

Learning Capacity of Totally Geodesic Maps into the Circle

Alper Ferudun*

25 October 2026

Abstract

We study the squared geodesic loss class of smooth totally geodesic maps from a connected Riemannian manifold into a circle. Its pseudodimension is finite precisely when the domain has no nonzero parallel one-form with integral periods. In that case the maps are constant and the pseudodimension is two; otherwise the loss class has infinite fat-shattering dimension at a fixed positive margin. For every closed domain admitting a nonconstant map, a Dirichlet-energy cutoff gives matching logarithmic orders for fat-shattering dimension and pseudodimension. The upper bound counts integral parallel forms and phase intervals, while the lower bound explicitly encodes labels in base four. We also characterize the closure of the interpolation values of affine circle maps and give a sharp two-point obstruction to an unrestricted interpolation assertion in the motivating report. The infinite capacity of the unrestricted circle-to-circle family was already known. The results here concern circle targets and do not classify totally geodesic maps between arbitrary pairs of manifolds.

1 The capacity question and the loss class

Hein, reporting joint work with Steinke and Schölkopf [1, pp. 1688–1689], asked which pairs of Riemannian manifolds have a finite-capacity family of totally geodesic maps. The motivation is regression with squared geodesic loss: a regularizer based on the covariant Hessian vanishes on every such map. The subsequent paper [2] develops this regression framework. The question is recorded as OWR-2041-002 in UnsolvedMath 1.6.0 [3].

The circle-to-circle family already illustrates infinite capacity in the source report. We retain that attribution. Our purpose is to give an exact geometric criterion for *circle targets*, to quantify the effect of an energy cutoff on every closed domain, and to distinguish shattering from arbitrary interpolation at a prescribed input set. These are results for a specified loss class, rather than a claim that every possible notion of statistical capacity agrees.

Throughout, M is a nonempty connected smooth Riemannian manifold without boundary. The target is $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, with circumference one and distance

$$d_{\mathbb{T}}(s, t) = \min_{j \in \mathbb{Z}} |s - t - j|.$$

A smooth map $F : M \rightarrow \mathbb{T}$ is totally geodesic when $\nabla dF = 0$, using the induced connection. Write $\mathcal{G}(M, \mathbb{T})$ for these maps and set

$$\mathcal{L}(M, \mathbb{T}) = \{\ell_F : (p, y) \mapsto d_{\mathbb{T}}(F(p), y)^2 : F \in \mathcal{G}(M, \mathbb{T})\}.$$

Thus $0 \leq \ell_F \leq 1/4$.

We use the strict-inequality definition of pseudodimension. Points $z_1, \dots, z_m$ are pseudoshattered by a real function class if there are thresholds t_i such that every binary vector is

*Mercury Software GmbH. Email: alper@mercurycodelab.com. GitHub: [AlperTheKing](https://github.com/AlperTheKing).

realized by strict inequalities above or below those thresholds. This is equivalent to the usual convention with a weak inequality on one side: choose one witness per labeling and slightly decrease each threshold, preserving the finitely many strict lower inequalities. For $\gamma > 0$, replace the strict inequalities by $f(z_i) \geq t_i + \gamma$ and $f(z_i) \leq t_i - \gamma$ to define γ -fat shattering. In particular $\text{fat}_\gamma \leq \text{Pdim}$.

Let

$$\Lambda(M) = \left\{ \alpha \in \Omega^1(M) : \nabla \alpha = 0, \int_c \alpha \in \mathbb{Z} \text{ for every closed piecewise smooth curve } c \right\}.$$

This is an additive group of actual parallel forms. We do not assume that the whole space of parallel forms is spanned by integral-period forms.

Theorem 1 (Circle-target dichotomy). *If $\Lambda(M) = \{0\}$, all maps in $\mathcal{G}(M, \mathbb{T})$ are constant and*

$$\text{Pdim } \mathcal{L}(M, \mathbb{T}) = 2.$$

If $\Lambda(M) \neq \{0\}$, then

$$\text{fat}_{1/24} \mathcal{L}(M, \mathbb{T}) = \infty.$$

Consequently, finite pseudodimension, finite fat-shattering dimension at every positive scale, and $\Lambda(M) = \{0\}$ are equivalent for this loss class.

2 Parallel forms and a constructive shattering argument

Lemma 2. *Every totally geodesic $F : M \rightarrow \mathbb{T}$ determines $\alpha = F^*dt \in \Lambda(M)$. Conversely, each $\alpha \in \Lambda(M)$ determines such a map up to an additive constant in $\mathbb{T}$. A nonconstant map is a submersion.*

Proof. In a local real lift f of F , the form is df . The target connection is flat, so $\nabla dF = 0$ is precisely $\nabla \alpha = 0$. Lifting the image of any closed curve shows that its period is an integer.

Conversely, a parallel one-form is closed. Fix $p_0 \in M$ and $b \in \mathbb{T}$, and define

$$F(p) = b + \int_{p_0}^p \alpha \pmod{\mathbb{Z}}.$$

Integral periods give path independence modulo $\mathbb{Z}$. Local primitives show smoothness, $F^*dt = \alpha$, and total geodesicity. Two maps with this derivative differ by a constant because M is connected. The norm of a parallel form is constant. If it is nonzero, the derivative of F has rank one everywhere. $\square$

Lemma 3 (Phase-pattern bound). *Suppose a family of maps $X \rightarrow \mathbb{T}$ is contained in the union of q translation families*

$$\{F_j + b : b \in \mathbb{T}\}, \quad 1 \leq j \leq q.$$

If $m \geq 1$ sample pairs in $X \times \mathbb{T}$ are pseudoshattered by their squared losses, then

$$2^m \leq 2mq.$$

In particular the pseudodimension is at most $2 \log_2 q + 4$.

Proof. Choose strict witnesses for all labelings. Each threshold must be strictly between 0 and $1/4$. For fixed j , the equation

$$d_{\mathbb{T}}(F_j(x_i) + b, y_i)^2 = t_i$$

has exactly two solutions in b . There are at most $2m$ boundary points on the phase circle, and at most $2m$ open components in their complement. A strict labeling is constant on each such component and none of its witnesses lies on a boundary. Summing over j gives the claimed count. If $m \geq 4$, then $\log_2 m \leq m/2$; taking logarithms gives $m \leq 2\log_2 q + 2$. The stated bound also covers $m < 4$. $\square$

Lemma 4 (Digit witnesses). *For every $m \geq 1$ and $b_1, \dots, b_m \in \{0, 1\}$, set*

$$k = \sum_{j=1}^m 2b_j 4^{j-1}, \quad x_i = 4^{-i}.$$

Then $0 \leq k < 4^m$, and

$$\begin{cases} d_{\mathbb{T}}(kx_i, 0)^2 \leq 1/36, & b_i = 0, \\ d_{\mathbb{T}}(kx_i, 0)^2 \geq 1/9, & b_i = 1. \end{cases}$$

These values realize the labels with threshold $5/72$ and margin $1/24$.

Proof. Modulo one,

$$k4^{-i} = \frac{b_i}{2} + \sum_{j < i} 2b_j 4^{j-1-i}.$$

The nonnegative tail is smaller than $\sum_{s \geq 1} 2 \cdot 4^{-s-1} = 1/6$. For a zero leading digit, the circle distance to zero is at most $1/6$; for a one leading digit, it is at least $1/2 - 1/6 = 1/3$. Squaring gives the two estimates. Finally,

$$\frac{5}{72} - \frac{1}{24} = \frac{1}{36}, \quad \frac{5}{72} + \frac{1}{24} = \frac{1}{9}.$$

$\square$

Proof of Theorem 1. If $\Lambda(M) = 0$, Lemma 2 gives only constant maps. Lemma 3 with $q = 1$ gives $2^m \leq 2m$, which fails for $m \geq 3$. For the matching lower bound take the two sample pairs $(p, 0)$ and $(p, 1/4)$, thresholds $1/36$, and constant outputs $1/8, 0, 1/4, 5/8$. Their losses have the four strict threshold patterns.

Otherwise let F be nonconstant. Since it is a submersion, a local real lift of its image contains an interval $(t_0, t_0 + \eta)$. Fix an integer $L \geq 0$ with $4^{-L-1} < \eta$. For each m select distinct points p_i satisfying

$$F(p_i) = t_0 + 4^{-L-i} \pmod{\mathbb{Z}}.$$

For any labeling, use the integer k from Lemma 4 and the totally geodesic map

$$F_b = 4^L k F - 4^L k t_0.$$

At p_i its value is $k4^{-i}$ modulo one. Take all target coordinates to be zero. The digit lemma gives arbitrary $1/24$ -fat-shattered sample sizes. $\square$

3 Energy cutoffs on closed domains

Here M is closed, meaning compact without boundary, and connected. Use the Dirichlet energy normalization

$$E(F) = \frac{1}{2} \int_M |dF|^2 dV.$$

Let $\mathcal{L}_E(M, \mathbb{T})$ denote the losses of maps with $E(F) \leq E$.

Theorem 5 (Logarithmic energy capacity). *If $\Lambda(M) = 0$, then $\text{Pdim } \mathcal{L}_E(M, \mathbb{T}) = 2$ for every $E \geq 0$. Otherwise $\Lambda(M)$ is a lattice in its real span, of some rank $r \geq 1$. Let*

$$\lambda = \min_{\alpha \in \Lambda(M) \setminus \{0\}} \|\alpha\|_{L^2}, \quad E_0 = \lambda^2/2.$$

Then for every $E \geq E_0$,

$$\left\lfloor \log_{16} \frac{E}{E_0} \right\rfloor \leq \text{fat}_{1/24} \mathcal{L}_E(M, \mathbb{T}) \leq \text{Pdim } \mathcal{L}_E(M, \mathbb{T}) \leq 2r \log_2 \left(1 + 2\sqrt{\frac{E}{E_0}} \right) + 4. \quad (1)$$

Thus both dimensions have order $\log E$ as $E \rightarrow \infty$, with M fixed and admitting a nonconstant map.

Proof. The vector space V of parallel one-forms is finite dimensional: evaluation at one point is injective by parallel transport. Integration over a basis of the free part of $H_1(M; \mathbb{Z})$ defines a linear period map from V into a finite-dimensional real vector space. Its kernel is zero. Indeed, vanishing periods make $\alpha = df$ globally; at a maximum of f , $df = 0$, so parallelness and connectedness imply $\alpha = 0$ everywhere. The preimage of the integer lattice under this injective linear map is discrete. It is precisely $\Lambda(M)$. A discrete subgroup of a finite-dimensional real vector space is a lattice in its span, so the shortest nonzero length λ exists and is positive.

By Lemma 2, maps of energy at most E form one phase circle for each

$$\alpha \in \Lambda(M), \quad \|\alpha\|_{L^2} \leq \sqrt{2E}.$$

Write $Q(E)$ for the number of these forms. In the r -dimensional span of the lattice, open balls of radius $\lambda/2$ about distinct lattice points are disjoint. Those centered in the radius- $\sqrt{2E}$ ball lie in the radius- $\sqrt{2E} + \lambda/2$ ball. Comparing volumes gives

$$Q(E) \leq \left(1 + \frac{2\sqrt{2E}}{\lambda} \right)^r = \left(1 + 2\sqrt{E/E_0} \right)^r.$$

The upper bound follows from Lemma 3.

For the lower bound, choose a map F corresponding to a shortest nonzero form, so $E(F) = E_0$. It is a submersion. Its image is open, and compactness makes that image closed; therefore $F(M) = \mathbb{T}$. Choose p_i with $F(p_i) = 4^{-i}$. For every binary labeling the digit construction uses the map kF , whose energy is $k^2 E_0 < 16^m E_0$. It is admissible whenever $16^m E_0 \leq E$. This proves the lower bound in (1). The zero-lattice case follows from the constant-map part of Theorem 1. $\square$

Remark 6. *The lattice rank need not be replaced by the dimension of the whole parallel-form space. The proof counts only integral-period forms. Closed domains with $H^1(M; \mathbb{R}) = 0$ give the constant case. Compactness is essential to that shortcut: on $M = \mathbb{R}$, the nonzero parallel form dx has integral periods and gives nonconstant circle maps despite vanishing first cohomology.*

For the unit circle as domain, all totally geodesic maps are

$$F_{k,b}(x) = kx + b, \quad k \in \mathbb{Z}, \quad b \in \mathbb{T}.$$

Indeed a lift to the real line has zero second derivative, and its unit-period compatibility forces an integer slope. In this case $E(F_{k,b}) = k^2/2$. If $\mathcal{L}_K$ is the loss class with $|k| \leq K$, the proofs give the more direct bounds

$$\lfloor \log_4 K \rfloor \leq \text{fat}_{1/24} \mathcal{L}_K \leq \text{Pdim } \mathcal{L}_K \leq 2 \log_2(2K + 1) + 4, \quad K \geq 1. \quad (2)$$

A hard energy constraint therefore restores finite capacity. By contrast, the covariant-Hessian penalty vanishes on every $F_{k,b}$; it imposes no winding bound. These are statements about the constrained function class, not convergence guarantees for a particular training algorithm.

4 Interpolation at prescribed inputs

Infinite shattering capacity is an existence statement about input configurations. It does not imply arbitrary accurate interpolation at *every* fixed input configuration.

Proposition 7 (Exact closure of the evaluation set). *Fix $x = (x_1, \dots, x_m) \in \mathbb{T}^m$. The closure of the vectors*

$$(kx_1 + b, \dots, kx_m + b), \quad k \in \mathbb{Z}, \quad b \in \mathbb{T},$$

is

$$H_x = \left\{ y \in \mathbb{T}^m : n \cdot y = 0 \text{ in } \mathbb{T} \text{ whenever } n \in \mathbb{Z}^m, \sum_i n_i = 0, \quad n \cdot x = 0 \text{ in } \mathbb{T} \right\}. \quad (3)$$

The same closure is obtained with positive integers k .

Proof. Let H be the closure of the subgroup generated by x and the diagonal circle. The character $\chi_n(y) = \exp(2\pi i n \cdot y)$ is trivial on H exactly when $n \cdot x = 0$ in $\mathbb{T}$ and $\sum_i n_i = 0$. A closed subgroup of a torus is the common kernel of its annihilating characters. For completeness, if $y \notin H$, separate the disjoint compact cosets H and $y + H$ by a continuous function. Approximate that function uniformly by a trigonometric polynomial, and average over H with its normalized Haar measure. A nontrivial character on H averages to zero, by invariance. If every character trivial on H also took value one at y , the two averages would be equal, contradicting the chosen separation. This proves the character-kernel assertion and hence the description (3).

Finally there are positive integers $q_j \rightarrow \infty$ with $q_j x \rightarrow 0$ in $\mathbb{T}^m$. One may obtain them by applying compactness to increasingly fine finite covers and taking differences of return times; if a bounded return time recurs, the cyclic subgroup is finite and its multiples suffice. For each fixed integer a , eventually $q_j + a > 0$, and $(q_j + a)x \rightarrow ax$. Thus positive powers have the same closure as all integer powers. $\square$

Equivalently, the target vector y can be fitted with arbitrarily small maximum geodesic error exactly when it satisfies the integer relations in (3). In particular, if $1, x_2 - x_1, \dots, x_m - x_1$, using real representatives, are linearly independent over $\mathbb{Q}$, then there are no nonzero relations and every target vector can be approximated. This hypothesis holds for Lebesgue-almost every input vector; it is not automatic for prescribed data.

Proposition 8 (Sharp two-point obstruction). *For inputs $x_1 = 0, x_2 = 1/2$ and targets $y_1 = 0, y_2 = 1/4$,*

$$\inf_{k \in \mathbb{Z}} \max_{b \in \mathbb{T}} \max_{i=1,2} d_{\mathbb{T}}(kx_i + b, y_i) = \frac{1}{8}.$$

The infimum is also attained with positive k .

Proof. The difference between the two outputs is 0 or $1/2$; its circle distance from the target difference $1/4$ is therefore $1/4$. If e_i are the two pointwise errors, translation invariance and the triangle inequality give $1/4 \leq e_1 + e_2 \leq 2 \max(e_1, e_2)$. Equality is achieved by $k = 2, b = 1/8$. $\square$

Theorem 1 on page 1689 of the report [1] asserts arbitrarily accurate affine-circle interpolation for unrestricted training data. Proposition 8 supplies distinct inputs contradicting that quantifier even when the phase is freely chosen. In its circumference- 2π convention the sharp lower bound is $\pi/4$. The almost-everywhere input condition above repairs the interpolation statement. The original infinite-capacity conclusion remains valid, as the explicit shattering proof shows.

5 Verification and limits

The proofs are analytic and exact. A separate rational-arithmetic script checks all 90 114 digit-margin inequalities for all binary labelings of lengths one through twelve. It also checks the two winding parities and equality witnesses in the interpolation obstruction, and the four strict patterns for constant-map losses. These finite checks are supporting diagnostics; they do not replace the proofs for arbitrary sample size or certify formal verification.

The main classification is complete for circle targets and squared geodesic loss. The energy bounds apply to closed connected domains. The arbitrary-target question in OWR-2041-002 remains unresolved here. The underlying integration, character, lattice-packing, and threshold-counting arguments are standard tools; no priority is claimed for those tools or for the known circle infinite-capacity example. A bounded source search did not identify the precise combined criterion and energy bounds stated here, but is not a certificate of absolute priority.

AI assistance was used in deriving arguments, searching sources, checking exact arithmetic, and preparing the manuscript. The author is responsible for the claims. This preprint has not undergone independent peer review.

References

- [1] M. Hein, joint work with F. Steinke and B. Schölkopf, Nonparametric Regression between Manifolds, in *Learning Theory and Approximation*, Oberwolfach Reports **5** (2008), no. 3, 1655–1706, contribution pp. 1688–1689. [doi:10.4171/OWR/2008/30](https://doi.org/10.4171/OWR/2008/30).
- [2] F. Steinke, M. Hein, and B. Schölkopf, Nonparametric Regression between General Riemannian Manifolds, *SIAM Journal on Imaging Sciences* **3** (2010), no. 3, 527–563. [doi:10.1137/080744189](https://doi.org/10.1137/080744189).
- [3] Ulam AI, *UnsolvedMath*, frozen version 1.6.0, record OWR-2041-002, record ID 30000994. <https://huggingface.co/datasets/ulamai/UnsolvedMath>. Snapshot commit c6e7f41b4eb6ce876feb83e7eb626be1b243fd7e.