Diffusion Generator Convergence on Compact Riemannian Manifolds Without Boundary

Jacob Busby

November 2025

1 The Discrete Diffusion Framework

We follow the classical diffusion–maps construction of Coifman–Lafon (Coifman and Lafon 2006); for an informal overview see the Wikipedia article on diffusion maps (*Diffusion map* n.d.).

Consider a random walk on a large dataset, and fix a measure space $(X, \mathcal{A}, \mu)$, where X is the dataset, $\mathcal{A}$ its σ–algebra, and μ the sampling measure. We use a connectivity (affinity) kernel $k_\varepsilon : X \times X \to \mathbb{R}_{\geq 0}$ to weight nearby points more than distant ones. A common choice is the Gaussian kernel

$$k_\varepsilon(x, y) := \exp\left(-\frac{\|x - y\|^2}{\varepsilon}\right), \qquad \varepsilon > 0,$$

which is symmetric and nonnegative (Coifman and Lafon 2006; *Diffusion map* n.d.).

Definition 1.1. For $x \in X$, set

$$d(x) := \int_X k(x, y)\, d\mu(y),$$

which is the mass of the kernel centered at x with respect to the sample measure μ.

Definition 1.2 (normalized kernel)**.** Define

$$p(x, y) := \frac{k(x, y)}{d(x)}.$$

For each fixed x, $p(x, \cdot) \geq 0$ and $\int_X p(x, y)\, d\mu(y) = 1$; in general $p(x, y) \neq p(y, x)$.

Consider the transition matrix M whose entries M_{ij} are given by $p(x_i, x_j)$. Indeed, it can be shown that M constitutes a reversible discrete-time Markov chain, which we will show shortly.

Definition 1.3. For sample points $\{x_i\}_{i=1}^n \subset X$, define

$$L_{ij} := k(x_i, x_j).$$

and define the diagonal matrix $D \in \mathbb{R}^{n\times n}$ by $D_{ii} := \sum_{j=1}^n L_{ij}$.

Definition 1.4 (α–normalized kernel)**.** Fix $\alpha \in \mathbb{R}$. With $D_{ii} := \sum_{j=1}^n L_{ij}$ and $D = \mathrm{diag}(D_{11}, \dots, D_{nn})$, set

$$L_{ij}^{(\alpha)} := \frac{L_{ij}}{(D_{ii}D_{jj})^\alpha}, \qquad \text{equivalently} \qquad L^{(\alpha)} = D^{-\alpha} L D^{-\alpha}.$$

Definition 1.5. Define

$$M = (D^{(\alpha)})^{-1} L^{(\alpha)}, \qquad \text{where } D_{ii}^{(\alpha)} = \sum_j L_{ij}^{(\alpha)}.$$

We also write $M^{(\alpha)}$ for M to emphasize the dependence on α.

Lemma 1.6. The t–step transition probability of the stochastic random walk on the data is given by

$$p(x_j, t \mid x_i) = (M^t)_{ij}.$$

Proof. We will show that M is a reversible Markov chain. Since $L^{(\alpha)} = (D^{-\alpha} L D^{-\alpha}) \geq 0$ (as L_{ij} and D_{ii} are each nonnegative), we have $M_{ij} \geq 0$. Consider the sum of an arbitrary row of M:

$$\sum_j M_{ij} = \frac{1}{D_{ii}^{(\alpha)}} \sum_j L_{ij}^{(\alpha)} = \frac{D_{ii}^{(\alpha)}}{D_{ii}^{(\alpha)}} = 1.$$

To prove reversibility, we must show that M satisfies detailed balance. Indeed, we claim that for a stationary distribution

$$\Pi_i = \frac{D_{ii}^{(\alpha)}}{\sum_\ell D_{\ell\ell}^{(\alpha)}},$$

we have

$$\Pi_i M_{ij} = \Pi_j M_{ji}.$$

First note that L and D are symmetric. $L^{(\alpha)}$ is also symmetric, since

$$(L^{(\alpha)})^T = (D^{-\alpha} L D^{-\alpha})^T = D^{-\alpha} L^T D^{-\alpha} = D^{-\alpha} L D^{-\alpha} = L^{(\alpha)}.$$

We leverage this symmetry to verify detailed balance. Define

$$C := \sum_{\ell=1}^n D_{\ell\ell}^{(\alpha)}, \qquad \Pi_i := \frac{D_{ii}^{(\alpha)}}{C}.$$

Then

$$\Pi_i M_{ij} = \frac{D_{ii}^{(\alpha)}}{C} \frac{L_{ij}^{(\alpha)}}{D_{ii}^{(\alpha)}} = \frac{L_{ij}^{(\alpha)}}{C} = \frac{L_{ji}^{(\alpha)}}{C} = \frac{D_{jj}^{(\alpha)}}{C} \frac{L_{ji}^{(\alpha)}}{D_{jj}^{(\alpha)}} = \Pi_j M_{ji}.$$

Thus M satisfies detailed balance and is reversible.

□

2 Spectral Decomposition of the Transition Matrix

We have that
$$M_{ij}^t = \sum_\ell \lambda_\ell^t \, \psi_\ell(x_i) \, \phi_\ell(x_j),$$
where $\{\lambda_\ell\}$ is the sequence of eigenvalues of M, and $\{\psi_\ell\}$ and $\{\phi_\ell\}$ are the sequences of biorthogonal left and right eigenvectors, respectively.

Proof. Let $M \in \mathbb{R}^{n\times n}$ be the transition matrix built from the kernel, and consider the matrix representation of the stationary distribution
$$\Pi = \operatorname{diag}(\pi_1, \dots, \pi_n).$$
Reversibility implies that $\Pi M = M^\top \Pi$, which in turn implies $M^\top = \Pi M \Pi^{-1}$.

Define $A := \Pi^{1/2} M \Pi^{-1/2}$. We have that
$$A^\top = \Pi^{-1/2} M^\top \Pi^{1/2} = \Pi^{-1/2}(\Pi M \Pi^{-1})\Pi^{1/2} = \Pi^{1/2} M \Pi^{-1/2} = A,$$
so A is symmetric.

By the spectral theorem, A is diagonalizable with an orthonormal basis of eigenvectors. Indeed, since $A \in \mathbb{R}^{n\times n}$ is real symmetric, there exists an orthogonal matrix U such that
$$U^\top A U = \Lambda,$$
where Λ is diagonal with entries given by the eigenvalues of A. Denote the orthonormal basis of eigenvectors of A by $\{u_\ell\}$ and the associated set of eigenvalues by $\{\lambda_\ell\}$.

Then we will have that $Au_\ell = \lambda_\ell u_\ell$ and $u_\ell^\top u_m = \delta_{\ell m}$. Define
$$\psi_\ell := \Pi^{-1/2} u_\ell, \qquad \phi_\ell := \Pi^{1/2} u_\ell.$$
Then,
$$M\psi_\ell = \Pi^{-1/2} A \Pi^{1/2}(\Pi^{-1/2} u_\ell) = \Pi^{-1/2} A u_\ell = \lambda_\ell \Pi^{-1/2} u_\ell = \lambda_\ell \psi_\ell.$$
Similarly,
$$\phi_\ell^\top M = u_\ell^\top \Pi^{1/2} M = u_\ell^\top (\Pi^{1/2} M \Pi^{-1/2})\Pi^{1/2} = (Au_\ell)^\top \Pi^{1/2} = \lambda_\ell u_\ell^\top \Pi^{1/2} = \lambda_\ell \phi_\ell^\top.$$
Hence $\{\psi_\ell\}$ and $\{\phi_\ell\}$ form the sets of right and left eigenvectors of M, respectively, and they are biorthogonal, since
$$\phi_\ell^\top \psi_m = u_\ell^\top u_m = \delta_{\ell m}.$$

Let Ψ and Φ denote the matrices whose columns are ψ_ℓ and ϕ_ℓ, respectively, and let Λ be the diagonal matrix of eigenvalues of M. Then,
$$M\Psi = \Psi\Lambda, \qquad \Phi^\top M = \Lambda \Phi^\top, \qquad \Phi^\top \Psi = I.$$

It follows that

$$M = \Psi\Lambda\Phi^\top, \quad M^t = \Psi\Lambda^t\Phi^\top.$$

We may then express the matrix elements as

$$M^t_{ij} = e_i^\top M^t e_j = e_i^\top \Psi\Lambda^t\Phi^\top e_j = \sum_{\ell=1}^{n} \lambda_\ell^t \psi_\ell(i)\phi_\ell(j),$$

as was to be shown.

□

3 The Diffusion Operator

Following Coifman and Lafon (2006), the one–step transition kernel $p : X \times X \to \mathbb{R}_{\geq 0}$ induces a diffusion operator $P : L^2(X,\mu) \to L^2(X,\mu)$ defined by

Definition 3.1 (Diffusion operator)**.** For $f \in L^2(X,\mu)$,

$$(Pf)(x) := \int_X p(x,y)\, f(y)\, d\mu(y).$$

Coifman–Lafon write the averaging kernel as $a(x,y)$; here we take $a(x,y) = p(x,y)$. The normalization $\int_X p(x,y)\, d\mu(y) = 1$ implies $P\mathbf{1} = \mathbf{1}$ and positivity preservation.

The integral operator P is the continuous analogue of the discrete diffusion matrix M. Indeed, in the discrete case,

$$(Pf)(x_i) = \sum_j M_{ij} f(x_j).$$

4 Continuum Limit and Kernel Asymptotics

We keep the Gaussian bandwidth ε in $k_\varepsilon(x,y) = \exp(-\|x-y\|^2/\varepsilon)$. Assume the data are sampled i.i.d. from a density ρ on $\mathbb{R}^d$ (with respect to Lebesgue measure dy). Define the following in this continuum setting:

$$d_\varepsilon(x) := \int_{\mathbb{R}^d} k_\varepsilon(x,y)\, \rho(y)\, dy.$$

As $\varepsilon \to 0$,

$$d_\varepsilon(x) = (\pi\varepsilon)^{d/2}\, \rho(x) + \frac{(\pi\varepsilon)^{d/2}}{4}\, \varepsilon\, \Delta\rho(x) + O\big(\varepsilon^{d/2+3/2}\big),$$

in particular $d_\varepsilon(x) \sim C\, \varepsilon^{d/2}\, \rho(x)$ with $C = \pi^{d/2}$.

Proof. We defined

$$d_\varepsilon(x) = \int_{\mathbb{R}^d} k_\varepsilon(x,y)\,\rho(y)\,dy, \qquad k_\varepsilon(x,y) = \exp\!\left(-\frac{\|x-y\|^2}{\varepsilon}\right).$$

Let $y = x + \sqrt{\varepsilon}\,z$, so that $dy = \varepsilon^{d/2}\,dz$. Then

$$d_\varepsilon(x) = \varepsilon^{d/2} \int_{\mathbb{R}^d} e^{-\|z\|^2}\,\rho(x + \sqrt{\varepsilon}\,z)\,dz.$$

Expanding $\rho(x + \sqrt{\varepsilon}\,z)$ via a multivariate Taylor expansion,

$$\rho(x + \sqrt{\varepsilon}\,z) = \rho(x) + \sqrt{\varepsilon}\,\nabla\rho(x)\cdot z + \frac{\varepsilon}{2}\,z^\top H_\rho(x)\,z + O\big(\varepsilon^{3/2}\|z\|^3\big),$$

where $H_\rho(x)$ denotes the Hessian of ρ. Substituting and using the symmetry of $e^{-\|z\|^2}$ to eliminate odd terms yields

$$d_\varepsilon(x) = \pi^{d/2}\,\varepsilon^{d/2}\,\rho(x) + \frac{\pi^{d/2}}{4}\,\varepsilon^{d/2+1}\,\Delta\rho(x) + O\big(\varepsilon^{d/2+3/2}\big).$$

Thus $d_\varepsilon(x) \sim C_\varepsilon\,\rho(x)$ with $C_\varepsilon = \pi^{d/2}\,\varepsilon^{d/2}$. ☐

Definition 4.1. Define the diffusion generator

$$(\mathcal{L}_\varepsilon^{(\alpha)} f) := \frac{(M^{(\alpha)} f - f)}{C\,\varepsilon},$$

where $M^{(\alpha)}$ is the α–normalized transition matrix, ε is the Gaussian bandwidth, and C is a scalar constant independent of ε.

Proposition 4.2. We will show that as the data becomes dense and $\varepsilon \to 0$ in the continuum, $\mathcal{L}_\varepsilon^{(\alpha)} f \to \Delta f$.

Proof. Recall that $M^{(\alpha)} = (D^{(\alpha)})^{-1} L^{(\alpha)}$, so that

$$(M^{(\alpha)} f)(x_i) = \frac{1}{D_{ii}^{(\alpha)}} \sum_j L_{ij}^{(\alpha)} f(x_j).$$

Hence

$$(\mathcal{L}_\varepsilon^{(\alpha)} f)(x_i) = \frac{1}{C\varepsilon} \left[\frac{1}{D_{ii}^{(\alpha)}} \sum_j L_{ij}^{(\alpha)} f(x_j) - f(x_i) \right].$$

which equals

$$(\mathcal{L}_\varepsilon^{(\alpha)} f)(x_i) = \frac{1}{C\varepsilon D_{ii}^{(\alpha)}} \left[\sum_j L_{ij}^{(\alpha)} f(x_j) - D_{ii}^{(\alpha)} f(x_i) \right].$$

Since $D_{ii}^{(\alpha)} := \sum_j L_{ij}^{(\alpha)}$, we see that the numerator equals

$$\sum_j L_{ij}^{(\alpha)}(f(x_j) - f(x_i)),$$

so we obtain

$$(\mathcal{L}_\varepsilon^{(\alpha)} f)(x_i) = \frac{1}{C\varepsilon D_{ii}^{(\alpha)}} \sum_j L_{ij}^{(\alpha)}(f(x_j) - f(x_i)).$$

We know that

$$L^{(\alpha)} = D^{-\alpha} L D^{-\alpha}, \quad \text{so } L_{ij}^{(\alpha)} = \frac{k_\varepsilon(x_i, x_j)}{D_{ii}^\alpha D_{jj}^\alpha}.$$

Thus,

$$(\mathcal{L}_\varepsilon^{(\alpha)} f)(x_i) = \left(\frac{1}{C\varepsilon}\right) \left[\frac{\sum_j \frac{k_\varepsilon(x_i,x_j)}{D_{ii}^\alpha D_{jj}^\alpha}(f(x_j) - f(x_i))}{D_{ii}^{(\alpha)}}\right].$$

which equals

$$\left(\frac{1}{C\varepsilon}\right) D_{ii}^{-\alpha} \sum_j k_\varepsilon(x_i, x_j)\, D_{jj}^{-\alpha}\big(f(x_j) - f(x_i)\big) \Big/ D_{ii}^{(\alpha)}.$$

Since

$$D_{ii}^{(\alpha)} = D_{ii}^{-\alpha} \sum_j k_\varepsilon(x_i, x_j)\, D_{jj}^{-\alpha},$$

we have that

$$(\mathcal{L}_\varepsilon^{(\alpha)} f)(x_i) = \left(\frac{1}{C\varepsilon}\right) \frac{\sum_j k_\varepsilon(x_i, x_j) D_{jj}^{-\alpha}(f(x_j) - f(x_i))}{\sum_j k_\varepsilon(x_i, x_j) D_{jj}^{-\alpha}}.$$

Using $D_{kk} \asymp C^* \varepsilon^{d/2} \rho(x_k)$ for a scalar constant C^*, we obtain the following expression for $(\mathcal{L}_\varepsilon^{(\alpha)} f)(x_i)$ in the limit:

$$(\mathcal{L}_\varepsilon^{(\alpha)} f)(x) = \left(\frac{1}{C\varepsilon}\right) \frac{\int k_\varepsilon(x, y) \rho(y)^{1-\alpha}(f(y) - f(x))\, dy}{\int k_\varepsilon(x, y) \rho(y)^{1-\alpha}\, dy}.$$

Now take $y = x + v$ for $v \in \mathbb{R}^d$ and Taylor expand each of $f(y) - f(x)$ and $\rho(y)^{1-\alpha}$ about x.

Define $g(t) = f(x + tv)$. Then,

$$g'(t) = \frac{d}{dt} f(x+tv) = \sum_{i=1}^d \frac{\partial f(x + tv)}{\partial x_i} \frac{d(x_i + tv_i)}{dt} = \sum_{i=1}^d \frac{\partial f(x + tv)}{\partial x_i} v_i = \langle \nabla f(x+tv), v \rangle,$$

where $\langle \cdot, \cdot \rangle$ denotes the dot product. Then $g'(0) = \langle \nabla f(x), v \rangle$.

Take

$$g''(t) = \sum_{i=1}^{d} \frac{d}{dt}\left(\frac{\partial f(x+tv)}{\partial x_i}\right) v_i = \sum_{i=1}^{d}\sum_{j=1}^{d} \frac{\partial^2 f(x+tv)}{\partial x_i \partial x_j} v_i v_j = v^T H_f v.$$

At $t = 0$, $g''(0) = v^T H_f(x) v$.

Now Taylor expand $g(t)$ about $t = 0$ and then let $t = 1$ to obtain

$$g(1) = f(x+v) = f(x) + \langle \nabla f(x), v\rangle + \tfrac{1}{2} v^T H_f(x) v + O(\|v\|^3).$$

Thus,

$$f(y) - f(x) = \langle \nabla f(x), v\rangle + \tfrac{1}{2} v^T H_f(x) v + O(\|v\|^3),$$

which is the desired Taylor expansion for $f(y) - f(x)$.

Now take $\rho(y)^{1-\alpha} = \rho(x+v)^{1-\alpha}$ and Taylor expand.

Define $g(t) = \rho(x+tv)$. Then,

$$g'(t) = \langle \nabla \rho(x+tv), v\rangle,$$

and at $t = 0$ we obtain

$$g'(0) = \langle \nabla \rho(x), v\rangle.$$

Thus to first order we obtain

$$\rho(x+v) = \rho(y) \approx \rho(x) + \langle \nabla \rho(x), v\rangle.$$

Use the binomial approximation $(1+u)^p \approx 1 + pu$ for small u as follows:

$$\rho(x+v)^{1-\alpha} \approx \rho(x)^{1-\alpha}\left(1 + \frac{\langle \nabla\rho(x), v\rangle}{\rho(x)}\right)^{1-\alpha} \approx \rho(x)^{1-\alpha}\left(1 + (1-\alpha)\frac{\langle \nabla\rho(x), v\rangle}{\rho(x)}\right).$$

We may now substitute these expressions into $(\mathcal{L}_\varepsilon^{(\alpha)} f)(x)$. We get

$$(\mathcal{L}_\varepsilon^{(\alpha)} f)(x) = \left(\frac{1}{C\varepsilon}\right) \frac{\int k_\varepsilon(v)\,\rho(x)^{1-\alpha}\left(1 + (1-\alpha)\frac{\langle\nabla\rho(x),v\rangle}{\rho(x)}\right)\left(\langle \nabla f(x), v\rangle + \frac{1}{2} v^T H_f(x) v + O(\|v\|^3)\right) dv}{\int k_\varepsilon(v)\,\rho(x)^{1-\alpha}\left(1 + (1-\alpha)\frac{\langle\nabla\rho(x),v\rangle}{\rho(x)}\right) dv}.$$

Note that, in keeping with our choice to let $y = x + v$, we have expressed $k_\varepsilon(x, y)$ as $k_\varepsilon(v)$. Note also that k is a symmetric function and the domain of integration is symmetric about the origin, and so we need only consider even order terms of v in the integrals in the numerator and denominator of $(\mathcal{L}_\varepsilon^{(\alpha)} f)(x)$.

Indeed, $(\mathcal{L}_\varepsilon^{(\alpha)} f)(x)$ reduces to:

$$\left(\frac{1}{dC\varepsilon}\right) \frac{\rho(x)^{1-\alpha}\left[(1-\alpha)\langle \nabla f(x), \nabla \log\rho(x)\rangle \int k_\varepsilon(v)\,\|v\|^2\,dv + \frac{1}{2}\Delta f(x)\int k_\varepsilon(v)\,\|v\|^2\,dv\right]}{\rho(x)^{1-\alpha}\int k_\varepsilon(v) dv}.$$

Let $m_0 := \int k_\varepsilon(v)\,dv$ and $m_2 := \int k_\varepsilon(v)\,\|v\|^2\,dv$, so that

$$(\mathcal{L}_\varepsilon^{(\alpha)} f)(x) = \left(\frac{1}{dC\varepsilon}\right)\frac{m_2}{m_0}\left[(1-\alpha)\langle\nabla f(x), \nabla \log \rho(x)\rangle + \tfrac{1}{2}\Delta f(x)\right].$$

For $K := \frac{m_2}{dm_0}$, we obtain

$$(\mathcal{L}_\varepsilon^{(\alpha)} f)(x) = \left(\frac{K}{C\varepsilon}\right)\left[(1-\alpha)\langle\nabla f(x), \nabla \log \rho(x)\rangle + \tfrac{1}{2}\Delta f(x)\right].$$

Thus, for $\alpha = 1$ we have

$$(\mathcal{L}_\varepsilon^{(1)} f)(x) = \frac{K}{2C\varepsilon}\Delta f(x) + \text{lower–order terms.}$$

In particular, for the Gaussian kernel k_ε we have $K/\varepsilon = 1/2$, so taking $\alpha = 1$ and $C = 1/4$ (independent of ε) makes the prefactor $K/(2C\varepsilon) = 1$. In this case, the diffusion generator converges to the Laplacian in the space $L^2(\mathbb{R}^d, \rho(x)\,dx)$.

□

5 Diffusion on Smooth Manifolds

We can now extend this convergence result to a more useful setting. We will make the *manifold hypothesis*, which means that we assume the dataset lies on a smooth manifold. In particular, we hypothesize that the data lie on a smooth, compact Riemannian manifold without boundary. Even in this case, we will see that the generator of diffusion converges to the Laplace–Beltrami operator for $\alpha = 1$ and an appropriately chosen C.

The differential–geometric preliminaries and the use of normal coordinates in this section are standard; for background and a compatible choice of conventions see, for example, Lee (2012) and Lee (2018). The expansion of the metric in Riemannian normal coordinates and the associated expression for the volume element are standard consequences of these tools and the Jacobi field equation; we record them here in a self-contained way without attempting to track a specific source.

Definition 5.1. A **smooth manifold** is a topological space M which is Hausdorff, second countable, and locally Euclidean of fixed dimension n. It is also equipped with a maximal smooth atlas of charts $\mathcal{A}$.

Definition 5.2. A **maximal smooth atlas** of charts is a set of pairs

$$\mathcal{A} = \{(U_i, \phi_i)\}_{i\in I},$$

where each U_i is an open set of the topological space M, and $\phi_i : U_i \to \phi_i(U_i) \subset \mathbb{R}^n$ is a homeomorphism onto its image.

Each pair (U_i, ϕ_i) is called a *chart*. For $\mathcal{A}$ to be a maximal smooth atlas, we also require that for any two charts (U_i, ϕ_i) and (U_j, ϕ_j) in $\mathcal{A}$ such that $U_i \cap U_j \neq \emptyset$, the transition map

$$\phi_j \circ \phi_i^{-1} : \phi_i(U_i \cap U_j) \to \phi_j(U_i \cap U_j)$$

is a C^∞ diffeomorphism. The condition thus mentioned is called *smooth compatibility*. $\mathcal{A}$ is maximal if it contains every chart that is smoothly compatible with all charts in $\mathcal{A}$.

A smooth manifold is a pair $(M, \mathcal{A})$.

Definition 5.3. When we equip $(M, \mathcal{A})$ with the Riemannian metric tensor g, we obtain a **Riemannian manifold** $(M, \mathcal{A}, g)$.

Definition 5.4. A Riemannian metric on a smooth manifold M is a symmetric, positive-definite $(0, 2)$ tensor field g such that the assignment $x \mapsto g_x$ is smooth. For each $x \in M$, g defines a symmetric bilinear form

$$g_x : T_xM \times T_xM \to \mathbb{R},$$

where T_xM denotes the tangent space to the manifold at x.

Definition 5.5. A (p, q) tensor field on M is an assignment $x \mapsto T_x$ with

$$T_x : (T_x^*M)^{\otimes p} \times (T_xM)^{\otimes q} \to \mathbb{R},$$

so that T_x has p covector slots and q vector slots. A covector is an element of the algebraic dual space T_x^*M of T_xM, known as the cotangent space.

Definition 5.6. A smooth (p, q) tensor field on M is defined as before, except that the assignment $x \mapsto T_x$ is smooth. This means that in any coordinate chart $(U, (x^1, \dots, x^n))$, the component functions

$$T^{i_1 \dots i_q}{}_{j_1 \dots j_p}(x)$$

of T with respect to the coordinate bases $\left\{\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}\right\}$ and $\{dx^1, \dots, dx^n\}$ are smooth functions on U.

For instance, when T is a vector field (a $(1, 0)$ tensor field), we may write in local coordinates

$$V(x) = V^1(x)\left(\frac{\partial}{\partial x^1}\right)_x + \cdots + V^n(x)\left(\frac{\partial}{\partial x^n}\right)_x,$$

and the smoothness of V means that each component function $V^i(x)$ is smooth on U.

Definition 5.7. Broadly speaking, a (smooth) vector field is an assignment of a tangent vector to every point on the manifold. A (smooth) vector field is a smooth section $\sigma : M \to TM$ with $\pi \circ \sigma = \mathrm{id}_M$. Here TM denotes the tangent bundle of M.

Definition 5.8. Let M be a smooth n-dimensional manifold. For each $p \in M$, consider T_pM. We define the tangent bundle as follows:

$$TM := \bigsqcup_{p \in M} T_pM.$$

The tangent bundle TM is also equipped with a projection map

$$\pi : TM \to M,$$

where $\pi(v_p) = p$, making the fiber over p equivalent to $\pi^{-1}(p) = T_pM$. TM is also equipped with a smooth structure making the projection map π into a smooth submersion. This also makes the tangent bundle TM into a smooth rank-n vector bundle over M.

Definition 5.9. A smooth submersion is a smooth map $F : M \to N$ between smooth manifolds such that

$$dF_p : T_pM \to T_{F(p)}N$$

is surjective.

Let $(U, (x^1, \dots, x^n))$ be a chart on M with coordinate map $\phi : U \to \phi(U) \subset \mathbb{R}^n$. Every $V_p \in T_pM$ has unique components V^i in the coordinate basis $\left\{\frac{\partial}{\partial x^i}\right\}_p$, so $V_p = \sum_{i=1}^n V^i \left(\frac{\partial}{\partial x^i}\right)_p$. Define the local trivialization

$$\Phi_U : \pi^{-1}(U) \longrightarrow \phi(U) \times \mathbb{R}^n, \qquad \Phi_U(V_p) := \big(\phi(p), (V^1, \dots, V^n)\big).$$

In these coordinates the bundle projection is the coordinate projection

$$(\pi)_{|\pi^{-1}(U)} = \phi^{-1} \circ \mathrm{pr}_1 \circ \Phi_U, \qquad \text{i.e.} \quad (\phi(p), v) \mapsto \phi(p).$$

If $(V, (x'^1, \dots, x'^n))$ is another chart with map ψ and $\pi^{-1}(U \cap V) \neq \varnothing$, the transition map on overlaps is

$$\Phi_U \circ \Phi_V^{-1} : \ (\psi(p), v') \longmapsto \big(\ \phi(p),\ A(\phi(p))\, v'\ \big), \qquad A(\phi(p)) = \left(\frac{\partial x^i}{\partial x'^j}\right)_{i,j}(p).$$

Thus the base coordinates transform by the smooth change of variables $\phi \circ \psi^{-1}$, and the fiber coordinates transform linearly by the Jacobian A (smooth in p). Hence all transition maps are smooth diffeomorphisms, which endows TM with a canonical smooth manifold structure of dimension $2n$.

Now, earlier we stated that TM constitutes a smooth rank-n vector bundle over M. We define a smooth rank-n vector bundle now.

Definition 5.10 Let M be a smooth manifold. A smooth rank-n vector bundle over M consists of a smooth manifold N together with a smooth surjective submersion $\pi : N \to M$ such that for every $p \in M$ there exists an open set $U \subset M$ with $p \in U$ and a diffeomorphism

$$\Phi_U : \ \pi^{-1}(U) \ \longrightarrow \ U \times \mathbb{R}^n$$

satisfying

$$\pi_U \circ \Phi_U \;=\; \pi\big|_{\pi^{-1}(U)}$$

and for each $x \in U$, the induced map on the fiber

$$\Phi_U\big|_{E_x} :\; E_x = \pi^{-1}(x) \;\longrightarrow\; \{x\} \times \mathbb{R}^n \cong \mathbb{R}^n$$

is a vector space isomorphism. Moreover, on overlaps $U \cap V \neq \varnothing$, the transition map

$$\Phi_V \circ \Phi_U^{-1} :\; (x, v) \;\longmapsto\; \big(x,\; g_{VU}(x)\, v\big)$$

is of this form with $g_{VU} : U \cap V \to GL(n, \mathbb{R})$ smooth. Here π refers to the projection map.

Definition 5.11: Consider a structure (E, B, π, F) where E, F, and B are smooth manifolds and $\pi : E \to B$ is a smooth surjection. We call E the total space, B the base space, and F the model fiber. π is the projection map. Moreover, we must have that $\forall x \in B, \exists U \subset B$ such that $x \in U$ and a local trivialization $\Phi_U : \pi^{-1}(U) \to U \times F$ is defined. We will be able to draw the following commutative diagram:

$$\begin{tikzcd}
\pi^{-1}(U) \arrow[r, "\Phi_U"] \arrow[d, "\pi"'] & U \times F \arrow[dl, "\mathrm{pr}_1"] \\
U &
\end{tikzcd}$$

The model fiber F is isomorphic to all fibers given by $\pi^{-1}(x)$ for some $x \in B$. Then we call (E, B, π, F) a smooth fiber bundle.

Definition 5.12: Let (E, B, π, F) be a smooth fiber bundle. A smooth section of this fiber bundle is a smooth map $\sigma : B \to E$ that is a right inverse of the projection $\pi : E \to B$, such that $\forall p \in B$, we have that $\pi \circ \sigma(p) = p$. In other words, $\pi \circ \sigma = \mathrm{id}_B$.

We may now formulate an alternate definition of a rank-n vector bundle in the context of smooth fiber bundles.

Definition 5.13: A rank-n vector bundle is a smooth fiber bundle $(E, B, \pi, \mathbb{R}^n)$ (here $\mathbb{R}^n$ is the model fiber F) together with the local trivializations $\Phi_U : \pi^{-1}(U) \to U \times \mathbb{R}^n$ for each open $U \subset B$ such that $\pi \circ \Phi_U = \pi|_{\pi^{-1}(U)}$, and for each $b \in U$, $\Phi_U|_{E_b} : E_b \to \{b\} \times \mathbb{R}^n \cong \mathbb{R}^n$ is a vector space isomorphism. Note: $E_b = \pi^{-1}(b)$.

Having introduced the relevant differential geometry concepts, we return to the definition of the generator of diffusion. For a diffusion process on a smooth, compact Riemannian manifold without boundary, we define

$$(\mathcal{L}_\varepsilon^{(\alpha)} f)(x) := \left(\frac{1}{C\varepsilon}\right) \frac{\displaystyle\int_M k(d(x,y))\big(f(y) - f(x)\big)\rho(y)^{1-\alpha}\, dV_g(y)}{\displaystyle\int_M k(d(x,y))\rho(y)^{1-\alpha}\, dV_g(y)}.$$

With this definition of $(\mathcal{L}_\varepsilon^{(\alpha)} f)(x)$ in the manifold setting, we have that $d(x,y)$ is the geodesic distance between points $x, y \in M$. The kernel k is defined just as before, except that we replace the Euclidean norm $\|x - y\|$ with the geodesic distance $d(x,y)$.

Definition 5.10 (Geodesic distance in normal coordinates)**.** Let (M, g) be Riemannian. If y lies in a normal neighborhood U of $x \in M$, there is a unique $v \in T_x M$ with $y = \exp_x(v)$, where $\exp_x|_V : V \subset T_x M \to U$ is a diffeomorphism. The curve $\gamma(t) = \exp_x(tv)$, $t \in [0,1]$, is the unique minimizing geodesic from x to y with initial velocity v, and

$$d(x,y) = \|v\|_{g_x}.$$

We have that

$$dV_g(y) = \sqrt{\det(g_{ij}(y))}\, dy^1 \wedge \cdots \wedge dy^n$$

for a smooth Riemannian manifold of dimension n. The expression $\det(g_{ij}(y))$ denotes the determinant of the metric tensor g at point y. We will expand g so that this operation is concrete. We also adapt the density $\rho(y)$ to be defined on M. We require that

$$\int_M \rho(y)\, dV_g(y) = 1,$$

so that ρ represents a valid probability density with respect to the differential volume measure dV_g.

Definition 5.11. For a point $x \in M$, fix an orthonormal basis $\{e_1, \ldots, e_n\}$ of $T_x M$, which is orthonormal with respect to the inner product g. Let $y = \exp_x(v)$ for $v \in T_x M$, written as

$$v = \sum_{i=1}^{n} v^i e_i.$$

Then the **Riemannian normal coordinates** of y are $(v^1, \ldots, v^n)$. In these coordinates, we have that $g_{ij}(x) = \delta_{ij}$.

Now we will show that

$$\sqrt{\det(g(v))} = (\det(g(v)))^{1/2} \approx 1 - \tfrac{1}{6}\mathrm{Ric}_{cd}v^c v^d + \mathcal{O}(\|v\|^3)$$

where Ric_{cd} denotes the Ricci curvature tensor. To do this, we will expand g using Riemannian normal coordinates to obtain

$$g_{ij}(\gamma(t)) = \delta_{ij} - \tfrac{t^2}{3}\, R_{jcdi}(x)v^c v^d + O(t^3).$$

We will derive the expression for $g_{ij}(\gamma(t))$ from first principles.

Definition 5.12. A **geodesic** is a smooth curve on a Riemannian manifold whose tangent vector is parallel transported along itself. Formally, we have that for $\gamma(t)$ a smooth curve on a Riemannian manifold M parameterized by t, the following must hold in order for $\gamma(t)$ to be a geodesic:

$$\nabla_{\dot\gamma(t)}\dot\gamma(t) = 0,$$

where ∇ denotes the Levi–Civita connection, and $\dot\gamma(t)$ is the velocity vector field of $\gamma(t)$.

Definition 5.13. Also called the **connection**, the **covariant derivative** ∇ is defined on the tangent bundle of a smooth manifold M. The connection ∇ is an operator that takes two vector fields X, Y and produces a new vector field $\nabla_X Y$. The connection ∇ satisfies the following properties for all vector fields X, Y, Z and all smooth real-valued functions f, g:

1. $\nabla_{fX+gY}Z = f\nabla_X Z + g\nabla_Y Z$,
2. $\nabla_X(Y+Z) = \nabla_X Y + \nabla_X Z$,
3. $\nabla_X(fY) = (Xf)Y + f\nabla_X Y$.

We can make the connection operation concrete by expressing it in local coordinates. Indeed, in a local coordinate chart $(U, (x^1, \ldots, x^n))$ on M, the action of the connection is determined by a set of functions called the **Christoffel symbols** Γ^k_{ij}, defined by

$$\nabla_{\partial_i}\partial_j = \Gamma^k_{ij}\,\partial_k.$$

Using the connection axioms, we can evaluate the components of $\nabla_X Y$. We have $X = X^i\partial_i$ and $Y = Y^j\partial_j$ in local coordinates. Then

$$\nabla_X Y = X^i\nabla_{\partial_i}(Y^j\partial_j).$$

Applying the product rule, we obtain

$$X^i\nabla_{\partial_i}(Y^j\partial_j) = X^i(\partial_i Y^j)\partial_j + X^i Y^j\nabla_{\partial_i}\partial_j = X^i(\partial_i Y^j)\partial_j + X^i Y^j\Gamma^k_{ij}\partial_k.$$

Now we relabel $Y^j\partial_j$ as $Y^k\partial_k$, which yields

$$X^i(\partial_i Y^k)\partial_k + X^i Y^j\Gamma^k_{ij}\partial_k.$$

Factoring ∂_k gives the k-th component of the vector field $\nabla_X Y$:

$$(\nabla_X Y)^k = X^i\partial_i Y^k + X^i Y^j\Gamma^k_{ij}.$$

Definition 5.14. The **Levi-Civita connection** is the unique connection ∇ on a Riemannian manifold (M, g) which satisfies:

1. **Torsion-free:** $T(X, Y) = 0$, where
$$T(X, Y) := \nabla_X Y - \nabla_Y X - [X, Y].$$
Here $[X, Y]$ denotes the Lie bracket of vector fields, defined by
$$[X, Y](f) = X(Y(f)) - Y(X(f)) \quad \text{for } f \in C^\infty(M).$$
In local coordinates,
$$[X, Y] = \sum_i \left(X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j} \right) \partial_i.$$

2. **Metric compatibility:** $\nabla g = 0$, i.e.
$$(\nabla_Z g)(X, Y) := Z(g(X, Y)) - g(\nabla_Z X, Y) - g(X, \nabla_Z Y) = 0.$$

It follows from the fundamental theorem of Riemannian geometry that there exists one and only one connection that is both torsion-free and metric-compatible. This is the Levi-Civita connection.

Definition 5.15. Let $\gamma(t) = \exp_x(tv)$ be a geodesic on a Riemannian manifold (M, g). A **geodesic variation of** $\gamma(t)$ is a smooth map $f(s, t)$ where, for fixed s, the curve $t \mapsto f(s, t)$ is a geodesic, and $f(0, t) = \gamma(t)$.

Definition 5.16. We define the **Jacobi field** as
$$J(t) = \frac{\partial f}{\partial s}\bigg|_{s=0}.$$
The Jacobi field describes the infinitesimal separation between the geodesics along $f(s, t)$ for each fixed s. The Jacobi field is a vector field along the geodesic $\gamma(t) = \exp_x(tv)$. A vector field $X(t)$ along $\gamma(t) = \exp_x(tv)$ is a Jacobi field if and only if it satisfies the **Jacobi field equation**:
$$\nabla_{\dot\gamma} \nabla_{\dot\gamma} X(t) + R(X(t), \dot\gamma(t))\, \dot\gamma(t) = 0.$$

Definition 5.17. The **Riemannian curvature tensor** R is a type $(1, 3)$ tensor where, for vector fields X, Y, Z, we have
$$R(X, Y)Z := \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z.$$

We may view R as a type $(1, 3)$ or a type $(0, 4)$ tensor field. To show why it makes sense to view R as either type of tensor field, we introduce the concept of *musical isomorphisms.*

Fix $x \in M$. The curvature endomorphism R at x is a type $(1, 3)$ tensor,
$$R: \; T_xM \times T_xM \times T_xM \longrightarrow T_xM, \qquad (X, Y, Z) \mapsto R(X, Y)Z.$$

Now we may use $g : T_xM \times T_xM \to \mathbb{R}$ to obtain an equivalent characterization of R as a type $(0,4)$ tensor field. Indeed, we may let

$$R(X,Y,Z,W) := g(R(X,Y)Z,W)$$

for vectors $X,Y,Z,W \in T_xM$.

Definition 5.18. We define the **sharp map** $\sharp : T_x^*M \to T_xM$ which takes some $\omega \in T_x^*M$ and outputs a vector $\omega^\sharp$ defined to be the unique vector $\omega^\sharp \in T_xM$ such that

$$g(\omega^\sharp, V) = \omega(V)$$

for all $V \in T_xM$. The map $\sharp : T_x^*M \to T_xM$ is known as the **sharp isomorphism**.

Definition 5.19. Define the **flat map** $\flat : T_xM \to T_x^*M$, which takes a $V \in T_xM$ (such as $V = R(X,Y)Z$) and allows us to consider the covector $V^\flat$ induced by g, where

$$V^\flat(W) := g(V,W) \in T_x^*M.$$

Note that in this framework, $V \in T_xM$ behaves as if it were a covector in the algebraic dual space acting on any $W \in T_xM$. The map $\flat : T_xM \to T_x^*M$ is known as the **flat isomorphism**. When we let $R(X,Y,Z,W) := g(R(X,Y)Z,W)$, the metric g induces the covector $V^\flat \in T_x^*M$ from the vector $V = R(X,Y)Z \in T_xM$. Using the flat isomorphism lets us view R as a type $(0,4)$ tensor field.

The *flat* and *sharp* musical isomorphisms allow us to freely associate between the two characterizations of R via metric equivalence.

We evaluate the metric in Riemannian normal coordinates based at $x \in M$. Choose an orthonormal basis $\{e_i\}$ of T_xM and write $v = \sum_{i=1}^n v^i e_i \in T_xM$, so $y = \exp_x(v)$ has normal coordinates $(v^1,\ldots,v^n)$. The exponential map is smooth, so its differential at v,

$$(d\exp_x)_v : \ T_v(T_xM) \longrightarrow T_yM,$$

is well defined. Using the canonical identification $T_v(T_xM) \cong T_xM$, the coordinate vectors $\{\partial_{v^i}\}$ at v correspond to $\{e_i\}$ at x, and

$$g_{ij}(y) \ = \ g_y\big((d\exp_x)_v(\partial_{v^i}), \ (d\exp_x)_v(\partial_{v^j})\big).$$

Now, let us consider the family of Jacobi fields $\{J_i(t)\}_i$ along the geodesic $\gamma(t) = \exp_x(tv)$, defined with initial conditions $J_i(0) = 0$ and $\nabla_t J_i(0) = e_i$. By solving the Jacobi field equation with these initial conditions, we may find the Taylor expansion for $J_i(t)$.

Definition 5.20. We define a **frame** along $\gamma(t)$ to be a smoothly varying choice of basis for the tangent space to the manifold at each point along the path $\gamma(t)$. Formally, we fix a frame

$$\{E_1(t),\ldots,E_n(t)\}, \qquad E_i(t) \in T_{\gamma(t)}M \ \forall t.$$

We also require that $\{E_i(t)\}$ be an orthonormal basis so that

$$g(E_i(t), E_j(t)) = \delta_{ij}.$$

Additionally, we require that this orthonormal frame satisfy **parallel transport**, meaning

$$\nabla_{\dot{\gamma}} E_i(t) = 0 \qquad \forall i.$$

When this condition is satisfied, we say that the frame is *parallel transported* along $\gamma(t)$. We have thus defined the parallel orthonormal frame that we will use to express $J_i(t)$, and by extension $g_{ij}(\gamma(t))$.

Now we may let

$$J_i(t) = {A_i}^a(t) E_a(t),$$

as $\{E_a(t)\}$ is an orthonormal basis of the tangent space along $\gamma(t)$. Each ${A_i}^a(t)$ is a smooth coefficient function of t.

Because $\nabla_{\dot{\gamma}} E_a(t) = 0$, we have that

$$\nabla_{\dot{\gamma}} J_i(t) = (\dot{A_i}^a(t)) E_a(t),$$

and hence

$$\nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} J_i(t) = (\ddot{A_i}^a(t)) E_a(t).$$

Now recall that R is multilinear, and so when we express $R(X,Y)Z$ with respect to the frame, we may write

$$R\left(\sum_b \phi^b E_b,\ \sum_c \beta^c E_c\right)\left(\sum_d \psi^d E_d\right) = \phi^b \beta^c \psi^d\, R(E_b, E_c) E_d.$$

Proof. The Riemannian curvature tensor is indeed multilinear. Consider $R(X,Y)(fZ)$. We have

$$\nabla_X \nabla_Y (fZ) = \nabla_X \big(Y(f)Z + f\nabla_Y Z\big) = X(Y(f))Z + Y(f)\nabla_X Z + X(f)\nabla_Y Z + f\nabla_X \nabla_Y Z.$$

Likewise,

$$\nabla_Y \nabla_X (fZ) = Y(X(f))Z + X(f)\nabla_Y Z + Y(f)\nabla_X Z + f\nabla_Y \nabla_X Z.$$

Also,

$$\nabla_{[X,Y]}(fZ) = (X(Y(f)) - Y(X(f)))Z + f\nabla_{[X,Y]} Z.$$

Hence

$$R(X,Y)(fZ) = f(\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z) = fR(X,Y)Z,$$

proving linearity in Z.

To show linearity in X and Y, we use the property of the Lie bracket

$$[fX, Y] = f[X,Y] - (Yf)X, \qquad [X, fY] = f[X,Y] + (Xf)Y.$$

Then, for $R(fX,Y)Z$, we compute

$$R(fX,Y)Z = \nabla_{fX}\nabla_Y Z - \nabla_Y\nabla_{fX}Z - \nabla_{[fX,Y]}Z = f\nabla_X\nabla_Y Z - \nabla_Y(f\nabla_X Z) - \nabla_{f[X,Y]-(Yf)X}Z,$$

which expands to

$$f\nabla_X\nabla_Y Z - f\nabla_Y\nabla_X Z - f\nabla_{[X,Y]}Z = fR(X,Y)Z.$$

Likewise, by a similar calculation,

$$R(X,fY)Z = fR(X,Y)Z.$$

Thus R is multilinear (i.e., tensorial) in all its arguments. $\square$

We return now to our expansion of $J_i(t)$, where we had already set

$$J_i(t) = A_i{}^b(t)E_b(t).$$

We now work on expanding $A_i{}^b(t)$ by re-expressing the Jacobi field equation:

$$\nabla_{\dot\gamma}\nabla_{\dot\gamma}J_i(t) + R(J_i(t),\dot\gamma(t))\,\dot\gamma(t) = 0.$$

We have seen that

$$\nabla_{\dot\gamma}\nabla_{\dot\gamma}J_i(t) = \ddot A_i{}^a(t)E_a(t),$$

and we consider

$$R(J_i(t),\dot\gamma(t))\,\dot\gamma(t).$$

Let $\dot\gamma(t) = v^c E_c(t)$, so that

$$R(J_i(t),\dot\gamma(t))\,\dot\gamma(t) = A_i{}^b(t)v^c v^d\,R(E_b,E_c)E_d.$$

We let $R(E_b,E_c)E_d = R^a{}_{bcd}E_a$ when expressed with respect to the orthonormal basis $\{E_a(t)\}$.

Thus

$$R(J_i(t),\dot\gamma(t))\,\dot\gamma(t) = A_i{}^b(t)v^c v^d R^a{}_{bcd}(x)E_a.$$

Hence the Jacobi field equation becomes

$$\ddot A_i{}^a(t) + A_i{}^b(t)\,v^c v^d R^a{}_{bcd}(x) = 0.$$

We have that $J_i(0)=0$ and $\nabla_{\dot\gamma}J_i(0) = e_i$, which implies

$$A_i{}^a(0) = 0, \qquad \dot A_i{}^a(0) = \delta_i{}^a.$$

Now consider the Taylor expansion of $A_i{}^a(t)$ about $t=0$:

$$A_i{}^a(t) = A_i{}^a(0) + \dot A_i{}^a(0)t + \tfrac{1}{2}\ddot A_i{}^a(0)t^2 + \tfrac{1}{6}A_i{}^{a(3)}(0)t^3 + O(t^4).$$

We have that

$$\ddot A_i{}^a(0) = -R^a{}_{bcd}(x)\,v^c v^d A_i{}^b(0),$$

and differentiating again gives

$$A_i{}^{a(3)}(0) = -R^a{}_{bcd}(x)\, v^c v^d\, \dot{A}_i{}^b(0).$$

Then we have

$$A_i{}^{a(3)}(0) = -R^a{}_{icd}(x) v^c v^d.$$

Notice the shift in index: since $\dot{A}_i{}^b(0) = \delta_i{}^b$, the contraction $R^a{}_{bcd}(x) v^c v^d \dot{A}_i{}^b(0)$ is nonzero only when $b = i$. Thus, we may shift the index b to i without consequence—this operation is known as **tensor contraction**. In particular, the index b is contracted with the upper index of $\dot{A}_i{}^b$.

Therefore,

$$A_i{}^a(t) = \delta_i{}^a t - \tfrac{1}{6} R^a{}_{icd}(x)\, v^c v^d t^3 + O(t^4).$$

Thus,

$$J_i(t) = A_i{}^a(t) E_a(t) = \left(\delta_i{}^a t - \tfrac{1}{6} R^a{}_{icd}(x)\, v^c v^d t^3 + O(t^4)\right) E_a(t).$$

This is the desired Taylor expansion of $J_i(t)$.

Note that $E_a(t) = e_a(t) + O(t^2)$, so we let $E_a(t) = e_a(t)$ when approximating to first order.

Now we fix $v \in T_x M$ and set $\gamma(t) = \exp_x(tv)$, and consider $F(s,t) = \exp_x(t(v + se_i))$. For each s, the map $t \mapsto F(s,t)$ is a geodesic with initial velocity $v + se_i$. Recall that

$$J_i(t) = \frac{\partial F}{\partial s}\Big|_{s=0} = d(\exp_x)_{tv}(te_i),$$

which implies that

$$\frac{1}{t} J_i(t) = d(\exp_x)_{tv}(e_i).$$

Thus,

$$g_{ij}(\gamma(t)) = g\big(d(\exp_x)_{tv}(e_i), d(\exp_x)_{tv}(e_j)\big) = \frac{1}{t^2} g(J_i(t), J_j(t)).$$

Then,

$$\frac{1}{t^2} g(J_i(t), J_j(t)) = \frac{1}{t^2} g\left(\left(e_i t - \tfrac{t^3}{6} R^a{}_{icd}(x) v^c v^d e_a + O(t^4)\right), \left(e_j t - \tfrac{t^3}{6} R^a{}_{jcd}(x) v^c v^d e_a + O(t^4)\right)\right).$$

Expanding g, which is bilinear, yields:

$$g_{ij}(\gamma(t)) = \delta_{ij} - \tfrac{t^2}{3}\, g(e_i, R^a{}_{jcd}(x) v^c v^d e_a) + O(t^3).$$

Recall the definition of R as a type $(0,4)$ tensor field, where

$$R(X,Y,Z,W) := g(R(X,Y)Z, W).$$

We have the following symmetries:

$$R(X,Y,Z,W) = -R(X,Y,W,Z), \quad R(X,Y,Z,W) = -R(Y,X,Z,W), \quad R(X,Y,Z,W) = R(Z,W,X,Y).$$

Proof. We will prove that $R(X,Y,Z,W) = -R(X,Y,W,Z)$. We first note that the covariant derivative acts on general tensor fields. For T a type (r,s) tensor field and X a vector field,

$$\begin{aligned}(\nabla_X T)(V_1,\ldots,V_r,\alpha_1,\ldots,\alpha_s) = {} & X(T(V_1,\ldots,V_r,\alpha_1,\ldots,\alpha_s)) \\ & - \sum_{i=1}^{r} T(V_1,\ldots,\nabla_X V_i,\ldots,V_r,\alpha_1,\ldots,\alpha_s) \\ & - \sum_{j=1}^{s} T(V_1,\ldots,V_r,\alpha_1,\ldots,\nabla_X \alpha_j,\ldots,\alpha_s).\end{aligned}$$

For a type $(0,2)$ tensor field, the metric compatibility condition of the Levi-Civita connection gives

$$(\nabla_X g)(Y,Z) = X(g(Y,Z)) - g(\nabla_X Y, Z) - g(Y, \nabla_X Z) = 0.$$

Hence

$$X(g(Y,Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z).$$

Now define

$$S := X(Y(g(Z,W))) - Y(X(g(Z,W))) - [X,Y](g(Z,W)).$$

Expanding and using the above identity yields

$$S = g(R(X,Y)Z,W) + g(R(X,Y)W,Z).$$

Thus,

$$R(X,Y,Z,W) + R(X,Y,W,Z) = 0,$$

which proves that

$$R(X,Y,Z,W) = -R(X,Y,W,Z),$$

as desired. □

6 Expansion of the Metric and Volume Element

Now, in $g_{ij}(\gamma(t)) = \delta_{ij} - \frac{t^2}{6}(g(e_i, R^a{}_{jcd}(x)v^c v^d) + g(e_j, R^a{}_{icd}(x)v^c v^d)) + \mathcal{O}(t^3)$, we have

$$g(e_i, R^a{}_{jcd}(x)v^c v^d) + g(e_j, R^a{}_{icd}(x)v^c v^d) = 2g(e_i, R^a{}_{jcd}(x)v^c v^d),$$

so that

$$g_{ij}(\gamma(t)) = \delta_{ij} - \tfrac{t^2}{3} g(e_i, R^a{}_{jcd}(x)v^c v^d) + \mathcal{O}(t^3).$$

Proof. Let $X := e_i$, $Y := e_j$, and write $V = v^c e_c \in T_x M$. Then

$$g(e_i,\, R^a{}_{jcd}(x)\, v^c v^d e_a) = g(R(Y,V)V,\, X), \qquad g(e_j,\, R^a{}_{icd}(x)\, v^c v^d e_a) = g(R(X,V)V,\, Y).$$

By the standard symmetries of R,

$$R(X, V, V, Y) = -R(V, X, V, Y) = R(Y, V, V, X),$$

justifying the previous step. We obtain

$$g_{ij}(v) = \delta_{ij} - \frac{1}{3} R_{jcdi}(x)\, v^c v^d + \mathcal{O}(\|v\|^3).$$

Now express $g_{ij}(v)$ in matrix form to obtain

$$g(v) = I + B$$

for $B_{ij} = -\frac{1}{3} R_{jcdi}(x) v^c v^d$.
For small B we have that $\det(I + B) \approx 1 + \mathrm{tr}(B)$, and

$$\mathrm{tr}(B) = \delta^{ij} B_{ij} = -\tfrac{1}{3}\, \delta^{ij} R_{jcdi}(x) v^c v^d.$$

$\mathrm{Ric}_{cd} := \delta^{ij} R_{jcdi}$ and is known as the Ricci tensor. In an orthonormal frame we may define it equivalently as $\mathrm{Ric}_{cd} := g^{ij} R_{jcdi}$. It is a contraction of the Riemannian curvature tensor in the i and j indices.
Remark: For $\mathrm{Ric}_{cd} : T_x M \times T_x M \to \mathbb{R}$, we may fix $u \in T_x M$ and define

$$\alpha_u : v \mapsto \mathrm{Ric}(u, v), \quad \alpha_u \in T_x^* M.$$

Then the sharp isomorphism enables

$$g_x(\mathrm{Ric}_x^{\#}(u), v) = \mathrm{Ric}_x(u, v) \quad \forall v \in T_x M,$$

since $\alpha_u := \mathrm{Ric}(u, \cdot) \in T_x^* M$ constitutes a linear map. Then the map

$$\mathrm{Ric}_x^{\#} : T_x M \to T_x M$$

is known as the Ricci endomorphism.
Thus,

$$\det(g(v)) \approx 1 - \tfrac{1}{3} \mathrm{Ric}_{cd} v^c v^d + \mathcal{O}(\|v\|^3)$$

Then

$$\sqrt{\det(g(v))} = (\det(g(v)))^{1/2} \approx 1 - \tfrac{1}{6} \mathrm{Ric}_{cd} v^c v^d + \mathcal{O}(\|v\|^3)$$

Now recall the definition of $\mathcal{L}_\varepsilon^\alpha f(x)$:

$$\mathcal{L}_\varepsilon^\alpha f(x) = \left(\frac{1}{C\varepsilon}\right) \frac{\displaystyle\int_M k(d(x, y))(f(y) - f(x))\rho(y)^{1-\alpha}\, dV_g(y)}{\displaystyle\int_M k(d(x, y))\rho(y)^{1-\alpha}\, dV_g(y)}.$$

In a normal neighborhood of x, write $y = \exp_x(v)$ with $v \in T_x M \cong \mathbb{R}^n$. Consider the function

$$g(t) := f(\exp_x(tv)).$$

Then $g(0) = f(x)$ and a Taylor expansion at $t = 0$ gives

$$g(1) - g(0) = \langle \nabla_g f(x), v \rangle + \tfrac{1}{2} v^T H_f(x) v + \mathcal{O}(\|v\|^3),$$

so

$$f(y) - f(x) = \langle \nabla_g f(x), v \rangle + \tfrac{1}{2} v^T H_f(x) v + \mathcal{O}(\|v\|^3).$$

and

$$\rho(y)^{1-\alpha} \approx \rho(x)^{1-\alpha} \left[1 + (1-\alpha) \frac{\langle \nabla_g \rho(x), v \rangle}{\rho(x)} \right] + \mathcal{O}(\|v\|^2).$$

Thus in $\mathcal{L}_\varepsilon^\alpha f(x)$ the numerator $N(x)$ becomes:

$$\begin{aligned} N(x) = \int_{\mathbb{R}^n} & k(\|v\|)\, \rho(x)^{1-\alpha} \left(1 + (1-\alpha) \frac{\langle \nabla_g \rho(x), v \rangle}{\rho(x)} \right) \\ & \times \left[\langle \nabla_g f(x), v \rangle + \tfrac{1}{2} v^T H_f(x) v + \mathcal{O}(\|v\|^3) \right] \left[1 - \tfrac{1}{6} \mathrm{Ric}_{cd}(x) v^c v^d \right] dv. \end{aligned}$$

and the denominator $D(x)$ becomes:

$$D(x) = \int_{\mathbb{R}^n} k(\|v\|) \rho(x)^{1-\alpha} \left(1 + (1-\alpha) \frac{\langle \nabla_g \rho(x), v \rangle}{\rho(x)} \right) \left[1 - \tfrac{1}{6} \mathrm{Ric}_{cd}(x) v^c v^d \right] dv.$$

The domain of integration is Euclidean, and so we need only consider even moments of v.
Indeed, $N(x)$ reduces to

$$N(x) = \frac{m_2}{n}\, \rho(x)^{1-\alpha} \left[\tfrac{1}{2} \Delta_g f(x) + (1-\alpha) \langle \nabla_g f(x), \nabla_g \log \rho(x) \rangle \right],$$

for $m_2 := \int_{\mathbb{R}^n} k(\|v\|)\, \|v\|^2\, dv$, and $D(x)$ reduces to

$$D(x) = \rho(x)^{1-\alpha} \int_{\mathbb{R}^n} k(\|v\|) dv = m_0\, \rho(x)^{1-\alpha}, \quad \text{for } m_0 := \int_{\mathbb{R}^n} k(\|v\|) dv.$$

Thus, $\mathcal{L}_\varepsilon^\alpha f(x)$ becomes

$$\mathcal{L}_\varepsilon^\alpha f(x) = \left(\frac{K}{C\varepsilon} \right) \left[\tfrac{1}{2} \Delta_g f(x) + (1-\alpha) \langle \nabla_g f(x), \nabla_g \log \rho(x) \rangle \right],$$

for $K := m_2/nm_0$. Since the Gaussian moments are computed in $T_xM \cong \mathbb{R}^n$, we have $K/\varepsilon = 1/2$ where $n = \dim M$. Choosing $\alpha = 1$ and $C = 1/4$ (independent of ε), we obtain

$$\mathcal{L}_\varepsilon^1 f \xrightarrow[\varepsilon \to 0]{L^2(M, \rho\, dV_g)} \Delta_g f,$$

in the sense that

$$\left\| \mathcal{L}_\varepsilon^1 f - \Delta_g f \right\|_{L^2(M, \rho\, dV_g)} \longrightarrow 0 \quad \text{for every } f \in C^\infty(M),$$

so the diffusion generator converges strongly to the Laplace–Beltrami operator on the manifold.

References

Coifman, Ronald R. and Stéphane Lafon (2006). "Diffusion Maps". In: *Applied and Computational Harmonic Analysis* 21.1. Special Issue: Diffusion Maps and Wavelets, pp. 5–30. DOI: 10.1016/j.acha.2006.04.006. URL: https://www.sciencedirect.com/science/article/pii/S1063520306000546.

Diffusion map (n.d.). https://en.wikipedia.org/wiki/Diffusion_map. Accessed: 2025-11-08.

Lee, John M. (2012). *Introduction to Smooth Manifolds*. 2nd ed. Graduate Texts in Mathematics 218. New York: Springer.

— (2018). *Introduction to Riemannian Manifolds*. 2nd ed. Graduate Texts in Mathematics 176. Cham: Springer.