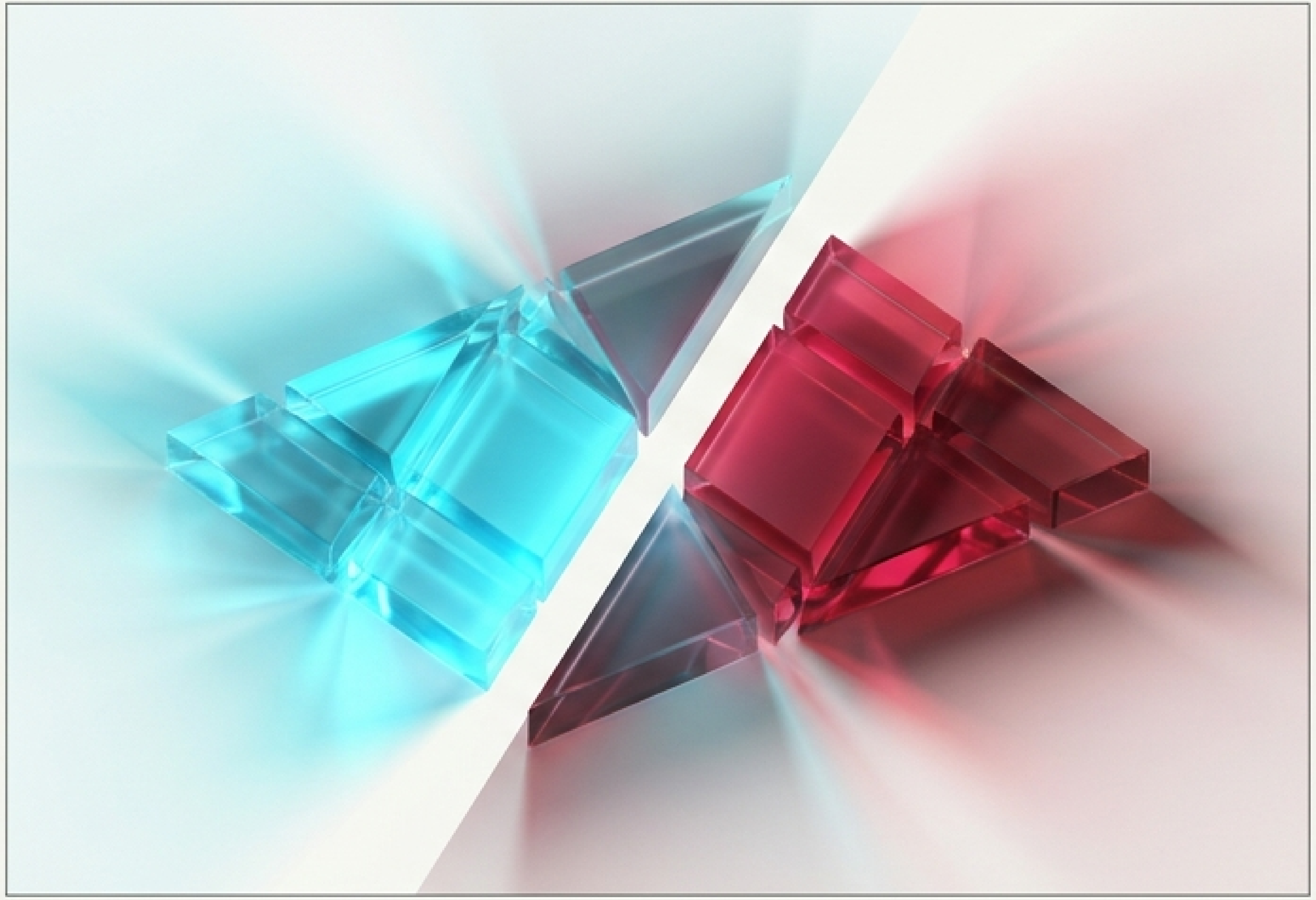


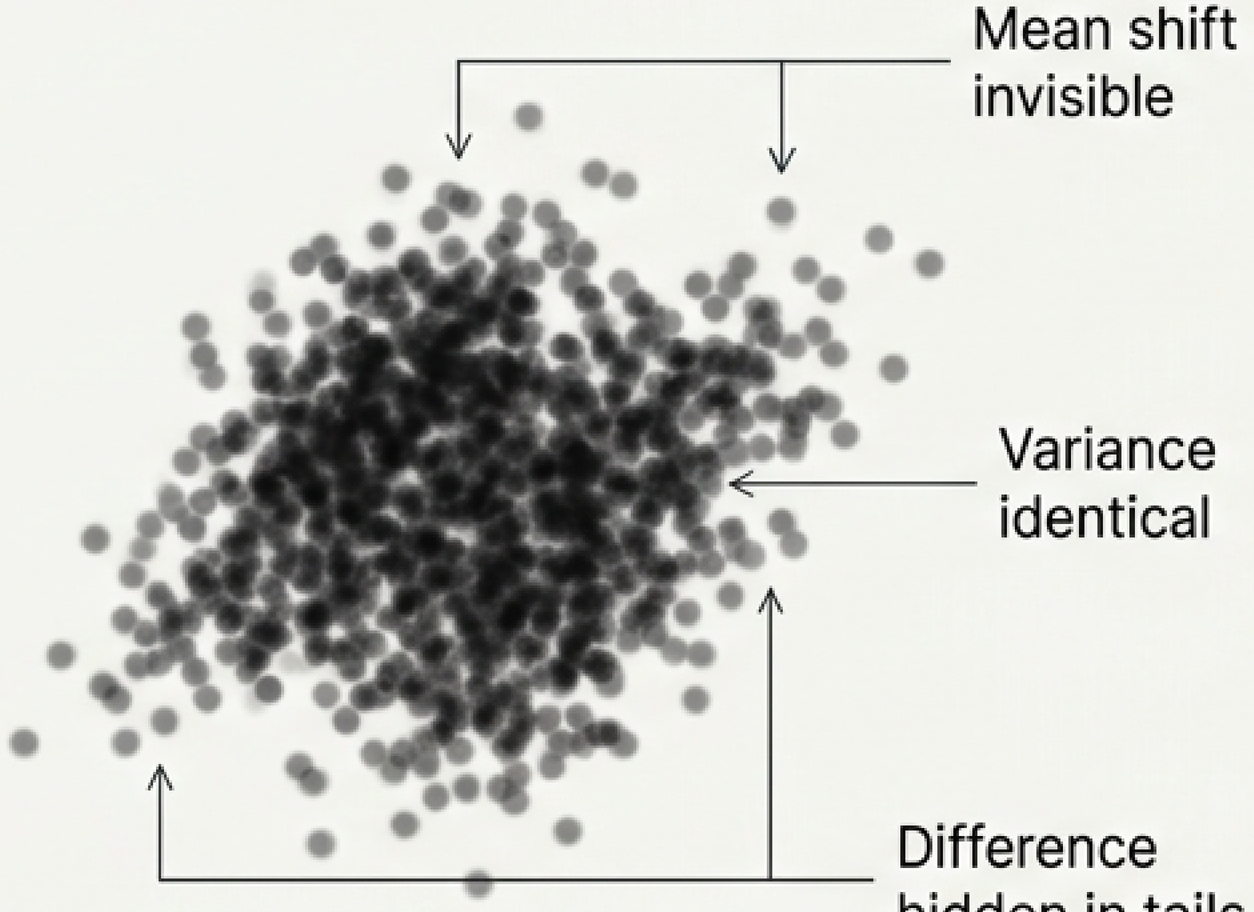
Kernel Embeddings & The Separation of Measure

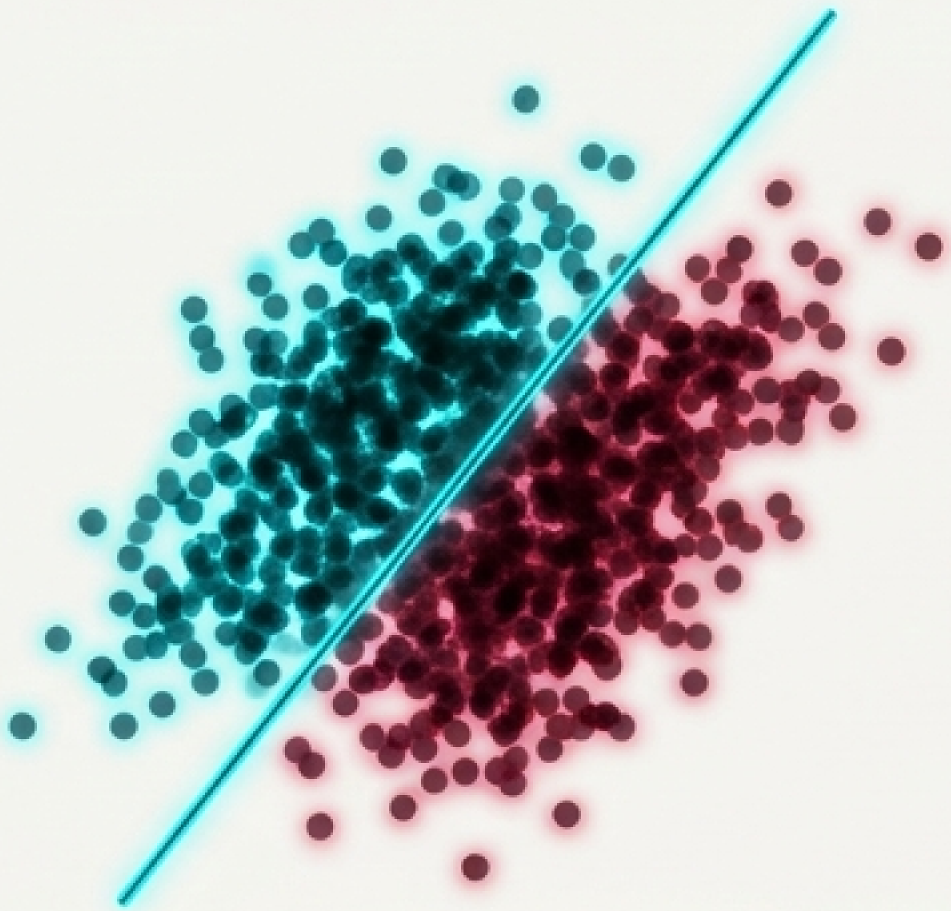
Translating an infinite-dimensional mathematical phenomenon into the geometric mechanics of distinction amplification.

Based on the monograph of the 2026 PNAS paper by Santoro, Waghmare, & Panaretos.

NotebookLM

The nonparametric alternative is too vast for scalar summaries.





The Challenge.

In complex spaces, two continuous distributions P and Q can differ in infinitely many ways. Any finite-dimensional summary compresses information, inevitably missing subtle, localized, or high-order deviations.

The Mystery.

Empirically, mapping these distributions into a Reproducing Kernel Hilbert Space (RKHS) yields inexplicably powerful two-sample tests. Why do kernel methods work so well, even when the differences are infinitesimally subtle?

Lifting data into the geometry of infinite dimensions.

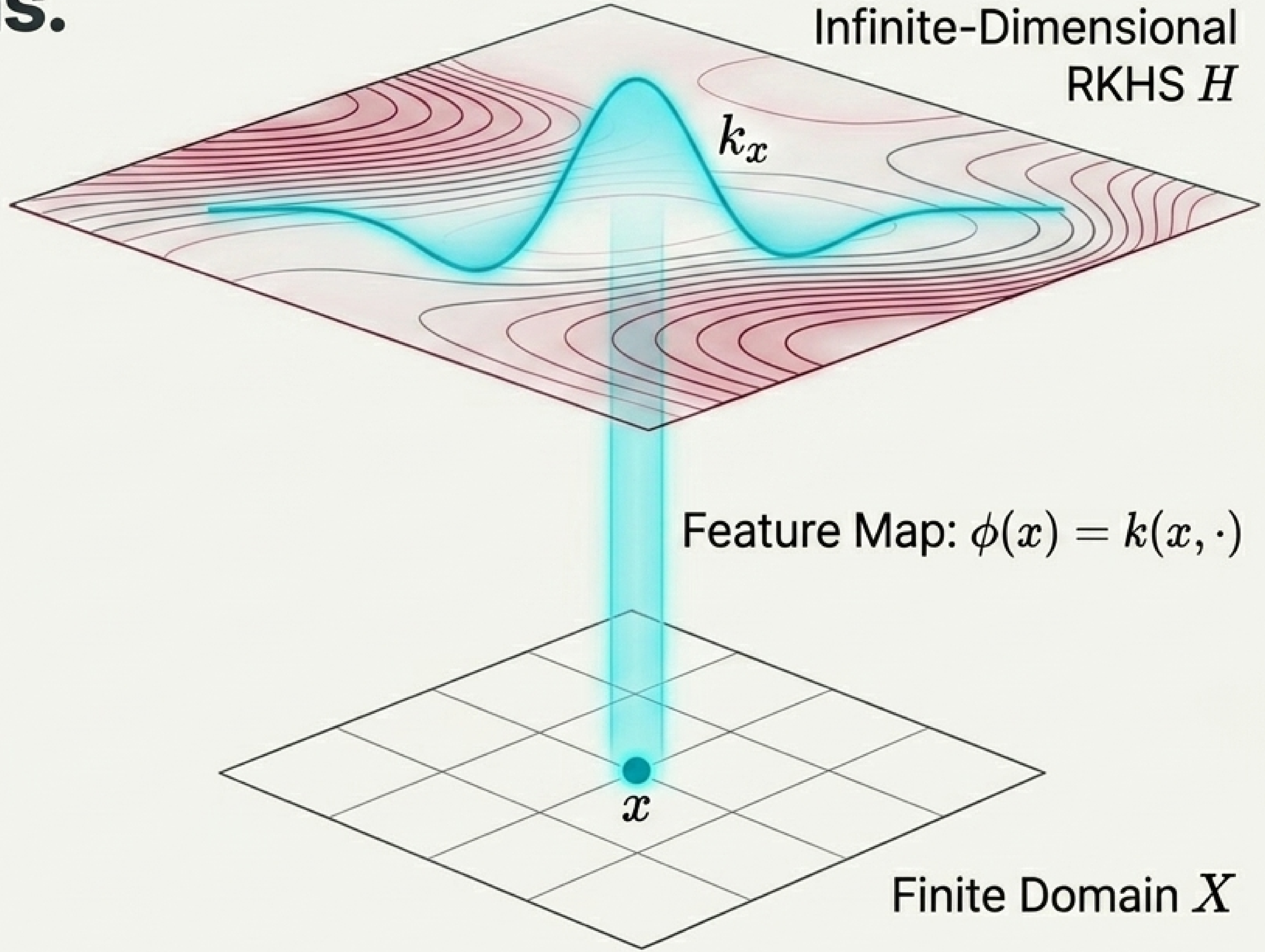
1. The Kernel

We replace raw data with functions. A positive semidefinite kernel (like the Gaussian RBF) computes similarities, acting as an inner product in a vastly larger space.

2. The RKHS

The Reproducing Kernel Hilbert Space H is the completion of these feature maps.

Here, functions cannot oscillate wildly—their values are strictly controlled by their geometric norm.



Inverting the intuition: The blessing of infinite dimensionality.

The Curse of Dimensionality	The Blessing of Infinite Dimensionality
Domain Finite $\mathbb{R}^d$ (d is large).	Domain Infinite-dimensional RKHS.
Primary Effect Sample complexity grows exponentially. Distinguishing distributions becomes practically impossible.	Primary Effect Distinct continuous distributions become perfectly separated.
Mechanism Too many parameters to estimate with limited data.	Mechanism Perturbations accumulate across infinite eigenmodes until their sum diverges.
Regime Statistical / Estimation focus.	Regime Information-theoretic / Structural focus.

The mean embedding preserves distinctions, but fails to amplify them

Maximum Mean Discrepancy

$$\mathrm{MMD}^2(P,Q) = \|\mathbf{m}_P - \mathbf{m}_Q\|_H^2$$

Under a characteristic kernel, the mean embedding is injective ($\mathbf{m}_P = \mathbf{m}_Q \Leftrightarrow P = Q$). This gives us a proper metric.

The Limitation

The MMD measures the distance between two vectors. Two distributions can differ drastically in their higher-order structure, yet their mean vectors might be arbitrarily close. Distinctions are preserved, but not mathematically amplified.

We need to look at operators, not just vectors.

The geometry of infinite Gaussians and the Cameron-Martin paradox

1. Covariance Operators

We embed P as a trace-class covariance operator $S_P = \int k_u \otimes k_u dP(u)$.

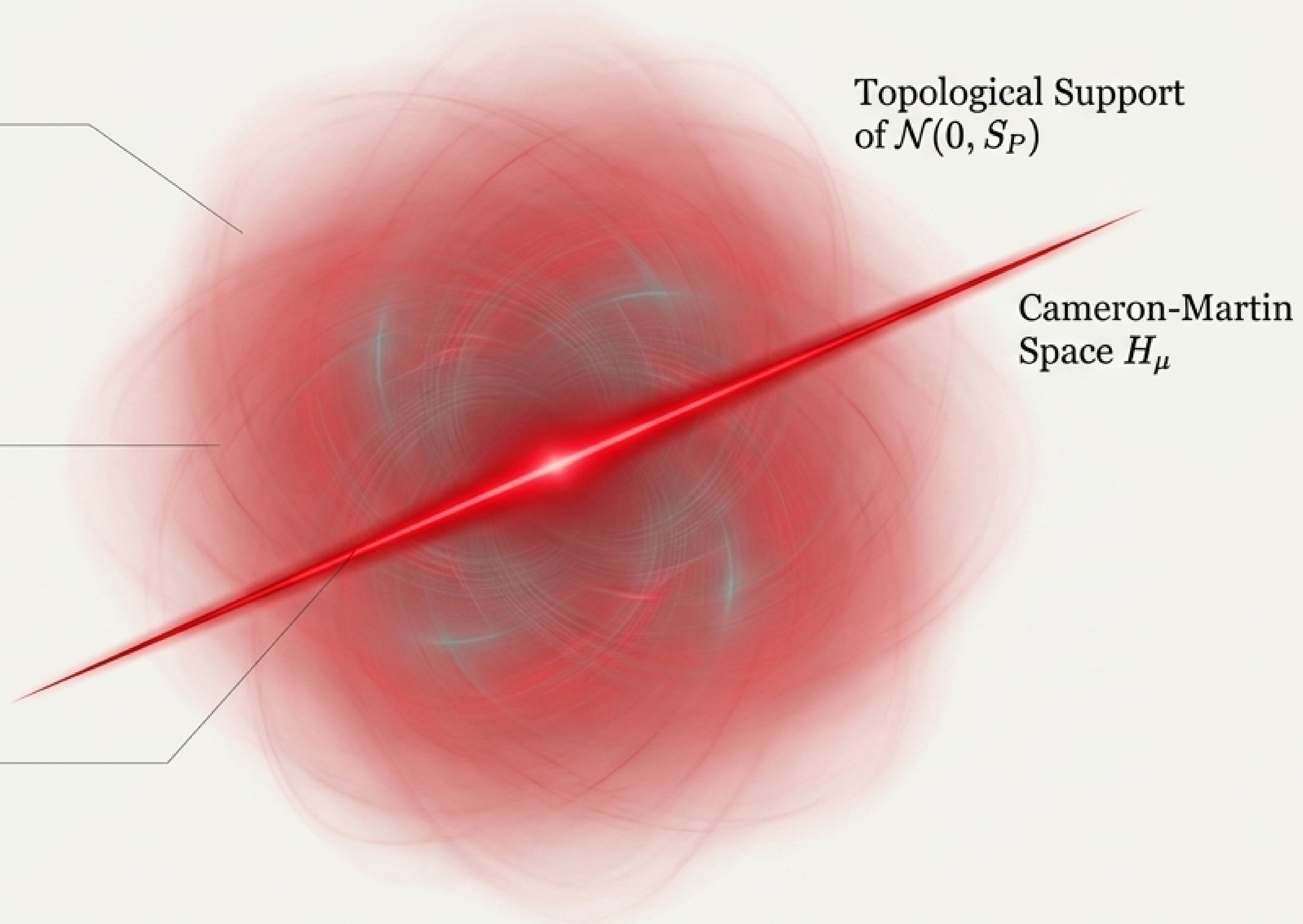
This captures the second-order structure of the distribution in the RKHS.

2. The Gaussian

We associate a Gaussian measure $\mathcal{N}(0, S_P)$ to this operator. Typical samples fall into this vast outer topological support.

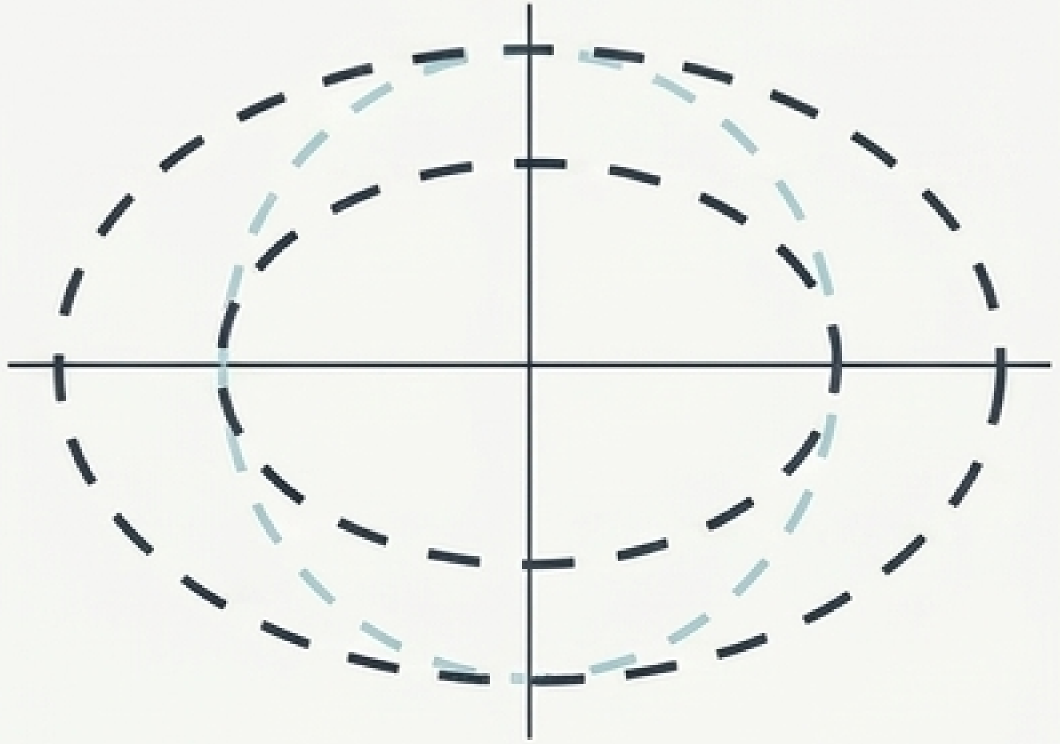
3. The Paradox

In infinite dimensions, a Gaussian assigns zero measure to its own Cameron-Martin space (its dense core). This strange geometry is the engine of the separation phenomenon.



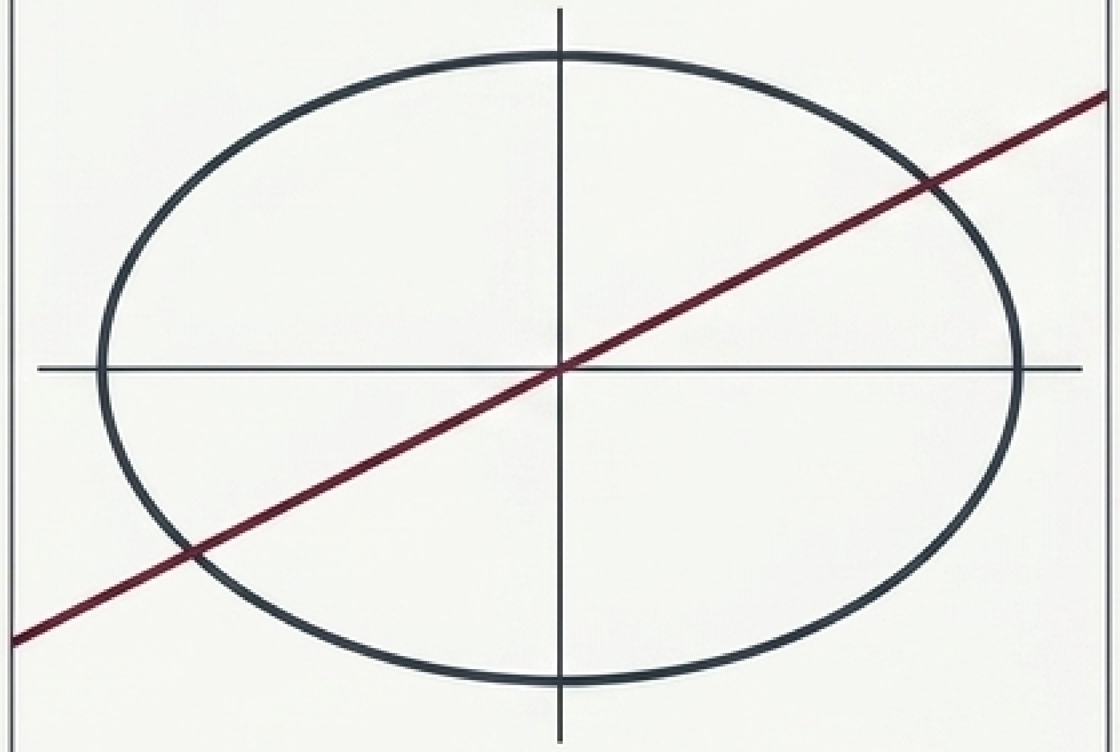
The Feldman-Hájek Dichotomy: Gaussians are either perfectly equivalent or mutually singular.

Equivalent



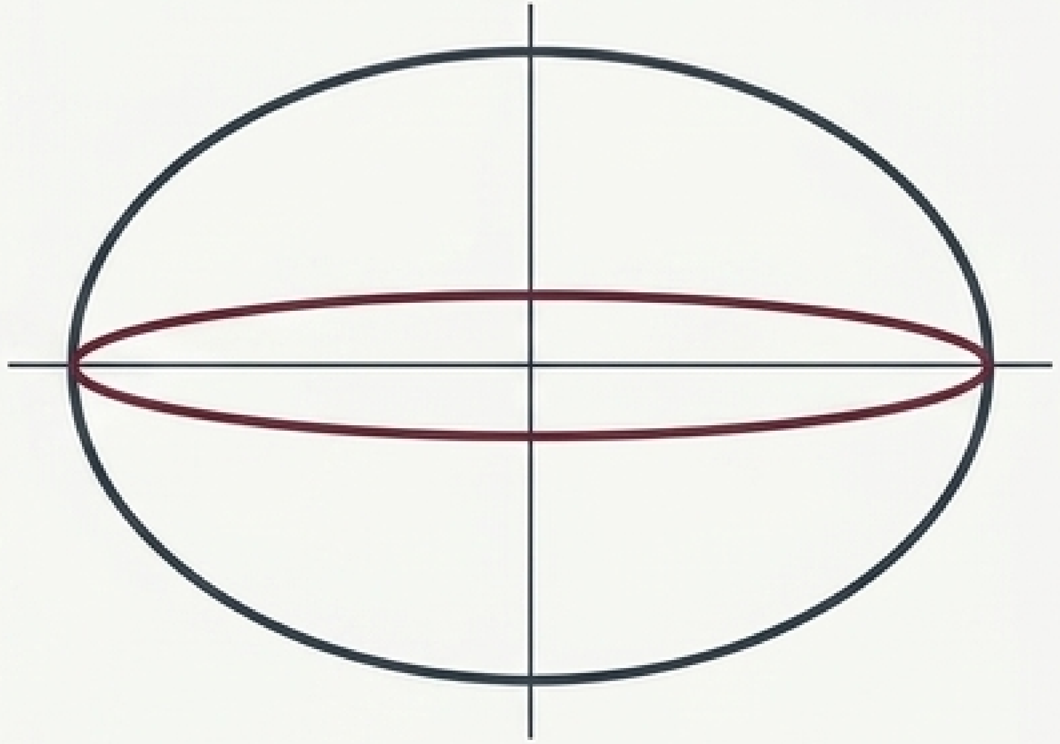
Same Support, Convergent Perturbation

Mutually Singular



Different Cameron-Martin Spaces

Mutually Singular



Divergent Hilbert-Schmidt Perturbation

In finite dimensions, non-degenerate Gaussians are always equivalent (they overlap). In infinite dimensions, there is no intermediate state. If the Cameron-Martin spaces differ, OR if the relative spectral perturbation fails to be square-summable, the measures become mutually singular—existing in totally disjoint universes.

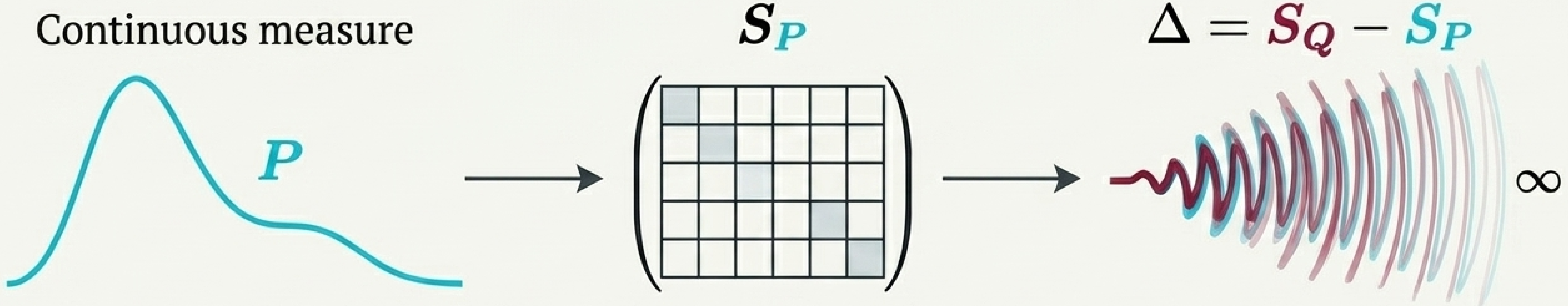
The Main Theorem: Separation of Measure.

Let P and Q be non-atomic (continuous) probability measures on a locally compact Polish space, embedded via a bounded universal kernel. Then:

$$P \neq Q \Leftrightarrow N(0, S_P) \perp N(0, S_Q)$$

Distinct continuous distributions don't just yield different covariance operators. Their Gaussian embeddings become completely, mutually singular. They live in non-overlapping regions of the infinite-dimensional space. Every continuous difference, no matter how subtle, translates to perfect geometric separation.

The Structural Miracle: Why the Hilbert-Schmidt condition fails.



The Trap

It is a common misconception that any two distinct trace-class operators yield singular Gaussians. They do not.

The Mechanism

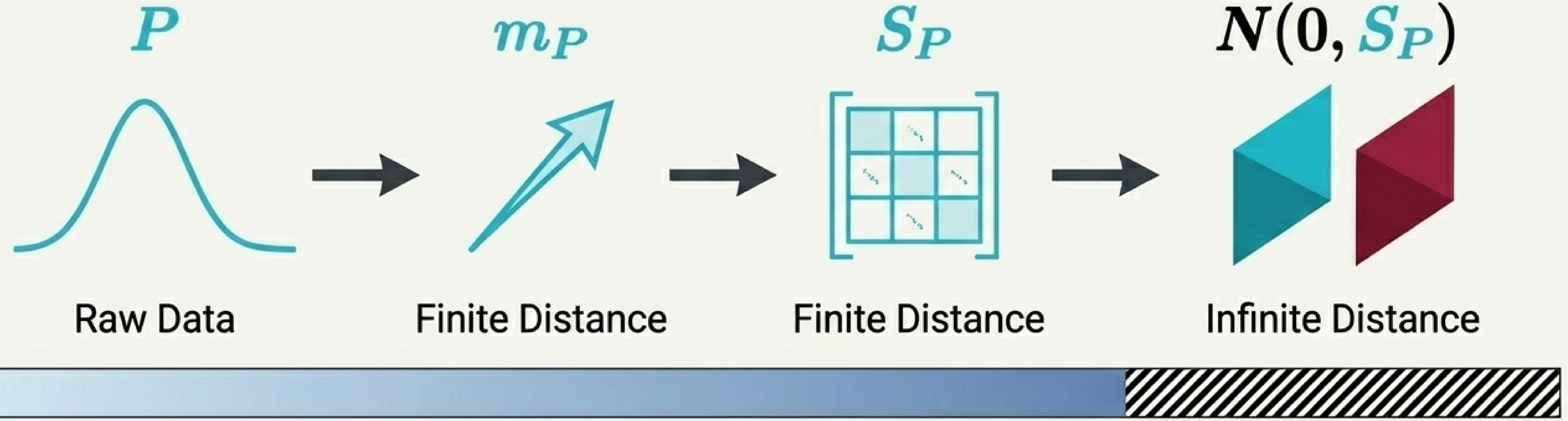
The separation is forced by the specific integral structure of kernel covariance operators. Because S_P and S_Q are generated by integrating continuous measures across the feature map, their difference Δ inherits the full measure-theoretic weight of $Q - P$.

The Result

This structural difference propagates across all infinite eigenmodes.

The relative perturbation $\sum(\lambda_{2,j}/\lambda_{1,j} - 1)^2$ is mathematically forced to diverge to $+\infty$.

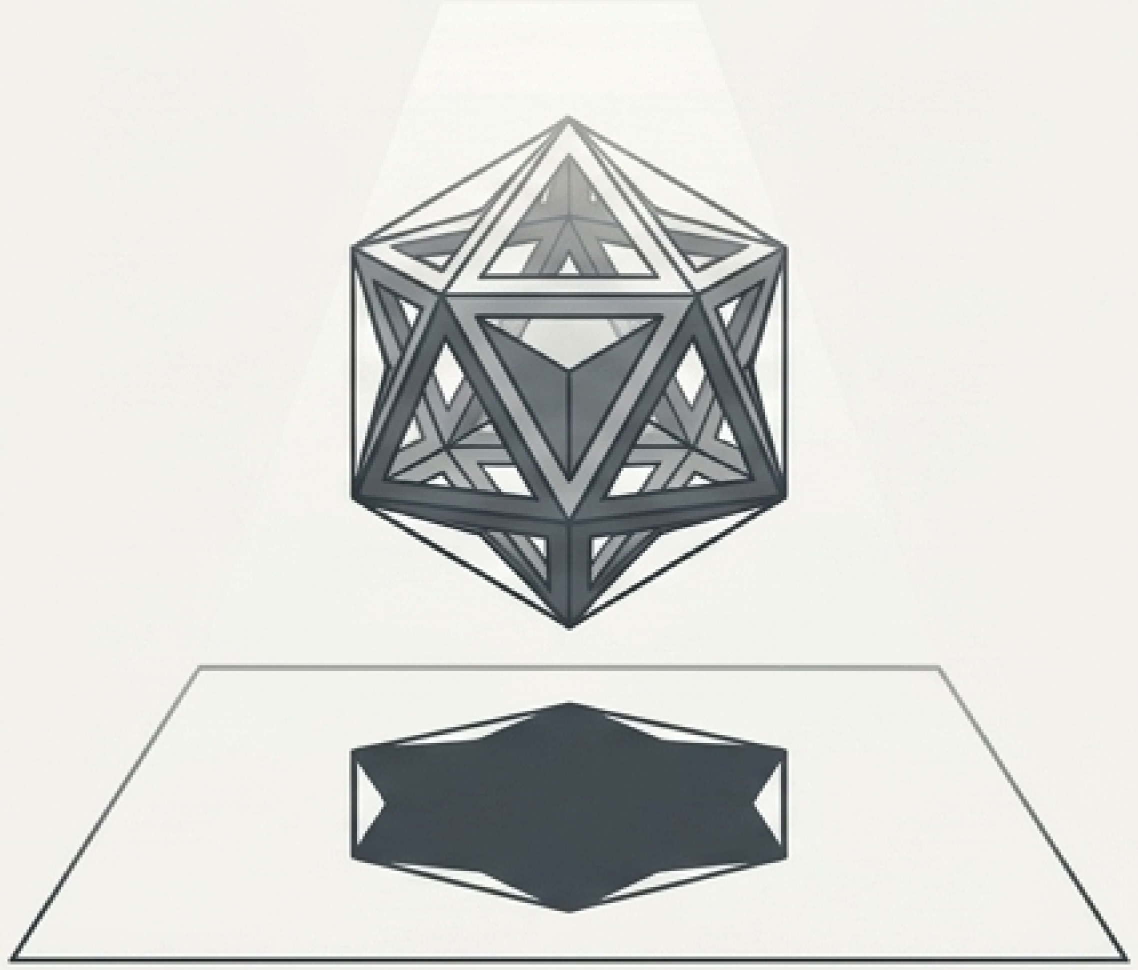
The Hierarchy of Representations: A geometric phase transition.



The raw data (information) remains identical at every step. What changes is the geometry of distinguishability. We move from linear metric spaces (where distances are finite and graded) into information-geometric space. The Gaussian embedding is a lens that triggers a phase transition from graded metrics to a binary singularity dichotomy.

Dualities of Representation: Projection vs. Amplification.

Projection



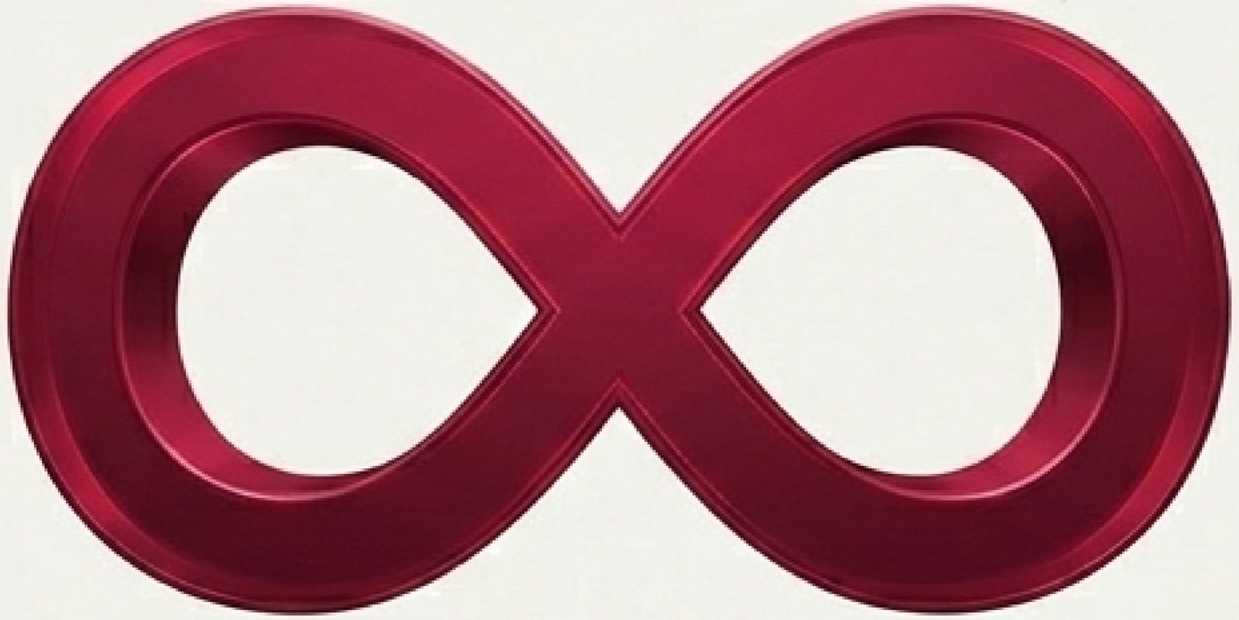
- **Dimension:** Dimensionality reduction (Lower).
- **Mapping Type:** Many-to-one.
- **Effect on Distinctions:** Collapses distances, erasing subtle information.
- **Information State:** Irreversibly lost.

Gaussian Embedding

- **Dimension:** Dimensional expansion (Infinite).
- **Mapping Type:** Strictly injective.
- **Effect on Distinctions:** Amplifies subtle differences into perfect singularity.
- **Information State:** Perfectly reorganized and magnified.

The 0 / ∞ Law for the Likelihood Ratio.

Null Hypothesis $P = Q$



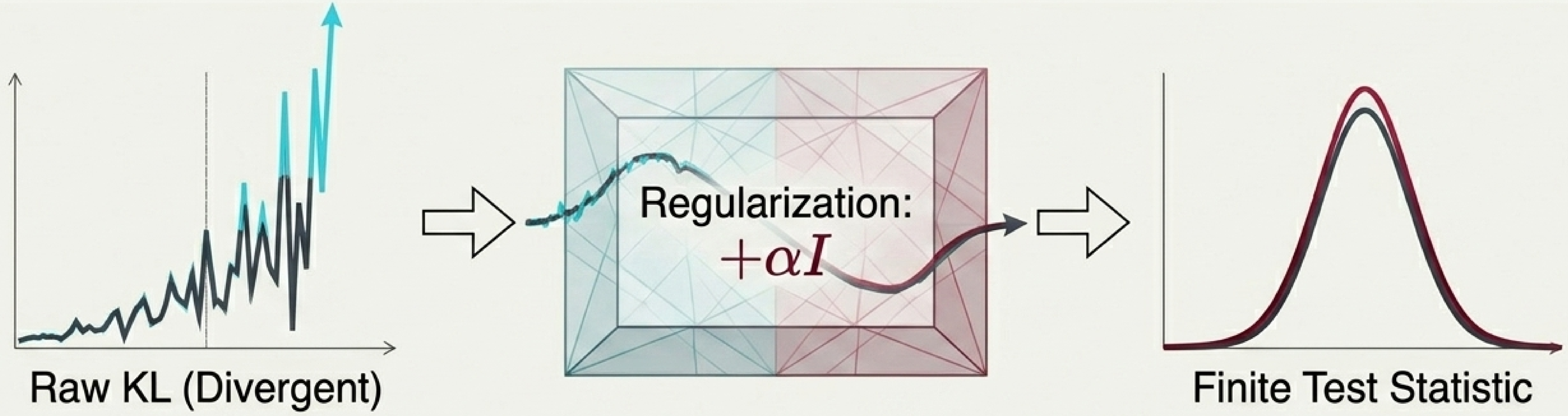
Alternative Hypothesis $P \neq Q$

The mutual singularity translates to a profound information-theoretic reality. The Kullback-Leibler (KL) divergence between the Gaussian embeddings satisfies a strict binary law:

- $\mathrm{KL} = 0$ when the distributions are identical.
- $\mathrm{KL} = +\infty$ when the distributions differ (continuous).

Takeaway: The embedded hypotheses are maximally separated. There is no intermediate distance in the population limit.

From population theory to finite-sample practice.



The Problem

Because the true likelihood ratio diverges to $+\infty$ under the alternative, the raw test statistic cannot be computed directly from finite data.

The Solution

We introduce spectral regularization, replacing the empirical covariance $\hat{S}_Q$ with $\hat{S}_Q + \alpha I$.

The Impact

This smooths the operators, ensuring finite Fredholm log-determinants. As $n \to \infty$ and $\alpha \to 0$, the test achieves remarkable uniform power, outperforming standard MMD tests by directly targeting the covariance separation.

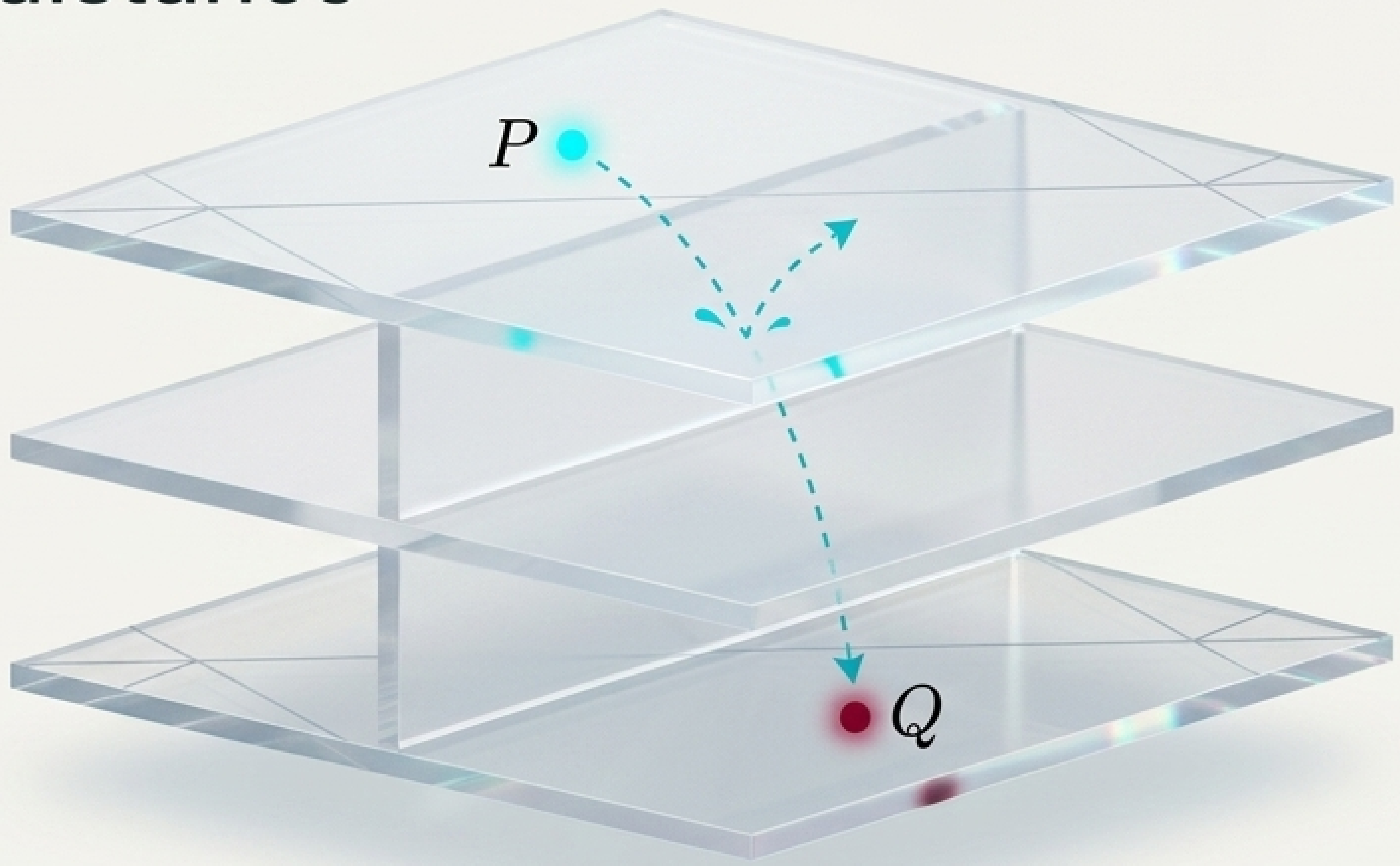
The Reachability Paradigm: Moving beyond metric distance

Classical statistics asks:

How far apart are P and Q?

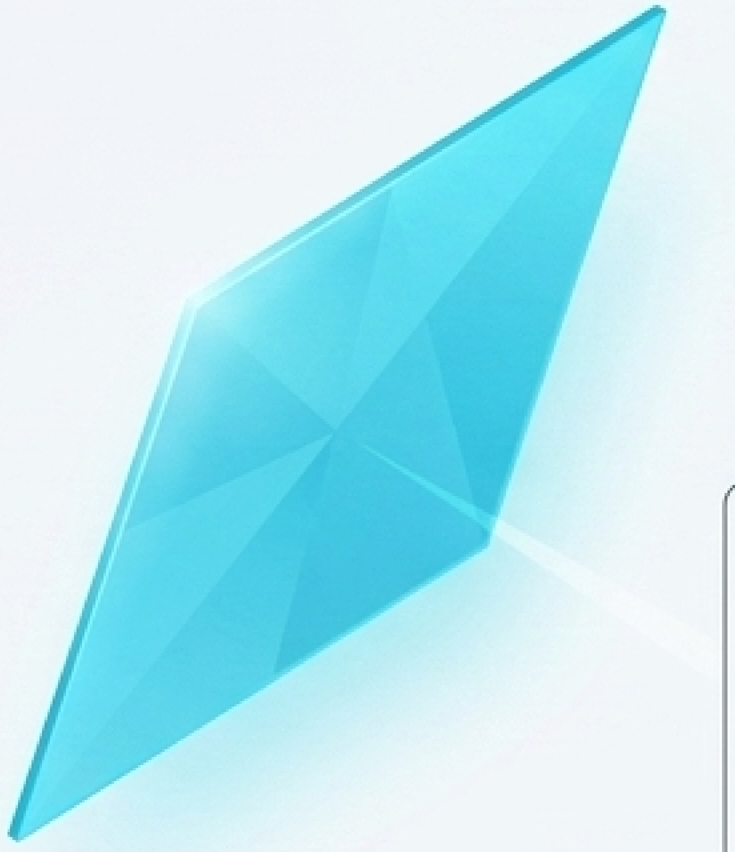
The separation of measure phenomenon asks a totally different question:

Do P and Q share the same admissible stratum?



Insight: Two measures are mutually singular when one cannot ‘reach’ the other by any absolutely continuous transformation. The kernel embedding assigns each distinct distribution its own incompatible geometric universe. They are separated not by vast distance, but by fundamental unreachable structure.

The True Power of Representation



The important effect is information amplification. A good representation does not merely retain distinctions. It can magnify them until they become geometrically unavoidable.

The Blessing of Infinite Dimensionality proves that the hardest nonparametric statistical problems can be perfectly resolved in the population limit. By choosing the right lens—the infinite Gaussian embedding—the invisible becomes perfectly separable.