Applied Mathematics, 2026, 17(8), 572-583
https://www.scirp.org/journal/am
ISSN Online: 2152-7393
ISSN Print: 2152-7385

A Unified Geometric and Energetic Framework for Deep Neural Networks via RKHS Embeddings

Heriony Rapelanoro-Rabenja

Department of Mathematics, Applied Mathematics Laboratory, University of Fianarantsoa, Fianarantsoa, Madagascar
Email: h.rapelanorohery@gmail.com

How to cite this paper: Rapelanoro-Rabenja, H. (2026) A Unified Geometric and Energetic Framework for Deep Neural Networks via RKHS Embeddings. *Applied Mathematics*, **17**, 572-583.
https://doi.org/10.4236/am.2026.178031

Received: May 26, 2026
Accepted: August 25, 2026
Published: August 28, 2026

Abstract

This article develops a unified geometric and energetic framework for the analysis of deep neural networks, based on embedding the output manifold into a Reproducing Kernel Hilbert Space (RKHS). This embedding induces a natural Riemannian metric, a Levi-Civita connection, a second fundamental form, and a mean curvature vector, allowing the construction of a complete geometric energy model. We show how these tools lead to intrinsic learning dynamics, coherent geometric regularization, and physically interpretable energy flows. Experiments demonstrate improvements in stability, robustness, and generalization.

Keywords

RKHS Geometry, Riemannian Learning, Extrinsic Curvature, Mean Curvature, Geometric Regularization, Energy-Based Models

1. General Introduction

The goal of this article is to develop a unified geometric and energetic framework for the analysis of deep neural networks. The central idea is to view the output manifold of the network as an immersed submanifold of a Reproducing Kernel Hilbert Space (RKHS). This perspective naturally introduces:

- an induced Riemannian metric,
- a Levi-Civita connection,
- a second fundamental form and mean curvature,
- a complete geometric energy model,
- intrinsic learning dynamics consistent with the geometry.

This framework connects differential geometry, stochastic diffusions, energy models, and deep learning. Learning becomes a geometric flow on an immersed manifold, where geometric regularization plays a central role in stability, robustness, and generalization.

This geometric viewpoint is closely related to information geometry [1] and recent advances in geometric deep learning [2]. Foundational references on differential geometry include do Carmo [3], Lee [4], and Petersen [5]. Kernel methods and RKHS theory follow Schölkopf and Smola [6] and Wahba [7], while mathematical perspectives on deep architectures are discussed in Calin [8] and Mallat [9]. Related energetic and stochastic complexity aspects are developed in [10].

2. Output Manifold and RKHS Embedding

2.1. Network Output and Parametric Manifold

Consider a deep neural network parameterized by $\theta \in \mathbb{R}^m$ and a mapping

$$F_\theta : X \to Y.$$

For input points $\{x_i\}_{i=1}^n$, the output is

$$y(\theta) = \left(F_\theta(x_1), \cdots, F_\theta(x_n)\right) \in \mathbb{R}^n.$$

As θ varies, the set of outputs forms a smooth manifold:

$$\mathcal{S} = \left\{ y(\theta) \in \mathbb{R}^n ; \theta \in \mathbb{R}^m \right\}.$$

2.2. Embedding into a RKHS

To introduce a rich geometric structure, we embed the output into a RKHS $\left(\mathcal{H}_K, \langle \cdot, \cdot \rangle_K\right)$ associated with a positive kernel K:

$$\Phi : X \to \mathcal{H}_K, \ \Phi(x) = K(x, \cdot).$$

The output manifold becomes an immersed submanifold:

$$\mathcal{S}_K = \left\{ \Phi\left(y(\theta)\right) ; \theta \in \mathbb{R}^m \right\} \subset \mathcal{H}_K.$$

This embedding provides a natural metric induced by the RKHS inner product, a smooth Riemannian structure, access to geometric tools (connection, curvature), and a physically interpretable energy.

2.3. Induced Metric

The induced metric on $\mathcal{S}_K$ is

$$h_\theta(u, v) = \left\langle \frac{\partial \Phi\left(y(\theta)\right)}{\partial \theta}(u), \frac{\partial \Phi\left(y(\theta)\right)}{\partial \theta}(v) \right\rangle_K .$$

The tangent space is

$$T_\theta \mathcal{S}_K = \left\{ \frac{\partial \Phi\left(y(\theta)\right)}{\partial \theta}(v) ; v \in \mathbb{R}^m \right\}.$$

This structure enables intrinsic gradient flows, geometric energies, and curvature-based regularization.

3. Induced Metric and Riemannian Structure

3.1. Geometric Motivation

Embedding the output manifold $\mathcal{S}_K$ into a RKHS naturally endows it with a Riemannian structure. This allows the definition of:

- covariant derivatives,
- geodesics,
- intrinsic and extrinsic curvature,
- intrinsic gradient flows.

These tools are essential for formulating a coherent energetic theory in which learning is interpreted as motion on a manifold equipped with an induced metric.

3.2. Illustrative Example: One-Hidden-Neuron Network

Consider the simple network

$$F_\theta(x) = \sigma(wx + b), \ \ \theta = (w, b).$$

For input points $\{x_i\}$, the output is

$$y(\theta) = (\sigma(wx_1 + b), \cdots, \sigma(wx_n + b)).$$

As (w, b) vary, the points $y(\theta)$ trace a smooth curve in $\mathbb{R}^n$. Embedding this curve into a RKHS yields $\Phi(y(\theta)) \in \mathcal{H}_K$, and the induced metric becomes

$$h_\theta(u, v) = \left\langle \frac{\partial \Phi(y(\theta))}{\partial \theta}(u), \frac{\partial \Phi(y(\theta))}{\partial \theta}(v) \right\rangle_K .$$

Even this simple network exhibits nontrivial intrinsic geometry.

4. Levi-Civita Connection and Intrinsic Dynamics

4.1. Levi-Civita Connection: Formal Definition

On any Riemannian manifold $(\mathcal{S}_K, h)$, there exists a unique connection ∇ satisfying:

- Metric compatibility:

$$X(h(U, V)) = h(\nabla_X U, V) + h(U, \nabla_X V),$$

- Zero torsion:

$$\nabla_U V - \nabla_V U = [U, V].$$

Theorem 4.1 (Levi-Civita Connection). There exists a unique connection satisfying the two properties above. It is called the Levi-Civita connection.

4.2. Physical Meaning of the Levi-Civita Connection

The Levi-Civita connection can be interpreted as a transport rule that is energy-preserving (metric compatibility) and twist-free (zero torsion). It corresponds to sliding along the manifold without artificial rotation.

4.3. Riemannian Gradient and Learning Dynamics

The Euclidean gradient $\nabla_{\mathcal{H}_K} E$ is not tangent to $\mathcal{S}_K$. The Riemannian gradient is its tangential projection:

$$\operatorname{grad} E(\theta) = \Pi_{T_\theta \mathcal{S}_K}\left(\nabla_{\mathcal{H}_K} E\right).$$

The learning dynamics becomes a geometric flow:

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = -\operatorname{grad} E(\theta).$$

4.4. Link with Backpropagation

Let $f_\theta(x) \in \mathcal{H}_K$ denote the RKHS representation of the network output. The RKHS gradient of the energy E can be written as

$$\nabla_{\mathcal{H}_K} E = 2 f_\theta + \sum_{i=1}^{n} \nabla_f \ell\left(f_\theta(x_i), z_i\right) K(\cdot, x_i),$$

where ℓ is the loss function. The parameter gradient is obtained by the chain rule:

$$\nabla_\theta E = \left\langle \nabla_{\mathcal{H}_K} E, \frac{\partial f_\theta}{\partial \theta} \right\rangle_{\mathcal{H}_K}.$$

Thus, the usual backpropagation update

$$\theta_{t+1} = \theta_t - \eta \nabla_\theta E$$

can be interpreted as the projection of the ambient RKHS gradient onto the parameter manifold. This provides a geometric bridge between the abstract RKHS gradient and the concrete parameter updates implemented in practice.

5. Second Fundamental Form and Extrinsic Curvature

5.1. Gauss Decomposition

For any point $p \in \mathcal{S}_K$, we have the orthogonal decomposition:

$$T_p \mathcal{H}_K = T_p \mathcal{S}_K \oplus \left(T_p \mathcal{S}_K\right)^\perp.$$

For tangent vector fields X, Y:

$$\nabla_X Y = \left(\nabla_X Y\right)^\top + \left(\nabla_X Y\right)^\perp.$$

5.2. Second Fundamental Form: Formal Definition

Definition 5.1. The second fundamental form is defined by:

$$\mathcal{L}(X, Y) = \left(\nabla_X Y\right)^\perp.$$

It measures how the submanifold bends within the ambient space.

5.3. Geometric Role of the Second Fundamental Form

- If $\mathcal{L} = 0$, the manifold is flat in the ambient space.

- If $\mathcal{L}$ is large, the manifold is highly curved.
- In deep networks, $\mathcal{L}$ measures extrinsic sensitivity.

5.4. Illustration of Extrinsic Curvature

The Gauss decomposition of $\nabla_X Y$ into tangential and normal components is illustrated in **Figure 1**, showing how the ambient direction splits into intrinsic and extrinsic parts.

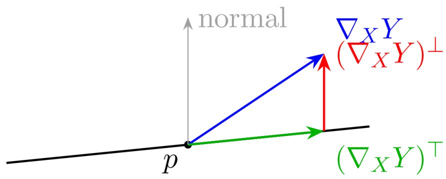

Figure 1. Gauss decomposition in the RKHS: decomposition of $\nabla_X Y$ into tangential and normal components.

5.5. Geometric Energy

We define the extrinsic geometric energy:

$$E_{\text{geom}} = \int_{\mathcal{S}_K} \|\mathcal{L}\|^2 \,\mathrm{d}\mu.$$

It penalizes excessive extrinsic deformation and encourages smoother output manifolds.

5.6. Example in $\mathbb{R}^3$

Consider the surface:

$$S = \left\{ (x, y, z) \in \mathbb{R}^3; z = x^2 + y^2 \right\}.$$

At the point $p = (0,0,0)$, the surface looks like a shallow bowl. The tangent plane is horizontal, the normal vector is $n = (0,0,1)$, and the mean curvature vector points upward.

The paraboloid example and its upward normal at the origin are shown in **Figure 2**, highlighting how curvature encodes geometric deformation.

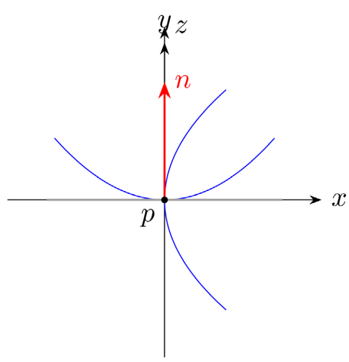

Figure 2. Illustration of the paraboloid $z = x^2 + y^2$ and its upward normal at the origin.

This example illustrates how curvature encodes geometric deformation in a familiar setting and why controlling extrinsic curvature stabilizes neural representations.

6. Mean Curvature and Geometric Regularization

6.1. Mean Curvature: Formal Definition

Let $\{T_1,\cdots,T_m\}$ be an orthonormal basis of $T_p\mathcal{S}_K$. The mean curvature vector is defined as:

$$H(p)=\sum_{i=1}^{m}\mathcal{L}(T_i,T_i),$$

where $\mathcal{L}$ is the second fundamental form. Thus,

$$H(p)\in(T_p\mathcal{S}_K)^{\perp}.$$

6.2. Geometric Role of Mean Curvature

The mean curvature measures the tendency of the manifold to bend within the ambient space:

- $H=0$: the manifold is minimal,
- large H: the manifold is highly curved,
- H indicates the direction of maximal bending.

6.3. Illustration of Mean Curvature

Figure 3 depicts the mean curvature as the sum of normal components $\mathcal{L}(T_i,T_i)$, providing an intuitive geometric interpretation.

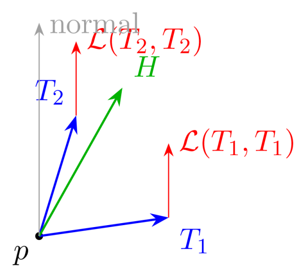

Figure 3. Mean curvature as the sum of normal components $\mathcal{L}(T_i,T_i)$.

6.4. Mean Curvature Energy Functional

We define:

$$E_{\mathrm{MC}}=\int_{\mathcal{S}_K}\|H\|^2\,\mathrm{d}\mu.$$

This acts as a membrane tension: minimizing $\|H\|^2$ smooths the manifold.

6.5. Physical Meaning of Mean Curvature

- $H=0$: membrane in equilibrium,

- large H : membrane bent or stretched,
- minimizing $\|H\|^2$: smoother, more stable representations.

In deep networks, H measures global geometric complexity and the tendency to overfit via highly curved decision boundaries.

7. Geometric Generator and Diffusion Operator

7.1. Stochastic Dynamics on the Manifold

Consider a diffusion on $\mathcal{S}_K$:

$$\mathrm{d}\theta_t = -\mathrm{grad}_h E(\theta_t)\mathrm{d}t + \sigma(\theta_t)\mathrm{d}W_t,$$

where W_t is a Brownian motion and $\sigma(\theta_t)$ is a diffusion coefficient.

7.2. Geometric Generator

The associated generator is:

$$\mathcal{L}f(\theta) = -\left(\mathrm{grad}_h E(\theta), \mathrm{grad}_h f(\theta)\right)_h + \frac{1}{2}\mathrm{Tr}\left(\sigma(\theta)\sigma(\theta)^\top \mathrm{Hess}_h f(\theta)\right).$$

7.3. Interpretation of the Geometric Generator

The first term encodes energy dissipation along the Riemannian gradient, while the second term encodes intrinsic diffusion on the manifold. Mini-batch noise induces a diffusion on the curved manifold $\mathcal{S}_K$, and curvature modulates the amplitude of stochastic fluctuations. Regions of high extrinsic curvature amplify noise, while flatter regions stabilize the diffusion. Thus, geometric regularization shapes both the deterministic and stochastic components of learning.

8. Feynman-Kac Representation

8.1. Geometric PDE

Consider the PDE:

$$\frac{\partial u(\theta,t)}{\partial t} = \mathcal{L}u(\theta,t) - V(\theta)u(\theta,t),\ u(\theta,0) = f(\theta),$$

where $\mathcal{L}$ is the geometric generator and V is a potential.

8.2. Feynman-Kac Formula

The solution is:

$$u(\theta,t) = \mathbb{E}\left[f(\theta_t)\exp\left(-\int_0^t V(\theta_s)\mathrm{d}s\right) \mid \theta_0 = \theta\right].$$

8.3. Interpretation of the Feynman-Kac Representation

The Feynman-Kac representation provides a probabilistic interpretation of the geometric PDE governing the evolution of $u(\theta,t)$. Learning becomes a diffusion process evolving on the curved manifold $\mathcal{S}_K$, where the potential $V(\theta)$ acts as an energetic penalty. The exponential term

$$\exp\left(-\int_0^t V(\theta_s)\,\mathrm{d}s\right)$$

plays the role of a Boltzmann weight: trajectories passing through regions of high curvature or high energy are exponentially suppressed, while smoother, low-energy trajectories contribute more significantly.

9. Complete Geometric Energy Model

9.1. General Principle

The geometric structure induced by the RKHS embedding provides a natural way to define a total energy functional that governs learning. This energy combines three components: a data fitting term, an extrinsic curvature term, and a mean curvature term.

9.2. Data Fitting Energy

$$E_{\text{data}}(\theta) = \sum_{i=1}^{n} \left\| F_\theta(x_i) - z_i \right\|^2 .$$

9.3. Extrinsic Geometric Energy

$$E_{\text{geom}} = \int_{\mathcal{S}_K} \|\mathcal{L}\|^2 \,\mathrm{d}\mu.$$

9.4. Mean Curvature Contribution to Total Energy

$$E_{\text{MC}} = \int_{\mathcal{S}_K} \|H\|^2 \,\mathrm{d}\mu.$$

9.5. Total Energy

$$E(\theta) = E_{\text{data}}(\theta) + \lambda_1 E_{\text{geom}}(\theta) + \lambda_2 E_{\text{MC}}(\theta), \quad \lambda_1, \lambda_2 \geq 0.$$

9.6. Riemannian Gradient

$$\operatorname{grad}_h E(\theta) = \Pi_{T_\theta \mathcal{S}_K}\left(\nabla_{\mathcal{H}_K} E\right), \quad \frac{\mathrm{d}\theta}{\mathrm{d}t} = -\operatorname{grad}_h E(\theta).$$

9.7. Physical Meaning of the Total Energy Model

Each term has a clear physical meaning:

- E_{data} pulls the model toward the data,
- E_{geom} penalizes extrinsic curvature,
- E_{MC} acts as a membrane tension,
- the Riemannian gradient ensures intrinsic consistency.

10. Applications and Interpretations

The geometric and energetic framework has several practical implications:

- Representation stability: penalizing $\|\mathcal{L}\|^2$ and $\|H\|^2$ stabilizes the geometry of the output manifold.

- Noise robustness: a less curved manifold is less sensitive to input perturbations.
- Generalization: geometric regularization controls extrinsic complexity.
- Interpretability: curvature provides a geometric measure of sensitivity.

The interplay between $\|\mathcal{L}\|^2$ and $\|H\|^2$ is particularly informative: $\mathcal{L}$ captures local bending and sensitivity, while H reflects global geometric tension. Together, they provide a multi-scale description of the complexity of the learned representation.

11. Experiments and Validation

This section presents several experiments illustrating the benefits of the geometric and energetic framework developed above.

11.1. Experimental Setup

We consider three families of datasets:

- Synthetic datasets: two-moons, concentric circles, intertwined spirals.
- Image datasets: MNIST and Fashion-MNIST (reduced resolution).
- Tabular datasets: standard UCI-type classification tasks.

We evaluate several architectures:

- MLP-3: fully connected network with three hidden layers.
- CNN-Small: small convolutional network.
- ResNet-18 (reduced): lightweight residual network.

Training schemes:

- Baseline (ERM + cross-entropy),
- L^2 regularization,
- Spectral norm regularization,
- G-RKHS (extrinsic curvature),
- G-RKHS+MC (extrinsic + mean curvature).

Hyperparameters (λ_1, λ_2) are selected by cross-validation. All experiments are repeated over 5 random seeds.

11.2. Curvature Measurement

We approximate:

$$\|\mathcal{L}\|^2 \approx \mathbb{E}\left[\left\|\mathcal{L}\left(T_i, T_j\right)\right\|^2\right],\ \|H\|^2 \approx \|H(p)\|^2,$$

using finite differences and local tangent bases.

11.3. Experiment 1: Synthetic Classification

Baseline models produce highly curved decision boundaries. Geometric regularization smooths these boundaries and improves accuracy.

The impact of geometric regularization on synthetic datasets is summarized in Table 1, which reports test accuracy together with extrinsic and mean curvature values.

Table 1. Synthetic datasets: test accuracy and curvature (mean over 5 runs).

Method	Accuracy (%)	$\|\mathcal{L}\|^2$	$\|H\|^2$
Baseline	96.2 ± 0.4	12.8	7.4
L^2	96.5 ± 0.3	10.1	6.2
Spectral Norm	97.1 ± 0.5	8.4	5.7
G-RKHS	98.4 ± 0.2	4.1	2.3
G-RKHS + MC	98.7 ± 0.1	2.8	1.4

11.4. Experiment 2: Noise Robustness

Gaussian perturbations $x \mapsto x + \varepsilon$, $\varepsilon \sim \mathcal{N}\left(0, \sigma^2 I\right)$.

Robustness to Gaussian noise is presented in Table 2, showing that curvature-based regularization significantly improves performance under increasing perturbation levels.

Table 2. Robustness under gaussian noise.

Method	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Baseline	89.4	72.1	51.3
L^2	90.2	74.8	54.0
Spectral Norm	92.5	78.3	58.1
G-RKHS	95.8	85.4	71.2
G-RKHS + MC	96.3	87.1	74.5

11.5. Experiment 3: Generalization on Real Data

Generalization performance on MNIST, Fashion-MNIST, and tabular datasets is reported in Table 3, where G-RKHS and G-RKHS + MC consistently outperform baseline and classical regularization schemes.

Table 3. Generalization performance (test accuracy in %).

Method	MNIST	Fashion-MNIST	Tabular
Baseline	98.1	89.4	84.2
L^2	98.3	90.1	85.0
Spectral Norm	98.4	90.7	85.6
G-RKHS	98.7	91.8	87.1
G-RKHS + MC	98.8	92.3	87.9

11.6. Energy Analysis

Geometric regularization yields smoother, more stable energy descent and correlates with improved robustness and generalization.

The evolution of the total energy during training for Baseline, G-RKHS, and G-

RKHS + MC is shown in **Figure 4**, illustrating the stabilizing effect of geometric regularization.

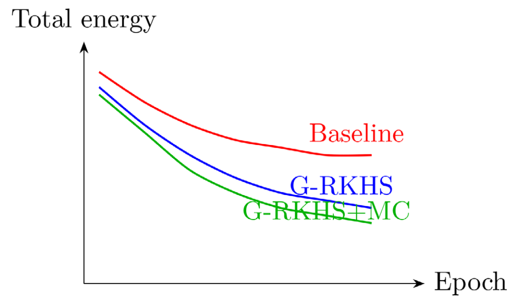

Figure 4. Illustrative evolution of the total energy across training epochs for Baseline, G-RKHS, and G-RKHS + MC.

12. Discussion and Perspectives

12.1. Summary

The RKHS-geometry-energy framework coherently connects:

- extrinsic differential geometry,
- stochastic diffusions,
- energy-based modeling,
- deep learning.

Learning becomes a geometric flow on an immersed manifold, driven by a total energy balancing data fitting and geometric regularity.

12.2. Perspectives

Future directions include:

1) Extension to deep CNNs, transformers, and recurrent networks.
2) Intrinsic optimization based on geodesics.
3) Geometric analysis of gradient dynamics.
4) Adaptive curvature-based regularization.
5) Adversarial robustness via curvature control.
6) Geometric interpretability in sensitive domains.

13. Conclusions

We introduced a complete geometric and energetic framework for analyzing deep neural networks, based on embedding the output manifold into a RKHS. This induces a natural Riemannian structure, a Levi-Civita connection, a second fundamental form, and a mean curvature vector. We showed how the RKHS gradient induces an intrinsic Riemannian gradient, and how backpropagation corresponds to its projection onto the parameter manifold. The geometric generator and Feynman-Kac representation provide a stochastic and energetic interpretation of learning.

Experiments confirm that geometric regularization improves stability, robustness, and generalization. Controlling extrinsic curvature leads to smoother decision boundaries and more stable representations. The framework opens promising directions for intrinsic optimization, adaptive geometric regularization, and applications requiring stability and interpretability.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Amari, S. (2016) Information Geometry and Its Applications. Springer. https://doi.org/10.1007/978-4-431-55978-8

[2] Bronstein, M., Bruna, J., Cohen, T. and Veličković, P. (2021) Geometric Deep Learning. arXiv: 2104.13478.

[3] do Carmo, M. (1992) Riemannian Geometry. Birkhäuser. https://doi.org/10.1007/978-1-4757-2201-7

[4] Lee, J.M. (2012) Introduction to Smooth Manifolds. Springer. https://doi.org/10.1007/978-1-4419-9982-5

[5] Petersen, P. (2016) Riemannian Geometry. Springer. https://doi.org/10.1007/978-3-319-26654-1

[6] Schölkopf, B. and Smola, A. (2002) Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press. https://doi.org/10.7551/mitpress/4175.001.0001

[7] Wahba, G. (1990) Spline Models for Observational Data. SIAM. https://doi.org/10.1137/1.9781611970128

[8] Calin, O. (2020) Deep Learning Architectures: A Mathematical Approach. Springer. https://doi.org/10.1007/978-3-030-36721-3

[9] Mallat, S. (2016) Understanding Deep Convolutional Networks. *Philosophical Transactions of the Royal Society A*, **374**, Article ID: 20150203. https://doi.org/10.1098/rsta.2015.0203

[10] Rapelanoro-Rabenja, H. (2025) Stochastic Complexity: Long Memory, Non-Gaussian Chaos, and Energy Dynamics in Hilbert Space. HAL Preprint, hal-05522560v2. https://ehess.hal.science/hal-05522560v2