Discrepancy Calculus: Foundations and Core Theory

A Referential Introduction to the Measure–Theoretic Framework for Singular Analysis and Structure–Aware Machine Learning

Roy S. Colca Jr.
Convergent Intelligence LLC: Research Division
convergentintel.com

March 2026

Abstract

We present the core definitions, axioms, and principal theorems of Discrepancy Calculus (DISC)—a measure–theoretic framework that treats singularities as primary mathematical structure rather than pathology. The central object is the *discrepancy operator*, which quantifies the mismatch between integration and differentiation on metric–measure spaces; classical calculus is recovered as a degenerate smooth limit. We state the eight axioms of DISC, prove the Mesh Fundamental Identity (the DISC replacement for the Fundamental Theorem of Calculus), introduce the counter-derivative construction that unfolds singular calculus into ordinary analysis, and establish three key results: (i) the Classical Shadow theorem showing exact recovery of Newton/Lagrange/Hamilton in smooth regimes, (ii) the DISC Incompleteness theorem proving classical Sobolev spaces cannot extend to gap–rich domains, and (iii) the Meta–Discrepancy theorem establishing a fundamental impossibility—when gap measure and discrepancy energy are both positive, the classical derivative/FTC/MVT package cannot hold on any set of positive measure. We demonstrate operational deployment across 49 published models on HuggingFace (22,500+ downloads) via Topological Knowledge Distillation, which uses the BV decomposition to preserve structural information that standard knowledge distillation provably destroys. The full proof apparatus, together with extensions to graph structures, quantum mechanics, and unified field theory, is developed in the companion monograph *"On the Formal Analysis of Discrepancy Calculus"* (Colca, 2026).

Contents

1 Introduction **3**
1.1 Motivation and Central Object . . . 3
1.2 Scope and Companion Monograph . . . 3

2 The BV Decomposition 3

3 The Mesh Fundamental Identity 4

4 The Counter-Derivative 4

5 The Eight Axioms of Discrepancy Calculus 5

6 Principal Theorems 6
6.1 Classical Recovery 6
6.2 DISC Incompleteness of Classical Sobolev Spaces 6
6.3 The Meta–Discrepancy Theorem 6

7 Separation Results 7

8 Application: Topological Knowledge Distillation 7
8.1 The Problem with Standard KD 7
8.2 TKD Pipeline 8
8.3 Empirical Deployment 8

9 Conclusion 8

1 Introduction

"Truth hides in the difference between what is measured and what is expected."
— R.S.C.

1.1 Motivation and Central Object

Discrepancy Calculus reconciles the mismatch between integration and differentiation in the presence of singularities, pathological oscillations, or measure concentration. It is measure–theoretic at heart, yet designed for symbolic and computational use: a rigorous analytic language for irregular domains and a formal apparatus for structure–aware inference.

The central object is the *discrepancy operator*

$$Df(x) \;=\; \lim_{\varepsilon\downarrow 0} \frac{1}{\varepsilon} \int_x^{x+\varepsilon} \frac{|f(t)-f(x)|}{|t-x|}\, dt, \tag{1}$$

whenever the limit exists (possibly $+\infty$). If f is C^1, then $Df(x) = |f'(x)|$. When f is rough, D quantifies the *average local slope* and, when divergent, localizes irregularity to null sets while preserving integral structure.

1.2 Scope and Companion Monograph

This paper presents the *core analytical foundations* of DISC: the axioms, principal theorems, and one application domain (machine learning). The full proof apparatus (203 pages, 41 chapters), together with extensions to graph structures (Part II), quantum mechanics (Part III), and unified field theory including the Theory of Other (Part IV), is developed in the companion monograph:

> R. S. Colca Jr., *"On the Formal Analysis of Discrepancy Calculus: A Measure–Theoretic and Symbolic Framework for Singular Structures and Stability,"* Convergent Intelligence LLC: Research Division, March 2026.

All theorem numbering in this paper matches the monograph for cross-referencing.

2 The BV Decomposition

Every function of bounded variation admits a canonical decomposition of its distributional derivative into three structurally distinct components. This decomposition is the foundation on which all of DISC is built.

For $f \in \mathrm{BV}(I)$ on a compact interval $I = [a, b]$, the distributional derivative Df is a finite signed Radon measure admitting the Lebesgue decomposition:

$$Df \;=\; f'\,\mathcal{L}^1 \;+\; D^j f \;+\; D^c f, \tag{2}$$

where $f' \in L^1(I)$ is the absolutely continuous (AC) part, $D^j f$ is the jump part (purely atomic, supported on the at most countable jump set J_f), and $D^c f$ is the Cantor part (singular-continuous, supported on a set of Lebesgue measure zero but possibly positive Hausdorff dimension).

The *singular masses* are $S_f^j := |D^j f|(I)$, $S_f^c := |D^c f|(I)$, and $S_f := S_f^j + S_f^c$. Classical calculus operates entirely in the regime where $S_f = 0$. DISC operates in the general case.

3 The Mesh Fundamental Identity

The Mesh Fundamental Identity is the DISC replacement for the classical Fundamental Theorem of Calculus.

Theorem 3.1 (Fundamental identity for BV). *For every $f \in \mathrm{BV}(I)$,*

$$f(b) - f(a) \; = \; \int_a^b f'(x)\, dx \; + \; \sum_{x \in J_f} \Delta f(x) \; + \; D^c f(I). \tag{3}$$

The classical FTC is recovered when the last two terms vanish—i.e., when f has no jumps and no Cantor part. The identity shows that total change is always the sum of three structurally distinct contributions: smooth accumulation, discrete jumps, and singular-continuous drift. Standard analysis accounts only for the first.

4 The Counter-Derivative

The counter-derivative is a novel construction that *unfolds* singular calculus into ordinary analysis on an expanded domain.

Definition 4.1 (Counter-derivative). *For $f \in BV(I)$, a* counter-derivative *is a function $\widetilde{f} \in C^0(\widetilde{I})$ such that:*

1. Trace: $\widetilde{f} = f$ *on* $I \setminus J_f$*;*
2. Per-gap affine connector: *for $x \in J_f$ with $L = f^-(x)$ and $R = f^+(x)$,*

$$\widetilde{f}|_{\Delta_x}(t) = L + \frac{R - L}{r_x - \ell_x}\,(t - \ell_x).$$

Theorem 4.2 (Existence, continuity, and AC regularity). *For every $f \in BV(I)$ there exists a counter-derivative $\widetilde{f} \in \mathrm{AC}(\widetilde{I})$ with $D\widetilde{f} = \dot{\widetilde{f}}\, d\sigma$ for some $\dot{\widetilde{f}} \in L^1(\widetilde{I})$.*

Theorem 4.3 (Projection of functions and derivatives). *With $\Phi : \widetilde{I} \to I$ as the collapse map, $\Phi_* \widetilde{f} = f$ and $\Phi_*(D\widetilde{f}) = Df$.*

The significance: singular calculus on I becomes *ordinary* calculus on $\widetilde{I}$. Compute on the unfolded domain where everything is smooth, then project back.

Theorem 4.4 (Counter-FTC). *If $\widetilde{f} \in \mathrm{AC}(\widetilde{I})$ with $D\widetilde{f} = \dot{\widetilde{f}}\, d\sigma$, then for any $\tilde{a}, \tilde{b} \in \widetilde{I}$,*

$$\widetilde{f}(\tilde{b}) - \widetilde{f}(\tilde{a}) = \mathcal{C}\int_{\tilde{a}}^{\tilde{b}} \dot{\widetilde{f}}\, d\sigma.$$

5 The Eight Axioms of Discrepancy Calculus

We axiomatize DISC on metric–measure spaces (X, d, μ) where (X, d) is complete and separable, and μ is a Borel Radon measure finite on bounded sets.

Axiom 1 (Discrepancy derivative (metric slope))**.** *For Borel $f : X \to \mathbb{R}$,*

$$Df(x) := \limsup_{r\downarrow 0} \sup_{0<d(x,y)<r} \frac{|f(y) - f(x)|}{d(x,y)} \in [0, \infty].$$

If X is a smooth Riemannian manifold and $f \in C^1$, then $Df(x) = \|\nabla f(x)\|$.

Axiom 2 (Discrepancy energy)**.** *Let $w : X \to (0, \infty)$ be measurable and essentially bounded above/below on bounded sets. Define*

$$E_{\mathrm{disc}}[f] := \frac{1}{2} \int_X w(x)\, (Df(x))^2\, d\mu(x),$$

and the Sobolev space $W^{1,D,2}(X)$ with norm $\|f\|^2_{W^{1,D,2}} := \|f\|^2_{L^2} + 2E_{\mathrm{disc}}[f]$.

Axiom 3 (DG-limit (discrepancy-guided limit))**.** *For $X = \mathbb{R}$ and $a \in \mathbb{R}$,*

$$\mathrm{D}\lim_{x\to a} f(x) \;:=\; \lim_{\varepsilon\downarrow 0} \frac{1}{2\varepsilon} \int_{a-\varepsilon}^{a+\varepsilon} f(t)\, dt$$

whenever the limit exists.

Axiom 4 (Gap geometry)**.** *For measurable $E \subset X$, define the* gap set

$$\Delta(E) := \{x \in X :\; \theta^{*E}(x) > \theta^{E}_{*}(x)\}.$$

The Position *map* $\mathrm{Position}(x) = (\theta^E_*(x), \theta^{*E}(x))$ *takes values in $P := \{(a, b) \in [0, 1]^2 : a < b\}$. Define $d_{\mathrm{gap}}(x, y) := \|\mathrm{Position}(x) - \mathrm{Position}(y)\|_2$ and $\mu_{\mathrm{gap}} := \mathrm{Position}_{\#}\mu$.*

Axiom 5 (Gap calculus)**.** *For $F : P \to \mathbb{R}$ measurable, define directional gap difference quotients $D^v_{\mathrm{gap},\varepsilon}F$ and, when limits exist in $L^2(P, \mu_{\mathrm{gap}})$, the gap gradient $\nabla_{\mathrm{gap}}F$ and Laplacian $\Delta_{\mathrm{gap}}F$.*

Axiom 6 (Function spaces and DG-absolute continuity)**.** *Define $W^{1,D,p}(X)$ via Df as generalized gradient. A curve γ is rectifiable if it has finite length; f is* DG-absolutely continuous *if for a.e. rectifiable γ, $|f(\gamma(1)) - f(\gamma(0))| \le \int_\gamma Df\, ds$.*

Axiom 7 (Fundamental discrepancy relation)**.** *For $f \in W^{1,D,1}_{\mathrm{loc}}(\mathbb{R})$ and any interval $[a, b]$,*

$$|f(b) - f(a)| \le \int_a^b Df(x)\, dx.$$

Consequently, $\frac{|f(b)-f(a)|}{|b-a|} \le \operatorname{ess\,sup}_{(a,b)} Df$. Moreover, secant slopes lie in the closed convex hull of the essential range of the metric differential.

Axiom 8 (Discrepancy Implicit Function Theorem (D–IFT))**.** *Let $F : U \times V \to \mathbb{R}$ with $U, V \subset \mathbb{R}^n$ open. Assume at (x_0, y_0):*

1. Vertical nondegeneracy*: $|F(x_0, y) - F(x_0, y')| \ge m\, \|y - y'\|$ for y, y' near y_0, $m > 0$;*
2. Horizontal calmness*: $|F(x, y) - F(x', y)| \le k\|x - x'\|$ near (x_0, y_0) with $0 < k < m$.*

Then there exist neighborhoods and a unique φ with $F(x, \varphi(x)) = 0$, $\mathrm{Lip}(\varphi) \le k/m$.

6 Principal Theorems

6.1 Classical Recovery

Theorem 6.1 (Classical shadow)**.** *On a smooth domain $\Omega \subset \mathbb{R}^n$, if $\mu_{\text{gap}} \equiv 0$ and $f \in C^1(\Omega)$, then* (a) *$Df = \|\nabla f\|$ a.e.;* (b) *E_{disc} reduces to the Dirichlet integral;* (c) *DG-limits equal classical limits at continuity points;* (d) *discrepancy mechanics reduce to Newton/Lagrange/Hamilton.*

This theorem establishes that classical analysis is a *degenerate smooth limit* of DISC. DISC is strictly more general; it contains classical analysis as a special case.

6.2 DISC Incompleteness of Classical Sobolev Spaces

Theorem 6.2 (DISC incompleteness of classical)**.** *There exist a compact $E \subset [0,1]$ with $\mu(E) > 0$ and $\mu_{\text{gap}}(E) > 0$ and a function $f \in W^{1,D,2}(E)$ such that no $g \in W^{1,2}([0,1])$ satisfies $g = f$ a.e. on E.*

Proof sketch. Construct a fat Cantor set E of positive measure. Define $f := \sum_k a_k \phi_k$ where ϕ_k are Lipschitz tent functions on the removed intervals I_k with amplitudes chosen so that $\sum a_k^2 < \infty$ and $\sum a_k^2 \ell_k^{-1} < \infty$. On E, the induced $Df \in L^2$ by the second series. But any $g \in W^{1,2}([0,1])$ matching f on E would require $\int |g'|^2 = +\infty$ by a Poincaré-type bound across scales. Full proof in the monograph, Chapter 11. □

6.3 The Meta–Discrepancy Theorem

Definition 6.3 (Gap-roughness condition (GRC))**.** *Let $E \subset X$ be measurable with $\mu_{\text{gap}}(E) > 0$. A function $f \in L^1_{\text{loc}}(E)$ satisfies* GRC *on E if there exist $c > 0$ and a set $A \subset E$ with $\mu(A) > 0$ such that for every $x \in A$ there are radii $r_k \downarrow 0$ and pairs of points $y_k^{\pm} \in B(x, r_k) \cap E$ with*

$$\left| \frac{f(y_k^+) - f(x)}{d(x, y_k^+)} - \frac{f(x) - f(y_k^-)}{d(y_k^-, x)} \right| \ \geq \ c.$$

Theorem 6.4 (Meta–Discrepancy)**.** *Let $E \subset X$ be measurable with $\mu_{\text{gap}}(E) > 0$. Let $f \in L^1_{\text{loc}}(E)$ with $E_{\text{disc}}[f; E] > 0$ and assume f satisfies GRC on a subset $A \subset E$ of positive measure. Then there do not exist a function $g \in L^1_{\text{loc}}(E)$ and a pointwise classical derivative operator $\mathcal{D}$ on E such that simultaneously:*

1. ***FTC pairing:*** *for a.e. rectifiable segment $[a, b] \subset E$, $f(b) - f(a) = \int_a^b g$, and $g = \mathcal{D}f$ a.e.;*
2. ***MVT/chain-rule structure:*** *on a set of positive measure of such segments, a classical mean-value identity holds.*

In particular, if such a package holds on E, then necessarily $\mu_{\text{gap}}(E) = 0$ and f is in the classical smooth regime a.e.

Proof sketch. The GRC at points $x \in A$ produces two-sided difference quotients separated by $\geq c > 0$. If the FTC/MVT package held, the pointwise derivative would satisfy the Darboux property along segments through x. But the separated quotients violate Darboux on any segment intersecting both cones of approach, contradicting condition (2). Full proof in the monograph, Chapter 11. $\square$

Consequence. Positive gap + positive discrepancy energy $\Rightarrow$ the classical derivative/FTC/MVT package is *impossible* on positive measure. This is not a heuristic—it is a mathematical impossibility result. Any method that implicitly assumes smooth teacher distributions (including standard knowledge distillation via KL divergence) provably cannot capture the structural information that DISC preserves.

7 Separation Results

The following table summarizes what DISC achieves that classical analysis provably cannot.

Statement	**DISC Status**	**Classical Status**
Implicit function without C^1	Theorem (D–IFT)	Inapplicable
Limit across jump discontinuity	DG-limit exists	Undefined
Mean value control at singularity	Axiom 7	MVT fails
Sobolev extension on gap sets	$W^{1,D,2}$ well-defined	No extension (Thm 6.2)
Energy functional on rough domains	E_{disc} finite	Dirichlet integral diverges
FTC/MVT on positive-gap sets	DISC Mesh Identity	Impossible (Thm 6.4)

This table constitutes the mathematical proof that DISC is a *strictly larger* framework than classical analysis—not a reformulation, but a proper extension.

8 Application: Topological Knowledge Distillation

We demonstrate operational deployment of DISC in machine learning through *Topological Knowledge Distillation (TKD)*, a methodology that applies the BV decomposition to knowledge transfer between neural networks.

8.1 The Problem with Standard KD

Standard knowledge distillation [4] minimizes KL divergence between teacher and student softmax distributions. This treats the teacher's output distribution as a smooth function and optimizes globally.

Language, however, is not smooth. Topic shifts, reasoning mode transitions, register changes, and logical pivots create discontinuities in the teacher's output distribution. Standard KD averages across these boundaries.

The Meta–Discrepancy Theorem (6.4) makes this precise: when the teacher's distribution has positive gap measure and positive discrepancy energy—which it does at every structural boundary—the smooth optimization package *provably cannot* capture the full structure.

8.2 TKD Pipeline

TKD treats the teacher's output distribution p_T over a concatenated token stream as a BV function and applies the Mesh Fundamental Identity:

$$p_T(b) - p_T(a) = \underbrace{\int_a^b p_T'(x)\,dx}_{\text{smooth KD}} + \underbrace{\sum_{x \in J_{p_T}} \Delta p_T(x)}_{\text{jump corrections}} + \underbrace{D^c p_T(I)}_{\text{drift corrections}}.$$

The pipeline computes:

1. **Discrepancy energy** $E_{\text{disc}}[p_T]$ over sliding windows to identify regions of high structural information density;

2. **Jump set** $J_{p_T} = \{x : Dp_T(x) > 3\sigma\}$ to locate conceptual boundaries;

3. **Gap energy density** over 64-token windows to capture Cantor-type drift invisible to both smooth and jump analysis;

4. **Topology-guided windowing**: training windows cut at low-discrepancy positions rather than fixed stride.

8.3 Empirical Deployment

TKD and DISC-informed training have been deployed across 49 published models on HuggingFace (`huggingface.co/reaperdoesntknow`), accumulating 22,500+ organic downloads. Key models include:

Model	DISC Application	Downloads
TopologicalQwen	Full TKD (BV decomposition, jump detection)	1,134
Qwen3-1.7B-Thinking-Distil	TKD with Thinking teacher	1,188
DiStil-Qwen3-1.7B-uncensored	Uncensored base for DISC chain	1,030
Qwen3-1.7B-Coder-Distilled-SFT	TKD with Coder teacher	966
DiscoverLM-70M	Metric slope attention, gap geometry	784
DualMind	Continuous Thought Dynamics (Ch. 19)	260
Qemma-GEI	Gap Envelope Integral fusion	423
SAGI	Discrepancy Mechanics (Ch. 16) swarm routing	503

Full methodology is documented in *"Structure Over Scale"* (DOI: 10.57967/hf/8165) and *"From Three Teachers to Dual Cognition"* (DOI: 10.57967/hf/8184).

9 Conclusion

Discrepancy Calculus provides a complete axiomatic framework for analysis on singular domains. The eight axioms (Axioms 1–8) extend classical calculus to regimes where smoothness fails, with:

- Classical recovery in smooth limits (Theorem 6.1);
- Strict separation from classical analysis (Section 7);
- A fundamental impossibility result (Theorem 6.4) proving that the classical FTC/MVT package cannot hold on positive-gap, positive-energy sets;
- Operational deployment in machine learning via TKD across 49 models.

The full theory—including graph-theoretic extensions (Measure–Theoretic Hamilton Cycles, Pattern–Field Resonance Graphs, Murphie's Discrepancy Theorem), quantum mechanics (Discrepancy–Schrödinger Equation, Distributed Anchors), and unified field theory (Theory of Other, No Assumptions Theory)—is developed in the companion monograph.

References

[1] R. S. Colca Jr., *On the Formal Analysis of Discrepancy Calculus: A Measure–Theoretic and Symbolic Framework for Singular Structures and Stability*, Convergent Intelligence LLC: Research Division, March 2026.

[2] R. S. Colca Jr., "Structure Over Scale: CPU-Native Training of Sparse Cognitive Architectures at $1.60 Per Model," HuggingFace, DOI: 10.57967/hf/8165, March 2026.

[3] R. S. Colca Jr., "From Three Teachers to Dual Cognition: Topology-Aware Multi-Teacher Distillation and Role-Conditioned Self-Critique at 1.7B Scale," HuggingFace, DOI: 10.57967/hf/8184, March 2026.

[4] G. Hinton, O. Vinyals, and J. Dean, "Distilling the Knowledge in a Neural Network," *NIPS Deep Learning Workshop*, 2015.