} \cdot \left| D^{\overrightarrow{\alpha}} \Phi_t(x) \right|^2 \\
&= (4\pi t)^{n/2} (2t)^{\left[\frac{|\overrightarrow{\alpha}|+|\overrightarrow{\alpha}|}{2}\right]} \cdot D_y^{\overrightarrow{\alpha}} D_x^{\overrightarrow{\alpha}} H(t,x,y)\Big|_{x=y} \to A(\overrightarrow{\alpha}, \overrightarrow{\alpha}),
\end{aligned}$$

as $t \to 0_+$, i.e.

$$\left| D^{\overrightarrow{\alpha}} \Phi_t(x) \right|^2 \to \frac{1}{(4\pi t)^{n/2}} \left(\frac{1}{2t}\right)^{|\overrightarrow{\alpha}|} A(\overrightarrow{\alpha}, \overrightarrow{\alpha}). \tag{22}$$

If $|\overrightarrow{\alpha}| + \left|\overrightarrow{\beta}\right|$ is even, by Proposition 3 we have

$$\begin{aligned}
&\lim_{t\to 0_+} \frac{\left\langle D^{\overrightarrow{\alpha}} \Phi_t(x), D^{\overrightarrow{\beta}} \Phi_t(x) \right\rangle}{\left| D^{\overrightarrow{\alpha}} \Phi_t(x) \right| \left| D^{\overrightarrow{\beta}} \Phi_t(x) \right|} \\
&= \lim_{t\to 0_+} \frac{(4\pi t)^{n/2} (2t)^{\frac{|\overrightarrow{\alpha}|+\left|\overrightarrow{\beta}\right|}{2}} \cdot \left\langle D^{\overrightarrow{\alpha}} \Phi_t(x), D^{\overrightarrow{\beta}} \Phi_t(x) \right\rangle}{\left((4\pi t)^{n/4} (2t)^{\frac{|\overrightarrow{\alpha}|}{2}} \left| D^{\overrightarrow{\alpha}} \Phi_t(x) \right| \right) \cdot \left((4\pi t)^{n/4} (2t)^{\frac{\left|\overrightarrow{\beta}\right|}{2}} \left| D^{\overrightarrow{\beta}} \Phi_t(x) \right| \right)} \\
&= \left[A\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) \right] \Big/ \left[A(\overrightarrow{\alpha}, \overrightarrow{\alpha}) A\left(\overrightarrow{\beta}, \overrightarrow{\beta}\right) \right]^{1/2} = B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right).
\end{aligned}$$

If $|\overrightarrow{\alpha}| + \left|\overrightarrow{\beta}\right|$ is odd, then some index must appear with odd multiplicity in $\{\overrightarrow{\alpha}\} \amalg \left\{\overrightarrow{\beta}\right\}$. By (20) and (22),

$$
\begin{aligned}
&\lim_{t\to 0_+} \frac{\left|\left\langle D^{\overrightarrow{\alpha}}\Phi_t(x), D^{\overrightarrow{\beta}}\Phi_t(x)\right\rangle\right|}{\left|D^{\overrightarrow{\alpha}}\Phi_t(x)\right|\left|D^{\overrightarrow{\beta}}\Phi_t(x)\right|} \\
&\leq \lim_{t\to 0_+} \frac{C\frac{1}{(4\pi t)^{n/2}}\left(\frac{1}{t}\right)^{\left[\frac{|\overrightarrow{\alpha}|+\left|\overrightarrow{\beta}\right|-1}{2}\right]}}{\frac{1}{(4\pi t)^{n/4}}\left(\frac{1}{2t}\right)^{\frac{|\overrightarrow{\alpha}|}{2}} A(\overrightarrow{\alpha},\overrightarrow{\alpha}) \cdot \frac{1}{(4\pi t)^{n/4}}\left(\frac{1}{2t}\right)^{\frac{\left|\overrightarrow{\beta}\right|}{2}} A\left(\overrightarrow{\beta},\overrightarrow{\beta}\right)} \\
&= O(t) \to 0 = B\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right).
\end{aligned}
$$

By Proposition 3, the convergence is uniform for all x on M. Theorem 2 is proved. □

Remark 3. If we only care about the leading order terms in $\left\langle D^{\overrightarrow{\alpha}}\Phi_t(x), D^{\overrightarrow{\beta}}\Phi_t(x)\right\rangle$ or $D_y^{\overrightarrow{\beta}} D_x^{\overrightarrow{\alpha}} H(t,x,y)\,|_{x=y}$ as $t \to 0_+$, the ordering of derivatives within $\overrightarrow{\alpha}$ and $\overrightarrow{\beta}$ are not important, and we can regard $\overrightarrow{\alpha}$ and $\overrightarrow{\beta}$ as *sets* of their indices, respectively. This is because when we switch the ordering of derivatives, say from $D_x^i D_x^j$ to $D_x^j D_x^i$, their difference is a lower order differential operator, and for lower order differential operator $D^{\overrightarrow{\alpha'}}$ with $\left|\overrightarrow{\alpha'}\right| < |\overrightarrow{\alpha}|$, $\left|D^{\overrightarrow{\alpha'}}\Phi_t(x)\right| = o\left(\left|D^{\overrightarrow{\alpha}}\Phi_t(x)\right|\right)$ as $t \to 0_+$ for all x on M by (10).

Remark 4. We remark that except the constant term $A\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$, the next order terms in $t^{\frac{n}{2}+\left[\frac{|\overrightarrow{\alpha}|+\left|\overrightarrow{\beta}\right|}{2}\right]} \cdot D_y^{\overrightarrow{\beta}} D_x^{\overrightarrow{\alpha}} H(t,x,y)\Big|_{x=y}$ do depend on g (or curvature), and are important for some geometric applications (e.g. Proposition 5), but to compute them is significantly harder.

3 Applications

In this section, we give several applications of the high-jet relations of ψ_t in Theorem 2 and Proposition 3. The applications to the mean curvature and the Riemannian curvature tensor are related to the *extrinsic geometry* and *intrinsic geometry* of $\psi_t(M) \subset l^2$ as $t \to 0_+$ respectively. We also discuss the *approximation* of the geometry of (M,g) by finitely many eigenfunctions.

3.1 Mean curvature of $\psi_t(M)$ in l^2

The first application is about the mean curvature of the embedded image $\psi_t(M) \subset l^2$. Intuitively speaking, as $t \to 0_+$, the extrinsic geometry of $\psi_t(M) \subset l^2$ locally looks the same at any point, and is evenly bumpy everywhere. More precisely we have

Proposition 4. *For the heat kernel embedding $\psi_t : M \to l^2$ and $x \in M$, let*

$$A(x,t) : T_{\psi_t(x)}\psi_t(M) \times T_{\psi_t(x)}\psi_t(M) \to T_{\psi_t(x)}\psi_t(M)^{\perp}$$

be the second fundamental form of the embedded image $\psi_t(M) \subset l^2$ and $H(x,t)$ be the mean curvature vector at $\psi_t(x)$. We have

1. *([WZ] Corollary 22) $\sqrt{t}A(x,t)$ converges to certain normal form for all x on M as $t \to 0_+$. The mean curvature vector $H(x,t)$, after scaled by a factor $\sqrt{t}$, converges to constant length:*

$$\lim_{t\to 0_+} \sqrt{t}\,|H(x,t)| = \sqrt{\frac{n+2}{2n}}. \tag{1}$$

2. *$\psi_t(M)$ is asymptotically umbilical in the mean curvature vector direction as $t \to 0_+$, in the following sense:*

$$t \cdot \lim_{t\to 0_+} S(x,t)\langle W,H\rangle = -\frac{3}{2}W, \text{ for any } W \in T_{\psi_t(x)}\psi_t(M),$$

where

$$S(x,t) : T_{\psi_t(x)}\psi_t(M) \times T_{\psi_t(x)}\psi_t(M)^{\perp} \to T_{\psi_t(x)}\psi_t(M)$$

is the second fundamental tensor on $\psi_t(M)$ defined by $S(X,\upsilon) = (\nabla_X \upsilon)^T$ for any $X \in T_{\psi_t(x)}\psi_t(M)$ and $\upsilon \in T_{\psi_t(x)}\psi_t(M)^{\perp}$, $\mathsf{T} : T_{\psi_t(x)}l^2 \to T_{\psi_t(x)}\psi_t(M)$ is the projection, and "$\perp$" means the orthogonal complement in l^2.

The convergence is uniform for all x on M.

Proof. For any $x \in M$, let $(x_1, \cdots, x_n)$ be a local coordinates near x such that the coordinate vectors $\left\{\frac{\partial}{\partial x^i}\right\}_{i=1}^n := \{V_i\}_{i=1}^n$ are orthonormal at x. The *second fundamental form* of the submanifold $\psi_t(M) \subset l^2$ can be written as

$$A(x,t) = \sum_{1\le i,j\le n} h_{ij}(x,t)\,dx^i dx^j$$

where $h_{jk}(x,t)$ $(1 \le j,k \le n)$ are vectors in l^2. Then as $t \to 0_+$, we have

$$\lim_{t\to 0_+} \left(h_{jk} - \nabla_{V_j}\nabla_{V_k}\psi_t\right)(x,t) = 0, \text{ for } 1 \le j,k \le n. \tag{2}$$

This is due to the 2-jet relation

$$\begin{aligned}
&\lim_{t\to 0_+} \left\langle \nabla_{V_i}\psi_t, \nabla_{V_j}\nabla_{V_k}\psi_t \right\rangle (x,t) \\
&= \lim_{t\to 0_+} (4\pi t)^{n/2} (2t) \left\langle \nabla_{V_i}\Phi_t, \nabla_{V_j}\nabla_{V_k}\Phi_t \right\rangle (x,t) = 0 \qquad (3)
\end{aligned}$$

from (15), so the second order terms in the Taylor expansion of $\psi_t : M \to l^2$ approximate the second fundamental form of $\psi_t(M) \subset l^2$ as $t \to 0_+$. Therefore, for the *mean curvature* $H(x,t) = \frac{1}{n}\Sigma_{k=1}^n h_{kk}(x)$, we have

$$\lim_{t\to 0_+} \left(H(x,t) - \frac{\Sigma_{k=1}^n \nabla_{V_k}\nabla_{V_k}\psi_t}{n} \right)(x,t) = 0. \qquad (4)$$

Item 1 was proved in [WZ] using the 2-jet relations of ψ_t in Theorem 1. For Item 2, using the 3-jet relations in Theorem 2 for $\overrightarrow{\alpha} = (i,k,k)$ and $\overrightarrow{\beta} = (j)$, we have

$$\begin{aligned}
&\frac{\left\langle \nabla_{V_i}\nabla_{V_k}\nabla_{V_k}\psi_t, \nabla_{V_j}\psi_t \right\rangle (x,t)}{\left|\nabla_{V_i}\nabla_{V_k}\nabla_{V_k}\psi\right| \left|\nabla_{V_j}\psi_t\right| (x,t)} \to B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) \\
&= \begin{cases} (-1)^{\frac{2}{2}} \cdot \frac{(4!/2!)}{[(6!/3!)(2!/1!)]^{1/2}} = -\sqrt{\frac{3}{5}}, & \text{if } i=j=k, \\ (-1)^{\frac{2}{2}} \cdot \frac{(2!/1!)}{[(2!/1!)(2!/1!)]^{1/2}} \cdot \frac{(2!/1!)}{[(4!/2!)(0!/0!)]^{1/2}} = -\frac{1}{\sqrt{3}}, & \text{if } i=j\neq k, \\ 0, & \text{if } i \neq j, \end{cases}
\end{aligned}$$

and

$$\begin{aligned}
\left|\nabla_{V_i}\nabla_{V_k}\nabla_{V_k}\psi_t(x,t)\right| &\to \sqrt{2}(4\pi)^{n/4} t^{\frac{n+2}{4}} \cdot \left(\frac{1}{4\pi t}\right)^{n/4} \left(\frac{1}{2t}\right)^{3/2} A\left(\overrightarrow{\alpha}, \overrightarrow{\alpha}\right)^{1/2} \\
&= \begin{cases} \frac{1}{2t}\sqrt{\frac{6!}{2^3 3!}} = \frac{1}{2t}\sqrt{15}, \text{ if } i=k \\ \frac{1}{2t}\sqrt{\frac{4!}{2^2 2!}\cdot\frac{2!}{2^1 2!}} = \frac{1}{2t}\sqrt{3}, \text{ if } i\neq k \end{cases}, \\
\left|\nabla_{V_j}\psi_t(x,t)\right| &\to \sqrt{2}(4\pi)^{n/4} t^{\frac{n+2}{4}} \cdot \left(\frac{1}{4\pi t}\right)^{n/4} \left(\frac{1}{2t}\right)^{1/2} A\left(\overrightarrow{\beta}, \overrightarrow{\beta}\right)^{1/2} = 1.
\end{aligned}$$

Therefore as $t \to 0_+$,

$$\left\langle \nabla_{V_i}\nabla_{V_k}\nabla_{V_k}\psi_t, \nabla_{V_j}\psi_t \right\rangle (x,t) \to \left\{ \begin{array}{l} -\sqrt{\frac{3}{5}} \cdot \frac{1}{2t}\sqrt{15} = -\frac{3}{2t}, \text{ if } i=j=k, \\ -\frac{1}{\sqrt{3}} \cdot \frac{1}{2t}\sqrt{3} = -\frac{3}{2t}, \quad \text{if } i=j\neq k. \end{array} \right\} = -\frac{3}{2t}.$$

By the definition of the second fundamental tensor,

$$
\begin{aligned}
& 2t \cdot \lim_{t \to 0_+} \left\langle S(x,t)\left(\nabla_{V_i}\psi_t, H\right), \nabla_{V_j}\psi_t \right\rangle (x,t) \\
&= 2t \cdot \lim_{t \to 0_+} \left\langle \nabla_{\nabla_{V_i}\psi_t} H\left(\psi_t(x), t\right), \nabla_{V_j}\psi_t \right\rangle \left(\psi_t(x), t\right) \\
&= 2t \cdot \lim_{t \to 0_+} \left\langle \nabla_{V_i} H(x,t), \nabla_{V_j}\psi_t \right\rangle (x,t) \\
&= 2t \cdot \lim_{t \to 0_+} \left\langle \nabla_{V_i} \left(\frac{1}{n} \sum_{k=1}^{n} \nabla_{V_k}\nabla_{V_k}\psi_t \right), \nabla_{V_j}\psi_t \right\rangle (x,t) \\
&= \frac{1}{n} 2t \cdot \sum_{k=1}^{n} \lim_{t \to 0_+} \left\langle \nabla_{V_i}\nabla_{V_k}\nabla_{V_k}\psi_t, \nabla_{V_j}\psi_t \right\rangle (x,t) \\
&= \begin{cases} \frac{1}{n} 2t \cdot \sum_{k=1}^{n} \left(-\frac{3}{2t}\right), & \text{if } i = j, \\ 0, & \text{if } i \neq j, \end{cases} \\
&= -3\delta_{ij}.
\end{aligned}
$$

Since $\left\{\nabla_{V_j}\psi_t\right\}_{1 \le j \le n}$ span $T_{\psi_t(x)}\psi_t(M)$, this means for any $W \in T_{\psi_t(x)}\psi_t(M)$,

$$
t \cdot \lim_{t \to 0_+} S(x,t) \langle W, H \rangle = -\frac{3}{2} W.
$$

So $\psi_t(M)$ is asymptotically umbilical in the mean curvature vector direction as $t \to 0_+$. □

Remark 5. In other words, for the image $\frac{1}{\sqrt{t}}\psi_t(M) \subset l^2$, as $t \to 0_+$ its mean curvature vector converges to constant length $\sqrt{\frac{n+2}{2n}}$, and $\frac{1}{\sqrt{t}}\psi_t(M)$ is asymptotically umbilical in the mean curvature vector direction. If we want to construct submanifolds in $\mathbb{R}^q$ with constant mean curvature length, perhaps we can truncate $\frac{1}{\sqrt{t}}\psi_t(M) \subset l^2$ to $\mathbb{R}^q$ by taking the first q components for large q, and then perturb it by the implicit function theorem.

3.2 Riemannian curvature tensor

In the second application we give a *heat kernel embedding* interpretation of the Levi-Civita connection and the Riemannian curvature tensor. The idea is that $\psi_t : M \to l^2$ is an almost isometric embedding, so the "extrinsic" constructions of the Levi-Civita connection and Gaussian curvature of surfaces in $\mathbb{R}^3$ using the ambient space can be mimicked here, with $\mathbb{R}^3$ replaced by l^2. The difficulty is that we need to take the limits as $t \to 0_+$ in our interpretation, and to show that these limits exist, we need the high-jet relations in Proposition 3 and some refinement.

As pointed to the author by L. I. Nicolaescu, there exists a *random function* interpretation of the Levi-Civita connection and the Riemannian curvature tensor in Section 12 of [AT] (equation (12.2.6) and Lemma 12.2.1, respectively). If the random function f in [AT] is taken as $f_t = \sum_{j\geq 1} u_j \phi_j(x)$ with the coefficients u_j being independent Gaussian random variables with the expectation $E(u_j) = 0$ and variance $Var(u_j) = e^{-\lambda_j t}$, then Lemma 2 and Proposition 5 follows from the results in [AT] by taking limits as $t \to 0$. (For this approach, see [Ni], appendix B.)

As before, for any $p \in M$, we choose the normal coordinates $\left(x^1, \cdots x^n\right)$ at p, and choose an orthonormal frame field $\{V_j\}_{j=1}^n$ near p such that $\{V_j\}_{i=1}^n|_p = \left\{\frac{\partial}{\partial x^j}\right\}$. Let ∇ be the Levi-Civita connection of (M, g), and $\nabla_{V_j} := \nabla_j$ for $1 \leq j \leq n$.

Lemma 2. *(Levi-Civita Connection) For any $X, Y \in T_pM$, we have*

$$\begin{aligned}\nabla_X Y(p) &= \lim_{t\to 0_+} \sum_{j=1}^n \left\langle \nabla_X \nabla_Y \psi_t, \nabla_{V_j}\psi_t \right\rangle (p) V_j(p) \qquad (5)\\ &= \lim_{t\to 0_+} d\psi_t^{-1}\left[\left(\nabla_X \nabla_Y \psi_t(p)\right)^{\top}\right],\end{aligned}$$

where $\top : T_{\psi_t(p)} l^2 \to T_{\psi_t(p)} \psi_t(M)$ is the orthogonal projection with respect to the standard metric on l^2, and on the right hand side ∇_X and ∇_Y are the derivatives for l^2-valued functions $f : M \to l^2$.

Proof. Since all quantities in (5) are linear in X, we may assume $X = V_i$. In a neighborhood of p, we write $Y(x) = \sum_{k=1}^n y^k(x) V_k(x)$. Since $\nabla_{V_i} V_j(p) = \delta_{ij}$, we have

$$\nabla_{V_i} Y(p) = \sum_{k=1}^n \nabla_i y^k(p) V_k(p),$$

where $\partial_i y^k(x) = \nabla_{V_i} y^k(x)$. On the other hand, by (3) and the 2-jet relation (3), we have

$$\begin{aligned}&\lim_{t\to 0_+} \sum_{i=1}^n \langle \nabla_i \nabla_Y \psi_t, \nabla_j \psi_t\rangle (p) V_j(p)\\ &= \lim_{t\to 0_+} \sum_{i=1}^n \left\langle \nabla_i \left(\sum_{k=1}^n y^k \nabla_k \psi_t\right), \nabla_j \psi_t \right\rangle (p) V_j(p)\\ &= \sum_{i=1}^n \sum_{k=1}^n \lim_{t\to 0_+} \nabla_i y^k(p) \langle \nabla_k \psi_t, \nabla_j \psi_t\rangle (p) V_j(p)\\ &\quad + y^k(p) \langle \nabla_i \nabla_k \psi_t, \nabla_j \psi_t\rangle (p) V_j(p)\\ &= \sum_{i=1}^n \sum_{k=1}^n \left(\partial_i y^k(p) \delta_{kj} V_j(p) + 0\right)\\ &= \nabla_i y^k(p) V_k(p).\end{aligned}$$

Therefore the first identity

$$\nabla_X Y(p) = \lim_{t\to 0_+} \sum_{j=1}^{n} \langle \nabla_X \nabla_Y \psi_t, \nabla_j \psi_t \rangle (p) V_j(p)$$

is proved. Since $\{V_j\}_{j=1}^n$ is an orthonormal basis of T_pM, and ψ_t is an almost isometric embedding in the sense of (3), $\{\nabla_{V_j}\psi_t(p)\}_{j=1}^n$ is asymptotically an orthonormal basis of $T_{\psi_t(p)}\psi_t(M)$ as $t \to 0_+$, and we have

$$\lim_{t\to 0_+}\left[(\nabla_X \nabla_Y \psi_t(p))^\top - \sum_{j=1}^{n} \langle \nabla_X \nabla_Y \psi_t, \nabla_{V_j}\psi_t \rangle (p) \nabla_{V_j}\psi_t(p)\right] = 0.$$

Applying $d\psi_t^{-1}$ on both sides, noting $d\psi_t^{-1}\left(\nabla_{V_j}\psi_t(p)\right) = V_j$, the second identity is proved. □

Suppose $f : (M, g) \to \mathbb{R}^q$ is an isometric embedding, and h is the second fundamental form of M. It is well-known that for vectors $X, Y, Z, W \in TM$, the Riemannian curvature tensor R of M satisfies

$$R(X, Y, Z, W) = \langle h(X, W), h(Y, Z)\rangle - \langle h(X, Z), h(Y, W)\rangle .$$

This is the *Gauss formula*, which relates the intrinsic curvature R of (M, g) with its extrinsic geometry in the ambient space. In our case $\psi_t : M \to l^2$ is asymptotically isometric as $t \to 0_+$, and by (2) the second fundamental form $h(X, Y)$ of $\psi_t(M) \subset l^2$ is approximated by $\nabla_X \nabla_Y \psi_t$ as $t \to 0_+$, so it is natural to expect the following "asymptotic" Gauss formula

Proposition 5. *(Riemannian curvature) For any X, Y, Z, W of TM, we have*

$$R(X, Y, Z, W) = \lim_{t\to 0+} \left[\langle \nabla_X \nabla_W \psi_t, \nabla_Y \nabla_Z \psi_t\rangle - \langle \nabla_X \nabla_Z \psi_t, \nabla_Y \nabla_W \psi_t\rangle\right]. \tag{6}$$

Proof. We only need to prove the above identity when X, Y, Z and W are taken from the orthonormal basis $\{V_j\}_{i=1}^n$ near x as above, say they are V_i, V_j, V_k and V_l respectively. For general vector fields, we can express $X = \sum_{j=1}^{n} a^j(x) V_j(x)$ etc., and plug them in (6). Then by the high-jet relations (3) and (3), as $t \to 0_+$ the terms of the type $\langle \nabla_k a^i \nabla_i \psi_t, \nabla_l b^j \nabla_j \psi_t\rangle$ will vanish or cancel each other, and the terms of the type $\langle \nabla_k a^i \nabla_i \psi_t, \nabla_j \psi_t \nabla_k \psi_t\rangle$ will converge to 0. For the same reason, switching the ordering of the covariant derivatives in (6), say from $\nabla_X \nabla_W \psi_t$ to $\nabla_W \nabla_X \psi_t$ etc., will not affect the limit in (6).

We first refine our estimate of $\lim_{t\to 0_+} \nabla^x_{V_i} \nabla^x_{V_l} \nabla^y_{V_j} \nabla^y_{V_k} H(t, x, y)|_{x=y}$ in Proposition 3. We have proved that

$$\lim_{t\to 0_+} (4\pi t)^{\frac{n}{2}} (2t)^2 D_x^{\overrightarrow{\alpha}} D_y^{\overrightarrow{\beta}} H(t, x, y)\Big|_{x=y} = A\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right)$$

for $\overrightarrow{\alpha}=(i,l)$ and $\overrightarrow{\beta}=(j,k)$, where $A\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$ is the leading (constant) term of $(4\pi t)^{\frac{n}{2}}(2t)^2 D_x^{\overrightarrow{\alpha}} D_y^{\overrightarrow{\beta}} H(t,x,y)\Big|_{x=y}$ as $t\to 0_+$, but to prove our theorem we will need the next order term, which is linear in t. Recall in Proposition 3, for $\vec{\gamma}=\left(\overrightarrow{\alpha},\overrightarrow{\overrightarrow{\beta}}\right)$, (where the notation $\overrightarrow{\overrightarrow{\beta}}$ means the derivative $D^{\overrightarrow{\beta}}$ is with respect to the y-variables), we write $D^{\overrightarrow{\gamma}}H(t,x,y)=\frac{1}{(4\pi t)^{n/2}}e^{-\frac{r^2}{4t}}P_{\overrightarrow{\gamma}}(t,x,y)$, where $P_{\overrightarrow{\gamma}}(t,x,y)$ is a polynomial in $D^{\overrightarrow{\mu_s}}\left(-\frac{r^2(x,y)}{4t}\right)$ and $D^{\overrightarrow{\eta}}U(t,x,y)\Big|_{x=y}$, i.e. each summand is of the form

$$\left(\prod_{s=1}^{h} D^{\overrightarrow{\mu_j}}\left(-\frac{r^2(x,y)}{4t}\right)\right)\Bigg|_{x=y}\cdot D^{\overrightarrow{\eta}}U(t,x,y)\Big|_{x=y}.$$

Using Lemma 1 and the total degree condition $|\vec{\gamma}|=4$, by the same method in Proposition 3 to find the leading terms in $P_{\overrightarrow{\gamma}}(t,x,y)$ of order $\left(\frac{1}{t}\right)^2$, we see the secondary terms of order $\left(\frac{1}{t}\right)$ must have $h=1$, $|\overrightarrow{\mu_1}|=2$, and $|\overrightarrow{\eta}|=2$, or $h=1$, $|\overrightarrow{\mu_1}|=4$, and $\overrightarrow{\eta}=0$, namely of the form

$$D^{\overrightarrow{\mu_1}}\left(-\frac{r^2(x,y)}{4t}\right)\Bigg|_{x=y}\cdot D^{\overrightarrow{\eta}}U(t,x,y)\Big|_{x=y},\ \text{with } \overrightarrow{\mu_1}\cup\overrightarrow{\eta}=\overrightarrow{\gamma},\ |\overrightarrow{\mu_1}|=|\overrightarrow{\eta}|=2,$$

or

$$D^{\overrightarrow{\gamma}}\left(-\frac{r^2(x,y)}{4t}\right)\Bigg|_{x=y}\cdot U(t,x,y)|_{x=y}=-\frac{1}{6t}\left(R_{ijkl}(x)+R_{ilkj}(x)\right),$$

where we have used (12). So as $t\to 0_+$, we have the refined convergence

$$\begin{aligned}
&(4\pi t)^{n/2}(2t)^2\nabla^x_{V_i}\nabla^x_{V_l}\nabla^y_{V_j}\nabla^y_{V_k}H(t,x,y)\,|_{x=y}\\
=\;&A\left(\overrightarrow{(i,l)},\overrightarrow{(j,k)}\right)\\
&-t\sum_{|\overrightarrow{\mu_1}|=|\overrightarrow{\eta}|=2,\ \overrightarrow{\mu_1}\cup\overrightarrow{\eta}=\left(\overrightarrow{(i,l)},\overrightarrow{\overline{(j,k)}}\right)}\left(D^{\overrightarrow{\mu_1}}\left(r^2(x,y)\right)\Big|_{x=y}\cdot D^{\overrightarrow{\eta}}U(t,x,y)\Big|_{x=y}\right)\\
&-\frac{2}{3}t\left(R_{ijkl}(x)+R_{ilkj}(x)\right)+O\left(t^2\right).
\end{aligned}\tag{7}$$

For $\overrightarrow{\gamma}=\left(\overrightarrow{(i,l)},\overrightarrow{\overline{(j,k)}}\right)$ and $\overrightarrow{\gamma}=\left(\overrightarrow{(i,k)},\overrightarrow{\overline{(j,l)}}\right)$, by considering all partitions of $\overrightarrow{\gamma}$ into $\overrightarrow{\mu_1}\cup\overrightarrow{\eta}$ with $|\overrightarrow{\mu_1}|=2$, and $|\overrightarrow{\eta}|=2$, and noticing $\nabla^a_x\nabla^b_y f(x,y)\,|_{x=y}=\nabla^a_y\nabla^b_x f(x,y)\,|_{x=y}$ for any indices a,b and functions f symmetric to x and y (here $r^2(x,y)$ and $U(t,x,y)$), we have

$$\sum_{\left|\overrightarrow{\mu_1}\right|=\left|\overrightarrow{\eta}\right|=2,\ \overrightarrow{\mu_1}\cup\overrightarrow{\eta}=\left(\overrightarrow{(i,l)},\overrightarrow{(j,k)}\right)}\left(D^{\overrightarrow{\mu_1}}\left(r^2(x,y)\right)\Big|_{x=y}\cdot D^{\overrightarrow{\eta}}U(t,x,y)\Big|_{x=y}\right)$$
$$=\sum_{\left|\overrightarrow{\mu_1}\right|=\left|\overrightarrow{\eta}\right|=2,\ \overrightarrow{\mu_1}\cup\overrightarrow{\eta}=\left(\overrightarrow{(i,k)},\overrightarrow{(j,l)}\right)}\left(D^{\overrightarrow{\mu_1}}\left(r^2(x,y)\right)\Big|_{x=y}\cdot D^{\overrightarrow{\eta}}U(t,x,y)\Big|_{x=y}\right).$$

We also have $A\left(\overrightarrow{(i,l)},\overrightarrow{(j,k)}\right)=A\left(\overrightarrow{(i,k)},\overrightarrow{(j,l)}\right)$ by the definition of $A\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$. Therefore, from (7) we have

$$\begin{aligned}
&\lim_{t\to 0_+}\left[\left\langle \nabla_{V_i}\nabla_{V_l}\psi_t,\nabla_{V_j}\nabla_{V_k}\psi_t\right\rangle(x)-\left\langle \nabla_{V_i}\nabla_{V_k}\psi_t,\nabla_{V_j}\nabla_{V_l}\psi_t\right\rangle(x)\right]\\
&=\lim_{t\to 0_+}\frac{1}{2t}\left[(4\pi t)^{n/2}(2t)^2\left(\nabla^x_{V_i}\nabla^x_{V_l}\nabla^y_{V_j}\nabla^y_{V_k}H(t,x,y)-\nabla^x_{V_i}\nabla^x_{V_k}\nabla^y_{V_j}\nabla^y_{V_l}H(t,x,y)\right)|_{x=y}\right]\\
&=\lim_{t\to 0_+}\frac{1}{2t}\frac{2t}{3}\left[-R_{ijlk}(x)-R_{iklj}(x)+R_{ijkl}(x)+R_{ilkj}(x)+O\left(t^2\right)\right]\\
&=\frac{1}{3}\lim_{t\to 0_+}\left[R_{ijkl}(x)+R_{ikjl}(x)+R_{ijkl}(x)+R_{kjil}(x)+O\left(t^2\right)\right]\\
&=\frac{1}{3}\lim_{t\to 0_+}\left[2R_{ijkl}(x)+\left(R_{ikjl}(x)+R_{kjil}(x)\right)+O\left(t^2\right)\right]\\
&=\frac{1}{3}\lim_{t\to 0_+}\left[2R_{ijkl}(x)+R_{ijkl}(x)+O\left(t^2\right)\right]\\
&=R_{ijkl}(x),
\end{aligned}$$

where in the last four identities we have used the symmetries of the Riemannian curvature tensor. □

Remark 6. There are several ways to express the curvature tensors by the eigenfunctions and the heat kernel on M, e.g. the Ricci curvature (8) and scalar curvature (3) in [BBG] and [Gil], but the expression (6) has a clear geometric meaning: the Riemannian curvature tensor on (M,g) is computed from the Gauss map of $\psi_t(M)\to S^\infty$ in l^2 as $t\to 0_+$. One more interesting point is that although our construction of the Riemannian curvature tensor via the embedding $\psi_t: M\to l^2$ looks “extrinsic”, it is actually “intrinsic”, since the embedding ψ_t is constructed by the eigenfunctions of Δ_g.

If we only want to prove the *existence* of the limit in (6), we only need the cancellation of the leading terms $A\left(\overrightarrow{(i,l)},\overrightarrow{(j,k)}\right)$ and $A\left(\overrightarrow{(i,k)},\overrightarrow{(j,l)}\right)$ in (7), but do not need the symmetries of the Riemannian curvature tensor. On the other hand, (6) *manifests the symmetries* of the Riemannian curvature tensor in a rather direct way. From (6), it is easy to see $R(X,Y)Z=-R(Y,X)Z$, $\langle R(X,Y)Z,W\rangle=-\langle R(X,Y)W,Z\rangle$, $\langle R(X,Y)Z,W\rangle=\langle R(Z,W)X,Y\rangle$, and $R(X,Y)Z+R(Y,Z)X+R(Z,X)Y=0$ (the first Bianchi identity). (Note that switching the order of differentiation in the second derivatives, say from $\nabla_X \nabla_Z \psi_t$ to $\nabla_Z \nabla_X \psi_t$, will not affect the limit in (6), as remarked in the beginning of the proof of Proposition 5.) With a little computation the second Bianchi identity also follows from our formula of R, as follows:

Lemma 3. *(Second Bianchi Identity) For any coordinates* $\left(x^1, \cdots, x^n\right)$ *near* x*, the Riemannian curvature tensor* R *satisfies*

$$\frac{\partial}{\partial x^h} R_{klij} + \frac{\partial}{\partial x^k} R_{lhij} + \frac{\partial}{\partial x^l} R_{hkij} = 0. \tag{8}$$

Proof. Since all terms on the left side are tensors, we only need to prove the identity when $\left(x^1, \cdots, x^n\right)$ is a normal coordinate near x. From (6) we have

$$\begin{aligned}
\frac{\partial}{\partial x^h} R_{klij}(x) &= \nabla_h \lim_{t \to 0+} \left[\langle \nabla_k \nabla_j \psi_t, \nabla_l \nabla_i \psi_t \rangle - \langle \nabla_k \nabla_i \psi_t, \nabla_l \nabla_j \psi_t \rangle \right](x) \\
&= \lim_{t \to 0+} \left[\langle \nabla_{hkj} \psi_t, \nabla_{li} \psi_t \rangle + \langle \nabla_{hli} \psi_t, \nabla_{kj} \psi_t \rangle \right. \\
&\quad \left. - \langle \nabla_{hki} \psi_t, \nabla_{lj} \psi_t \rangle - \langle \nabla_{hlj} \psi_t, \nabla_{ki} \psi_t \rangle \right](x),
\end{aligned}$$

where $\nabla_{hkj} := \nabla_h \nabla_k \nabla_j$. Similarly

$$\begin{aligned}
\frac{\partial}{\partial x^k} R_{lhij}(x) &= \lim_{t \to 0+} \left[\langle \nabla_{klj} \psi_t, \nabla_{hi} \psi_t \rangle + \langle \nabla_{khi} \psi_t, \nabla_{lj} \psi_t \rangle \right. \\
&\quad \left. - \langle \nabla_{kli} \psi_t, \nabla_{hj} \psi_t \rangle - \langle \nabla_{khj} \psi_t, \nabla_{li} \psi_t \rangle \right](x), \\
\frac{\partial}{\partial x^l} R_{hkij}(x) &= \lim_{t \to 0+} \left[\langle \nabla_{lhj} \psi_t, \nabla_{ki} \psi_t \rangle + \langle \nabla_{lki} \psi_t, \nabla_{hj} \psi_t \rangle \right. \\
&\quad \left. - \langle \nabla_{lhi} \psi_t, \nabla_{kj} \psi_t \rangle - \langle \nabla_{lkj} \psi_t, \nabla_{hi} \psi_t \rangle \right](x).
\end{aligned}$$

Adding the 3 identities together, (8) follows. There is a convergence issue of limits like $\lim_{t \to 0+} \langle \nabla_{hkj} \psi_t, \nabla_{li} \psi_t \rangle = O\left(\frac{1}{t}\right)$ from Proposition 3, but writing

$$\frac{\partial}{\partial x^h} R_{klij} + \frac{\partial}{\partial x^k} R_{lhij} + \frac{\partial}{\partial x^l} R_{hkij} = \frac{1}{2t} \left[2t \left(\frac{\partial}{\partial x^h} R_{klij} + \frac{\partial}{\partial x^k} R_{lhij} + \frac{\partial}{\partial x^l} R_{hkij} \right) \right]$$

and applying Proposition 3 easily fix that. □

3.3 Approximation of Riemannian geometry by eigenfunctions

For the embedding map $\psi_t : (M, g) \to l^2$, we can truncate the infinite dimensional l^2 to $\mathbb{R}^q$ by taking its first q components, and denote the resulted map by ψ_t^q. The following truncation lemma in [WZ] gives the bound of q that can preserve most good properties of ψ_t. It was proved by the Weyl's asymptotic formula of eigenvalues (see [Ch]) and C^0-estimates of eigenfunctions (and their derivatives) on compact Riemannian manifolds.

Lemma 4. *([WZ] Theorem 6 and Remark 7) Let $\rho > 0$ be a fixed small constant. For integers $q = q(t)$ of the order $t^{-\left(\frac{n}{2}+\rho\right)}$ as $t \to 0_+$, the truncated heat kernel mapping*

$$\psi_t^{q(t)} : (M, g) \to \mathbb{R}^{q(t)}$$

still satisfies the asymptote

$$\left(\psi_t^{q(t)}\right)^* g_0 = g + \frac{t}{3}\left(\frac{1}{2}S_g \cdot g - Ric_g\right) + O\left(t^2\right), \tag{9}$$

where g_0 is the standard metric in l^2. The truncation order $q(t) \sim t^{-\left(\frac{n}{2}+\rho\right)}$ is also valid to preserve all high-jet relations in Theorem 2 for higher derivatives of ψ_t.

Using the above truncation lemma, we can approximate the metric g, the Levi-Civita connection ∇, and the Riemannian curvature R by $q(t) \sim t^{-\frac{n}{2}-\rho}$ eigenfunctions by the expressions (3), (5), and (6), respectively, with the error of certain orders of t. This suggests the possibility of approximating the Riemannian geometry of (M, g) by finitely many eigenfunctions. Indeed, if the Schwartz function w is chosen with compact support, Theorem 1.6 of [Ni] gives one of such approximations.

It will be useful to get a precise control of the coefficients in the above remainder $O\left(t^2\right)$ and their higher derivative counterparts, so we know how large $q(t)$ (i.e. how small t) should be in order to yield a given accuracy. For recent progress on truncation dimension of heat kernel embedding maps, and for Riemannian metrics of lower regularity (only $C^{2,\alpha}$), see [CW] and [Po].

4 Algebraic structures in the ∞-jet space of $\Phi_t(x)$

In Section 2 we have obtained the angle and length relations of all derivative vectors $D^{\overrightarrow{\alpha}}\Phi_t(x)$ as $t \to 0_+$. These relations are independent on g, x and the choice of orthonormal basis, so it is worthwhile to investigate them more closely. The relations are expressed by the constants $A\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right)$ and $B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right)$. In this section we explore more relations between these constants, or in other words study the ∞-jet space of $\Phi_t(x)$ as $t \to 0_+$.

4.1 Inductive relations of $A\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right)$ and $B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right)$

It turns out the constants $A\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$ and $B\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$ involved in the high-jet relations in Theorem 2 can be defined inductively.

Definition 4. Suppose $j_r \in \{1,2,\cdots n\}$, but not necessarily in $\mathcal{S}=\{\overrightarrow{\alpha}\} \amalg \left\{\overrightarrow{\beta}\right\}$. Let $\overrightarrow{\alpha}_+ = (\overrightarrow{\alpha}, j_r)$, $\overrightarrow{\beta}_+ = \left(\overrightarrow{\beta}, j_r\right)$, and $\overrightarrow{\alpha_{+k}} = \left(\overrightarrow{\alpha}, \underbrace{j_r,\cdots,j_r}_{k}\right)$. If $j_r \in \overrightarrow{\beta}$, we let $\overrightarrow{\beta}_- = \overrightarrow{\beta} \setminus \{j_r\}$.

By definition it is easy to check $A\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$ satisfies the following four properties:

$$\text{Normalization:}\quad A\left(\overrightarrow{\varnothing},\overrightarrow{\varnothing}\right)=1,\ \text{(where }\varnothing\text{ is the empty set)} \tag{1}$$

$$\text{Symmetry:}\quad A\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)=A\left(\overrightarrow{\beta},\overrightarrow{\alpha}\right), \tag{2}$$

$$\text{Leibniz rule:}\quad A\left(\overrightarrow{\alpha}_+,\overrightarrow{\beta}_-\right)=-A\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right), \tag{3}$$

$$\text{Adding index:}\quad A\left(\overrightarrow{\alpha}_+,\overrightarrow{\beta}_+\right)=A\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)(a_r+b_r+1). \tag{4}$$

Proposition 6. *The above four properties completely characterize* $A\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$.

Proof. To obtain a general $A\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$, we can start from $\overrightarrow{\alpha}=\overrightarrow{\beta}=\overrightarrow{\varnothing}$, use (4) to obtain $A\left(\overrightarrow{\varnothing_{+k}},\overrightarrow{\varnothing_{+k}}\right)$, and then use (3) to obtain $A\left(\overrightarrow{\varnothing_{+(k+l)}},\overrightarrow{\varnothing_{+(k-l)}}\right)$ for $0\le l\le k$. Then we can use (4) to add different indices j' to $\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$ to get all $A\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$ (note if $j'\notin\{\overrightarrow{\alpha}\}\amalg\left\{\overrightarrow{\beta}\right\}$, then $A\left(\overrightarrow{\alpha_+},\overrightarrow{\beta_+}\right)=A\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$ by (4)). □

Similarly, $B\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$ is characterized by the following four properties:

$$\begin{gathered}B\left(\overrightarrow{\varnothing},\overrightarrow{\varnothing}\right)=1,\ B\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)=B\left(\overrightarrow{\beta},\overrightarrow{\alpha}\right),\\ B\left(\overrightarrow{\alpha}_+,\overrightarrow{\beta}_-\right)=-B\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)\left(\frac{2b_r-1}{2a_r+1}\right)^{1/2},\\ B\left(\overrightarrow{\alpha}_+,\overrightarrow{\beta}_+\right)=B\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)\frac{a_r+b_r+1}{\left[(2a_r+1)(2b_r+1)\right]^{1/2}}.\end{gathered} \tag{5}$$

4.2 Stabilization property from repeated differentials

For some index $j_r \in \{1, 2, \cdots n\}$ (but not necessarily in $\{\overrightarrow{\alpha}\} \amalg \left\{\overrightarrow{\beta}\right\}$), if we apply ∂_{j_r} repeatedly to $D^{\overrightarrow{\alpha}}\Phi_t(x)$ and $D^{\overrightarrow{\beta}}\Phi_t(x)$, then the the resulted derivative vectors become "more and more collinear", and the angle between them has certain stabilization property. More precisely, using the conventions in Definition 4, we have the following

Proposition 7. *$B\left(\overrightarrow{\alpha}_+, \overrightarrow{\beta}_+\right)$ and $B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right)$ always have the same sign, and*

$$\left|B\left(\overrightarrow{\alpha}_+, \overrightarrow{\beta}_+\right)\right| \geq \left|B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right)\right| , \tag{6}$$

where the equality holds if and only if j_r appears with the same multiplicity in $\overrightarrow{\alpha}$ and $\overrightarrow{\beta}$. Further more we have the following limit: if we let the multi-index $\overrightarrow{\alpha_{+k}} = \left(\overrightarrow{\alpha}, \underbrace{j_r, \cdots, j_r}_{k}\right)$, similarly for $\overrightarrow{\beta_{+k}}$, then

$$\lim_{k\to\infty} B\left(\overrightarrow{\alpha_{+k}}, \overrightarrow{\beta_{+k}}\right) := B\left(\overrightarrow{\alpha}_*, \overrightarrow{\beta}_*\right) , \tag{7}$$

where the multi-indices $\overrightarrow{\alpha_*}$ and $\overrightarrow{\beta}_*$ are obtained from deleting all j_r indices in $\overrightarrow{\alpha}$ and $\overrightarrow{\beta}$ respectively. We also have

$$\lim_{k\to\infty} B\left(\overrightarrow{\alpha}, \overrightarrow{\beta_{+k}}\right) := 0. \tag{8}$$

Proof. Since $|\overrightarrow{\alpha}_+| - \left|\overrightarrow{\beta}_+\right| = |\overrightarrow{\alpha}| - \left|\overrightarrow{\beta}\right|$, by definition $B\left(\overrightarrow{\alpha}_+, \overrightarrow{\beta}_+\right)$ and $B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right)$ always have the same sign. Let us assume $B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) \geq 0$ (the $B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) \leq 0$ case is similar). From (5) we have

$$\begin{aligned} B\left(\overrightarrow{\alpha}_+, \overrightarrow{\beta}_+\right) &= B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) \cdot \left(\frac{(a_r + b_r + 1)^2}{(2a_r + 1)(2b_r + 1)}\right)^{1/2} \\ &= B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) \cdot \left(1 + \frac{(a_r - b_r)^2}{(2a_r + 1)(2b_r + 1)}\right)^{1/2} \\ &\geq B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right), \end{aligned}$$

and the equality holds if and only if $a_r = b_r$. Recall

$$B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) = (-1)^{\frac{|\overrightarrow{\alpha}| - \left|\overrightarrow{\beta}\right|}{2}} \prod_{r=1}^{s} \frac{(2\sigma_r)!}{\sigma_r!} \Big/ \left[\frac{(2a_r)!}{a_r!} \frac{(2b_r)!}{b_r!}\right]^{1/2} .$$

Since $\sigma_r = \frac{a_r+b_r}{2}$, if we let $l_r = \frac{b_r-a_r}{2} \geq 0$ (assuming $b_r \geq a_r$), then

$$\frac{a_r!}{\sigma_r!}\frac{b_r!}{\sigma_r!} = \frac{(\sigma_r+1)\cdots(\sigma_r+l_r)}{(\sigma_r-1)\cdots(\sigma_r-l_r)}, \text{ and } \frac{(2a_r)!}{(2\sigma_r)!}\frac{(2b_r)!}{(2\sigma_r)!} = \frac{(2\sigma_r+1)\cdots(2\sigma_r+2l_r)}{(2\sigma_r-1)\cdots(2\sigma_r-2l_r)}.$$

So the r-th factor in $B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right)$ is

$$\begin{aligned} \frac{(2\sigma_r)!}{\sigma_r!} \Big/ \left[\frac{(2a_r)!}{a_r!}\frac{(2b_r)!}{b_r!}\right]^{1/2} &= \left[\frac{(2\sigma_r)!}{(2a_r)!}\frac{(2\sigma_r)!}{(2b_r)!}\frac{a_r!}{\sigma_r!}\frac{b_r!}{\sigma_r!}\right]^{1/2} \\ &= \left[\frac{(2\sigma_r-1)(2\sigma_r-3)\cdots(2\sigma_r-2l_r+1)}{(2\sigma_r+1)(2\sigma_r+3)\cdots(2\sigma_r+2l_r-1)}\right]^{1/2} \leq 1. \end{aligned} \tag{9}$$

As $k \to \infty$, in the multi-index $\left(\overrightarrow{\alpha_{+k}}, \overrightarrow{\beta_{+k}}\right)$, σ_r is changed to $\sigma_r + k$, but l_r is unchanged, so the factor involving j_r in $B\left(\overrightarrow{\alpha_{+k}}, \overrightarrow{\beta_{+k}}\right)$ becomes

$$\left[\frac{(2\sigma_r+2k-1)(2\sigma_r+2k-3)\cdots(2\sigma_r+2k-2l_r+1)}{(2\sigma_r+2k+1)(2\sigma_r+2k+3)\cdots(2\sigma_r+2k+2l_r-1)}\right]^{1/2} \to 1$$

as $k \to \infty$, and the other factors not involving j_r are unchanged. This gives the limit $B\left(\overrightarrow{\alpha_*}, \overrightarrow{\beta_*}\right)$.

For (8), w.l.o.g we can assume $|\overrightarrow{\alpha}| + \left|\overrightarrow{\beta}\right|$ is even (otherwise we can replace $\overrightarrow{\beta}$ by $\overrightarrow{\beta}_+$ before we take limit). In the multi-index $\left(\overrightarrow{\alpha}, \overrightarrow{\beta_{+2k}}\right)$, σ_r and l_r are changed to $\sigma_r + k$ and $l_r + k$ respectively, so the factor involving j_r in $B\left(\overrightarrow{\alpha}, \overrightarrow{\beta_{+2k}}\right)$ becomes

$$\left[\frac{(2\sigma_r+2k-1)(2\sigma_r+2k-3)\cdots(2\sigma_r-2l_r+1)}{(2\sigma_r+2k+1)(2\sigma_r+2k+3)\cdots(2\sigma_r+4k+2l_r-1)}\right]^{1/2} \leq \left[\frac{2\sigma_r-2l_r+1}{2\sigma_r+4k+2l_r-1}\right]^{1/2} \to 0$$

as $k \to \infty$. Therefore $B\left(\overrightarrow{\alpha}, \overrightarrow{\beta_{+2k}}\right) \to 0$. By our definition $B\left(\overrightarrow{\alpha}, \overrightarrow{\beta_{+(2k+1)}}\right) = 0$. Combining the two cases, $\lim_{k\to\infty} B\left(\overrightarrow{\alpha}, \overrightarrow{\beta_{+k}}\right) = 0$. $\square$

Corollary 1. $-1 < B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) \leq 1$, *and* $B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) = 1$ *if and only if* $\overrightarrow{\alpha} = \overrightarrow{\beta}$.

Proof. From (9) we see for any multi-indices $\overrightarrow{\alpha}$ and $\overrightarrow{\beta}$,

$$\left|B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right)\right| \leq 1. \tag{10}$$

If $B\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)=1$ but $\overrightarrow{\alpha}\neq\overrightarrow{\beta}$, then there must be some index j_r appearing with *different* multiplicities in $\overrightarrow{\alpha}$ and $\overrightarrow{\beta}$. We can construct $\overrightarrow{\alpha_+}$ and $\overrightarrow{\beta_+}$ by adding the index j_r. By (6) (notice the equality can *not* hold by the condition on j_r), we have

$$\left|B\left(\overrightarrow{\alpha}_+,\overrightarrow{\beta}_+\right)\right|>\left|B\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)\right|=1,$$

contradicting with (10). Therefore $\overrightarrow{\alpha}=\overrightarrow{\beta}$.

We claim $B\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)\neq-1$. Otherwise, by the sign of $B\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$, we see $\frac{|\overrightarrow{\alpha}|-|\overrightarrow{\beta}|}{2}$ must be odd, especially $\overrightarrow{\alpha}\neq\overrightarrow{\beta}$. Therefore there is some $l_r\neq 0$. By (9), this makes the r-th factor in $B\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$ have absolute value strictly less than 1, so $\left|B\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)\right|<1$. But this contradicts with $\left|B\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)\right|=|-1|=1$. □

4.3 Lattice geometry on $\mathbb{Z}_+^n$

In this subsection, we use the high-jet relations of Φ_t in Theorem 2 to put a metric d on the lattice space

$$\mathbb{Z}_+^n=\left\{(m_1,\cdots,m_n)\,|m_j\in\mathbb{Z},\,m_j\geq 0\text{ for }1\leq j\leq n\right\}.$$

The motivation is to visualize these relations by the geometry of $\left(\mathbb{Z}_+^n,d\right)$.

For any multi-index $\overrightarrow{\alpha}$ with indices in $\{1,2,\cdots,n\}$, taking the multiplicity of its each index, we obtain a point $z\in\mathbb{Z}_+^n$; conversely we can use z to recover $\overrightarrow{\alpha}$. For example, if we let z_1,z_2,z_3 be the three axes of $\mathbb{Z}_+^3$, then the multi-index $\overrightarrow{\alpha}=(1,1,2,3)$ has 1 with multiplicity 2, 2 with multiplicity 1, and 3 with multiplicity 1. So $\overrightarrow{\alpha}$ correspond to $z=(2,1,1)\in\mathbb{Z}_+^3$. With this identification, we have

Definition 5. (Angle distance) For any two multi-indices $\overrightarrow{\alpha}$ and $\overrightarrow{\beta}$ in $\mathbb{Z}_+^n$, we define their *angle distance* to be

$$d\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)=\cos^{-1}\left(B\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)\right).\tag{11}$$

By Theorem 2, $d\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$ is the *limiting distance* of $\frac{D^{\overrightarrow{\alpha}}\Phi_t(x)}{\left|D^{\overrightarrow{\alpha}}\Phi_t(x)\right|}$ and $\frac{D^{\overrightarrow{\beta}}\Phi_t(x)}{\left|D^{\overrightarrow{\beta}}\Phi_t(x)\right|}$ on the unit sphere $S^\infty\subset l^2$ as $t\to 0_+$. Let $\theta_t\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$ be the minimal distance of $\frac{D^{\overrightarrow{\alpha}}\Phi_t(x)}{\left|D^{\overrightarrow{\alpha}}\Phi_t(x)\right|}$ and $\frac{D^{\overrightarrow{\beta}}\Phi_t(x)}{\left|D^{\overrightarrow{\beta}}\Phi_t(x)\right|}$ on S^∞, then $0\leq\theta_t\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)\leq\pi$, and

$$\lim_{t\to 0_+} \theta_t\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) = d\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right). \tag{12}$$

Lemma 5. *The angle distance d defined in* (11) *is a metric on $\mathbb{Z}_+^n$. For any $\overrightarrow{\alpha}$ and $\overrightarrow{\beta} \in \mathbb{Z}_+^n$, $0 \le d\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) < \pi$.*

Proof. To prove d is a metric, we only need to verify that it is *nondegenerate*, since the triangle inequality of $\theta_t\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right)$ on S^∞ is preserved when we take the limit in (12) as $t \to 0_+$, and $d\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) = d\left(\overrightarrow{\beta}, \overrightarrow{\alpha}\right)$ follows from the definition of $B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right)$. Suppose $d\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) = 0$, then $B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) = 1$. From Corollary 1, we have $\overrightarrow{\alpha} = \overrightarrow{\beta}$. By Corollary 1, $B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) \in (-1, 1]$, so $0 \le d\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) < \pi$. $\square$

We can interpret the stabilization property in (7) of Proposition 7 in terms of the distance d on $\mathbb{Z}_+^n$:

Lemma 6. *(Stabilization) For any coordinate vector $v = \frac{\partial}{\partial z^j}$, and any two vectors $\overrightarrow{\alpha}, \overrightarrow{\beta} \in \mathbb{Z}_+^n$, we have*

$$\left|d\left(\overrightarrow{\alpha} + v, \overrightarrow{\beta} + v\right) - \frac{\pi}{2}\right| \ge \left|d\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) - \frac{\pi}{2}\right|,$$
$$\lim_{k\to\infty} d\left(\overrightarrow{\alpha} + kv, \overrightarrow{\beta} + kv\right) = d\left(\overrightarrow{\alpha}_*, \overrightarrow{\beta}_*\right),$$
$$\lim_{k\to\infty} d\left(\overrightarrow{\alpha}, \overrightarrow{\beta} + kv\right) = \frac{\pi}{2},$$

where $\overrightarrow{\alpha}_*$ and $\overrightarrow{\beta}_*$ are obtained by setting the j-th component of $\overrightarrow{\alpha}$ and $\overrightarrow{\beta}$ to be zero.

The stabilization properties in the above lemma make each axis of $\left(\mathbb{Z}_+^n, d\right)$ somehow looks like a real projective line. It will be interesting to say more about the geometry of $\left(\mathbb{Z}_+^n, d\right)$ as $|\overrightarrow{\alpha}| \to \infty$.

In the following we are interested in the orthogonal relations of $D^{\overrightarrow{\alpha}}\Phi_t(x)$ as $t \to 0_+$. From our definition, $B\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) = 0$ if and only if some index appears with odd multiplicity in $\{\overrightarrow{\alpha}\} \amalg \left\{\overrightarrow{\beta}\right\}$. This can be reformulated as

Lemma 7. *(Orthogonal relations) $d\left(\overrightarrow{\alpha}, \overrightarrow{\beta}\right) = \frac{\pi}{2}$ if and only if $\overrightarrow{\alpha} - \overrightarrow{\beta} \notin (2\mathbb{Z})^n$. Consequently,*

1. *Any two adjacent lattice points in $\mathbb{Z}_+^n$ have the angle distance $\frac{\pi}{2}$;*
2. *$\mathbb{Z}_+^n / (2\mathbb{Z}_+)^n$ has 2^n "orthogonal" cosets $A_1, \cdots, A_{2^n} \subset \mathbb{Z}_+^n$, such that for any two elements in different cosets, their angle distance is $\frac{\pi}{2}$.*

This $\mathbb{Z}_+^n/(2\mathbb{Z}_+)^n$-grading of the ∞-jet space of Φ_t as $t\to 0$ may have future applications. Locally, in an orthonormal coordinates $(x_1,\cdots,x_n)$ around $x\in M$, the Weyl algebra spanned by $\left\{x_1,\cdots,x_n,\frac{\partial}{\partial x_1},\cdots,\frac{\partial}{\partial x_n}\right\}$ acts on the ∞-jet space $j^\infty(x)$ of $\Phi_t(x)$, and when $t\to 0_+$, by the above lemma, $j^\infty(x)$ splits into 2^n asymptotically orthogonal summands that are permuted by the Weyl algebra action.

In the following we give some estimate of $d\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$. Let d_0 be the "*set distance*" function on $\mathbb{Z}_+^n$: for $\overrightarrow{\alpha}=(a_1,\cdots,a_n)$ and $\overrightarrow{\beta}=(b_1,\cdots,b_n)$,

$$d_0\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)=\sum_{j=1}^n|a_j-b_j|=\left|\overrightarrow{\alpha}\bigtriangleup\overrightarrow{\beta}\right|,$$

where $\bigtriangleup$ is the difference of two sets A and B: $A\bigtriangleup B=(A\backslash B)\cup(B\backslash A)$.

Intuitively, if $\left|\overrightarrow{\alpha}\bigtriangleup\overrightarrow{\beta}\right|$ is very large relative to $|\overrightarrow{\alpha}|+\left|\overrightarrow{\beta}\right|$, then $\overrightarrow{\alpha}$ and $\overrightarrow{\beta}$ are very different, and we expect the angle $d\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$ between the derivative vectors $D^{\overrightarrow{\alpha}}\Phi_t(x)$ and $D^{\overrightarrow{\beta}}\Phi_t(x)$ to be not too close to 0 or π as $t\to 0_+$. Indeed, we have

Proposition 8. *(Distance comparison) There exists a constant $\delta>0$ independent on n, such that for any multi-indices $\overrightarrow{\alpha}$ and $\overrightarrow{\beta}\in\mathbb{Z}_+^n$,*

$$\left|\cos d\left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)\right|\le 1-\delta\frac{\left|\overrightarrow{\alpha}\bigtriangleup\overrightarrow{\beta}\right|}{|\overrightarrow{\alpha}|+\left|\overrightarrow{\beta}\right|}. \tag{13}$$

Proof. Inequality (13) is equivalent to the following combinatorial inequality

$$\prod_{r=1}^s\frac{(2\sigma_r-1)(2\sigma_r-3)\cdots(2\sigma_r-2l_r+1)}{(2\sigma_r+1)(2\sigma_r+3)\cdots(2\sigma_r+2l_r-1)}\le 1-\delta\frac{\Sigma_{r=1}^s l_r}{\Sigma_{r=1}^s\sigma_r}, \tag{14}$$

where $\sigma_r\ge l_r\ge 0$ are integers. We first prove (14) when $s=1$. Let $f(x)=\ln\left(\frac{1-x}{1+x}\right)$, then $\acute{f}''(x)\le 0$ on $[0,1)$. Therefore

$$\frac{f\left(\frac{1}{2\sigma_r}\right)+\cdots+f\left(\frac{2l_r-1}{2\sigma_r}\right)}{l_r}\le f\left(\frac{\frac{1}{2\sigma_r}+\cdots+\frac{2l_r-1}{2\sigma_r}}{l_r}\right)=f\left(\frac{l_r}{2\sigma_r}\right),\text{ i.e.}$$

$$\left(\frac{1-\frac{1}{2\sigma_r}}{1+\frac{1}{2\sigma_r}}\right)\cdots\left(\frac{1-\frac{2l_r-1}{2\sigma_r}}{1+\frac{2l_r-1}{2\sigma_r}}\right)\le\left(\frac{1-\frac{l_r}{2\sigma_r}}{1+\frac{l_r}{2\sigma_r}}\right)^{l_r}\le\frac{1-\frac{l_r}{2\sigma_r}}{1+\frac{l_r}{2\sigma_r}}\le 1-\frac{l_r}{2\sigma_r}. \tag{15}$$

For $s>1$, using (15) we have

$$\prod_{r=1}^{s} \frac{(2\sigma_r - 1)(2\sigma_r - 3)\cdots(2\sigma_r - 2l_r + 1)}{(2\sigma_r + 1)(2\sigma_r + 3)\cdots(2\sigma_r + 2l_r - 1)}$$
$$\leq \prod_{r=1}^{s}\left(1 - \frac{l_s}{2\sigma_s}\right) \leq \left(\frac{\Sigma_{r=1}^{s}\left(1 - \frac{l_r}{2\sigma_r}\right)}{s}\right)^{s} \text{ (arithmetic-geometric inequality)}$$
$$= \left(1 - \frac{1}{2s}\Sigma_{r=1}^{s}\frac{l_r}{\sigma_r}\right)^{s} \leq \left(1 - \frac{1}{2s}\left(\frac{\Sigma_{r=1}^{s} l_r}{\Sigma_{r=1}^{s}\sigma_r}\right)\right)^{s} \text{ (since } \frac{a}{b} + \frac{c}{d} \geq \frac{a+c}{b+d} \text{ for } a,b,c,d > 0)$$
$$\leq 1 - \delta\frac{\Sigma_{r=1}^{s} l_r}{\Sigma_{r=1}^{s}\sigma_r},$$

where the last inequality is because there exists a $\delta > 0$ such that

$$\left(1 - \frac{1}{2s}x\right)^{s} \leq 1 - \delta x \tag{16}$$

on $[0,1]$ for any integer $s \geq 1$. This can be proved as follows: For each integer s, let $h_s(x) = \frac{1-\left(1-\frac{1}{2s}x\right)^{s}}{x}$. Defining $h_s(0) = \frac{1}{2}$ we see $h_s(x)$ is a continuous, positive function on $[0,1]$, thus h_s has a minimum $\delta_s > 0$ on $[0,1]$. Thus for all $x \in [0,1]$,

$$\left(1 - \frac{1}{2s}x\right)^{s} \leq 1 - \delta_s x.$$

Next we improve δ_s to a uniform constant independent on s. Note when $s \to \infty$, $h_s(x)$ tends to a constant function:

$$\lim_{s\to\infty} h_s(x) = \lim_{s\to\infty} \frac{1 - \left(1 - \frac{1}{2s}x\right)^{s}}{x} = \lim_{s\to\infty} \frac{1 - e^{-\frac{1}{2}x}}{x} = \frac{1}{2}.$$

Thus $\lim_{s\to\infty} \delta_s = \frac{1}{2}$, and the constant $\delta := \inf_{s\geq 1} \delta_s > 0$ makes the inequality (16) hold for all $x \in [0,1]$. $\square$

Inequality (13) gives a quantitative control of the linear independence of the derivative vectors $D^{\overrightarrow{\alpha}}\Phi_t(x)$ and $D^{\overrightarrow{\beta}}\Phi_t(x)$ as $t \to 0_+$. When $|\overrightarrow{\alpha}| = \left|\overrightarrow{\beta}\right| = 2$, the *uniform* linear independence of $D^{\overrightarrow{\alpha}}\Phi_t(x)$ and $D^{\overrightarrow{\beta}}\Phi_t(x)$ as $t \to 0_+$ played an important role in [WZ] to construct isometric embeddings by the implicit function theorem. It will be interesting to see if there are other geometric problems where the higher-jet linear independence property can play role.

References

[AT] R. Adler, R.J.E. Taylor, *Random Fields and Geometry*, Springer Monographs in Mathematics, Springer Verlag, 2007

[BBG] P. Bérard, G. Besson, S. Gallot, *Embedding Riemannian Manifolds by Their Heat kernel*, Geom. and Func. Analysis, vol. 4, No. 4, (1994)

[BeGaM] M. Berger, P. Gauduchon, E. Mazet, *Le spectre d'une variété Riemannianne*, Springer Lecture Notes in Math. 194, 1971

[BGV] N. Berline , E. Getzler , M. Vergne, *Heat Kernels and Dirac Operators*, SPRINGER VERLAG; 199, 2004

[Ch] I. Chavel, *Eigenvalues in Riemannian Geometry*, Vol 115, Pure and Applied Mathematics, Academic Press, 1984

[CW] C. Lin, H. Wu, *Embeddings of Riemannian manifolds with finite eigenvector fields of connection Laplacian,* Calculus of Variations and Partial Differential Equations volume 57, Article number: 126 (2018)

[De] B. DeWitt, *The Global Approach to Quantum Field Theory*, Vol 114, The International Series of Monographs on Physics, Oxford University Press, 2003

[Gil] P. Gilkey, *Asymptotic Formulae in Spectral Geometry,* (Studies in Advanced Mathematics), Chapman and Hall/CRC; 1st edition (December 17, 2003)

[Ni] L. I. Nicolaescu, *Random Morse functions and spectral geometry*, arXiv:1209.0639

[Po] J. Portegies, *Embeddings of Riemannian manifolds with heat kernels and eigenfunctions*, Communications on Pure and Applied Mathematics 69 (3), 478-518, 2016

[WZ] X. Wang, K. Zhu, *Isometric embeddings via heat kernel,* J. Diff. Geom., (99), 497-538, 2015