Shadow Math: A Calculus of Dimensional Identity

Version 1.0 (Foundations)

Abstract

Shadow Math formalizes the idea that a lower–dimensional "shadow" carries partial information about a higher–dimensional "self". A 2D shadow of a 3D object, a 3D self as a shadow of a 4D self, and so on, define a ladder of irreversible projections with constrained lifts, information loss, and layered identity. This text introduces:

- layered spaces and shadow projections,
- identity fibers and partial identity transmission,
- shadow channels as information–constrained maps,
- shadow differentials, integrals, and entropy,
- invariants that persist across projection.

The goal is not to model physics directly, but to provide a mathematical language for "identity across dimensions" and its irreversible projections.

Contents

1 Layers, Shadows, and Identity Fibers **2**
1.1 Layered spaces . . . 2
1.2 Shadows and identity fibers . . . 2
1.3 Identity as an equivalence class . . . 3

2 Shadow Measures and Information Loss **3**
2.1 Probability measures on layers . . . 3
2.2 Information loss and fiber entropy . . . 3

3 Shadow Channels and Lifts **4**
3.1 Shadow channels . . . 4
3.2 Lifts and partial identity reconstruction . . . 4

4 Shadow Differential and Integral **5**
4.1 Shadow observables . . . 5
4.2 Shadow integral . . . 5
4.3 Shadow differential . . . 5

5 Shadow Entropy and Invariants **6**
5.1 Shadow entropy . . . 6
5.2 Shadow invariants . . . 6

6 Chains of Shadows and Shadow Calculus **6**
6.1 Shadow chains . . . 6
6.2 Shadow derivative along a chain . . . 7

7 Narrative Summary and Further Directions **7**

1 Layers, Shadows, and Identity Fibers

1.1 Layered spaces

We start from the intuition:

> The 2D shadow is the information of the 3D self. The 3D self is the shadow of a 4D self. And so on, layer by layer.

Definition 1.1 (Layered space). *A* layered space *is a sequence of measurable spaces*

$$\mathcal{L}_0, \mathcal{L}_1, \mathcal{L}_2, \ldots$$

together with measurable maps

$$\pi_{k+1\to k} : \mathcal{L}_{k+1} \to \mathcal{L}_k, \quad k \geq 0,$$

called shadow projections*, satisfying the consistency condition*

$$\pi_{k+2\to k} = \pi_{k+1\to k} \circ \pi_{k+2\to k+1}.$$

We interpret $\mathcal{L}_{k+1}$ as a "higher" layer and $\pi_{k+1\to k}$ as casting a shadow down from layer $k+1$ to k.

Remark 1.2. A concrete example is $\mathcal{L}_k = \mathbb{R}^{n+k}$ with

$$\pi_{k+1\to k}(x_1, \ldots, x_{n+k+1}) = (x_1, \ldots, x_{n+k}),$$

the standard coordinate projection (dropping the last coordinate). More general projections may encode occlusion, compression, or coarse–graining.

1.2 Shadows and identity fibers

Definition 1.3 (Shadow of a point). *Given $x_{k+1} \in \mathcal{L}_{k+1}$, its* shadow in layer k *is*

$$\mathsf{Sh}_k(x_{k+1}) := \pi_{k+1\to k}(x_{k+1}) \in \mathcal{L}_k.$$

Definition 1.4 (Identity fiber). *For a point $x_k \in \mathcal{L}_k$, the* identity fiber above x_k *is*

$$\mathcal{F}_{k+1}(x_k) := \{x_{k+1} \in \mathcal{L}_{k+1} : \pi_{k+1\to k}(x_{k+1}) = x_k\}.$$

Intuition: $\mathcal{F}_{k+1}(x_k)$ is the set of all possible "selves" in layer $k+1$ that cast the same shadow x_k in layer k. Multiple higher identities may share the same lower–dimensional shadow.

Definition 1.5 (Irreversible projection). *We say the projection $\pi_{k+1\to k}$ is* irreversible *on a set $A \subseteq \mathcal{L}_k$ if for all $x_k \in A$,*

$$|\mathcal{F}_{k+1}(x_k)| > 1.$$

More generally, if $\pi_{k+1\to k}$ is not injective on any nontrivial region, we call it an irreversible shadow projection.

Remark 1.6. Irreversibility captures the idea that the 2D shadow does not uniquely determine the 3D object, nor does the 3D self uniquely determine its 4D ancestor. Identity information is partially lost under projection.

1.3 Identity as an equivalence class

Definition 1.7 (Shadow equivalence and identity class). *Given a layered space* $(\mathcal{L}_k, \pi_{k+1\to k})_{k\geq 0}$*, define an equivalence relation on* $\bigsqcup_{k\geq 0} \mathcal{L}_k$ *by:*

$$x \sim y \iff \exists k \text{ and } z \in \mathcal{L}_k \text{ such that } \pi(x) = \pi(y) = z$$

where π *denotes the appropriate chain of projections into layer* k*. The equivalence class* $[x]$ *is called the* identity class *of* x*.*

In words: two points in different layers are considered the same "identity" if there exists some lower layer where their shadows coincide.

2 Shadow Measures and Information Loss

2.1 Probability measures on layers

Definition 2.1 (Layer measure). *Let* $(\mathcal{L}_k, \mathcal{F}_k)$ *be a measurable space. A* layer measure *is a probability measure* μ_k *on* $(\mathcal{L}_k, \mathcal{F}_k)$*.*

Definition 2.2 (Pushforward shadow measure). *Given* μ_{k+1} *on* $\mathcal{L}_{k+1}$ *and a projection* $\pi_{k+1\to k} : \mathcal{L}_{k+1} \to \mathcal{L}_k$*, the* shadow measure *on layer* k *is the pushforward*

$$\mathsf{Sh}_k(\mu_{k+1}) := (\pi_{k+1\to k})_{\#}\mu_{k+1},$$

defined by

$$\mathsf{Sh}_k(\mu_{k+1})(A) = \mu_{k+1}(\pi_{k+1\to k}^{-1}(A)), \quad A \subseteq \mathcal{L}_k \text{ measurable.}$$

Remark 2.3. $\mathsf{Sh}_k(\mu_{k+1})$ encodes what an observer confined to layer k can see of a higher layer distribution μ_{k+1}. Different μ_{k+1} may shadow to the same $\mathsf{Sh}_k(\mu_{k+1})$.

2.2 Information loss and fiber entropy

Assume for simplicity that $\mathcal{L}_k$ and $\mathcal{L}_{k+1}$ are countable or have countable partitions so that entropy is well defined.

Definition 2.4 (Conditional fiber distribution). *Given* μ_{k+1} *on* $\mathcal{L}_{k+1}$ *and* $\pi_{k+1\to k}$*, define the family of conditional distributions*

$$\mu_{k+1}(\cdot \mid x_k), \quad x_k \in \mathcal{L}_k,$$

on each identity fiber $\mathcal{F}_{k+1}(x_k)$*, whenever this is well defined (e.g., via disintegration).*

Definition 2.5 (Fiber entropy). *The* fiber entropy *of the projection* $\pi_{k+1\to k}$ *under* μ_{k+1} *is*

$$H_{\text{fiber}}(\mu_{k+1}, \pi_{k+1\to k}) := \mathbb{E}_{x_k\sim \mathsf{Sh}_k(\mu_{k+1})}\big[H(\mu_{k+1}(\cdot \mid x_k))\big],$$

where $H(\cdot)$ *denotes Shannon entropy.*

Proposition 2.6 (Information loss under shadow). *Let* $H(\mu)$ *denote Shannon entropy. Then*

$$H(\mu_{k+1}) = H(\mathsf{Sh}_k(\mu_{k+1})) + H_{\text{fiber}}(\mu_{k+1}, \pi_{k+1\to k}).$$

In particular,

$$H(\mu_{k+1}) \geq H(\mathsf{Sh}_k(\mu_{k+1})),$$

with equality if and only if the fibers are almost surely degenerate.

Proof. This is the standard entropy decomposition:

$$H(X, Y) = H(Y) + H(X \mid Y)$$

with $Y = \pi_{k+1\to k}(X)$ and $X \sim \mu_{k+1}$. □

Remark 2.7. H_{fiber} is the quantitative measure of identity information lost when we move from the higher layer to its shadow. It is a formal version of "the shadow is not the full self".

3 Shadow Channels and Lifts

3.1 Shadow channels

Definition 3.1 (Channel between layers). *A* channel *between two layers is a Markov kernel*

$$\mathsf{Ch}_{k\to\ell}(y_\ell \mid x_k)$$

assigning to each $x_k \in \mathcal{L}_k$ a probability distribution over $\mathcal{L}_\ell$.

Definition 3.2 (Shadow–compatible channel). *A channel $\mathsf{Ch}_{k+1\to k+1}$ is* shadow–compatible *with the projection $\pi_{k+1\to k}$ if there exists a (possibly stochastic) channel $\mathsf{Ch}_{k\to k}$ such that the following diagram commutes in law:*

$$\begin{array}{ccc} \mathcal{L}_{k+1} & \xrightarrow{\mathsf{Ch}_{k+1\to k+1}} & \mathcal{L}_{k+1} \\ \downarrow \pi_{k+1\to k} & & \downarrow \pi_{k+1\to k} \\ \mathcal{L}_k & \xrightarrow{\mathsf{Ch}_{k\to k}} & \mathcal{L}_k \end{array}$$

That is, for any μ_{k+1}, we have

$$\mathsf{Sh}_k(\mu_{k+1}\mathsf{Ch}_{k+1\to k+1}) = \mathsf{Sh}_k(\mu_{k+1})\mathsf{Ch}_{k\to k}.$$

Remark 3.3. Shadow–compatible channels are dynamics on the higher layer that induce well–defined dynamics on the lower layer. The lower–layer observer can see a consistent evolution of shadows, even without access to the full higher–layer state.

3.2 Lifts and partial identity reconstruction

Definition 3.4 (Lift). *Given a projection $\pi_{k+1\to k}$, a* lift *is a map*

$$\mathsf{Lift}_{k\to k+1} : \mathcal{L}_k \to \mathcal{P}(\mathcal{L}_{k+1}),$$

assigning to each $x_k \in \mathcal{L}_k$ a probability distribution on the fiber $\mathcal{F}_{k+1}(x_k)$.

Definition 3.5 (Deterministic vs stochastic lift). *A lift is* deterministic *if $\mathsf{Lift}_{k\to k+1}(x_k)$ is a Dirac measure for all x_k. Otherwise it is* stochastic*, encoding a distribution over possible higher–dimensional selves consistent with the shadow.*

Proposition 3.6 (No lossless deterministic lift for irreversible projection). *If $\pi_{k+1\to k}$ is irreversible on a set $A \subseteq \mathcal{L}_k$ of positive $\mathsf{Sh}_k(\mu_{k+1})$-measure, then no deterministic lift $\mathsf{Lift}_{k\to k+1}$ can reconstruct the true higher–layer distribution μ_{k+1} from $\mathsf{Sh}_k(\mu_{k+1})$ on A.*

Proof. If $\pi_{k+1\to k}$ is irreversible, there exist x_{k+1}, x'_{k+1} with $\pi_{k+1\to k}(x_{k+1}) = \pi_{k+1\to k}(x'_{k+1})$ but $x_{k+1} \neq x'_{k+1}$ and both having positive probability under μ_{k+1}. Any deterministic lift must choose a single representative in the fiber, and cannot reproduce both probabilities simultaneously. □

4 Shadow Differential and Integral

4.1 Shadow observables

Definition 4.1 (Shadow observable). *An observable on layer k is a measurable function $f_k : \mathcal{L}_k \to \mathbb{R}$. Given a projection $\pi_{k+1\to k}$, the associated* shadow observable *on layer $k+1$ is*

$$f_k^{\uparrow}(x_{k+1}) := f_k(\pi_{k+1\to k}(x_{k+1})).$$

Remark 4.2. $f_k^{\uparrow}$ only depends on x_{k+1} through its shadow. Any two higher–layer points with the same projection are indistinguishable by $f_k^{\uparrow}$.

4.2 Shadow integral

Definition 4.3 (Shadow integral). *Let μ_{k+1} be a measure on $\mathcal{L}_{k+1}$ and f_k an observable on layer k. The* shadow integral *of f_k against μ_{k+1} is*

$$\int_{\mathcal{L}_{k+1}} f_k^{\uparrow}(x_{k+1})\, d\mu_{k+1}(x_{k+1}).$$

Proposition 4.4 (Shadow integral equals lower–layer integral). *We have*

$$\int_{\mathcal{L}_{k+1}} f_k^{\uparrow}\, d\mu_{k+1} = \int_{\mathcal{L}_k} f_k(x_k)\, d\mathsf{Sh}_k(\mu_{k+1})(x_k).$$

Proof. This is the standard change–of–variables formula for pushforward measures. □

Remark 4.5. The shadow integral expresses that any observable that "lives" in the shadow layer can be evaluated either from the higher–layer distribution directly or from the shadow measure alone. This is one precise sense in which the shadow carries all information relevant to observables that do not look "above" their layer.

4.3 Shadow differential

Now suppose each layer $\mathcal{L}_k$ is a smooth manifold and $\pi_{k+1\to k}$ is a smooth submersion.

Definition 4.6 (Shadow differential). *Let $f_k : \mathcal{L}_k \to \mathbb{R}$ be smooth. Define its* shadow differential *on $\mathcal{L}_{k+1}$ as*

$$d\mathsf{Sh}f_k(x_{k+1}) := df_k(\pi_{k+1\to k}(x_{k+1})) \circ d\pi_{k+1\to k}(x_{k+1}),$$

i.e., the pullback of the usual differential via the projection.

Remark 4.7. $d\mathsf{Sh}f_k$ annihilates directions tangent to the fibers $\mathcal{F}_{k+1}(x_k)$. In other words, it is blind to "vertical" changes that do not alter the shadow.

Proposition 4.8 (Kernel of the shadow differential). *At any $x_{k+1} \in \mathcal{L}_{k+1}$, the kernel of $d\mathsf{Sh}f_k(x_{k+1})$ contains the tangent space of the fiber:*

$$T_{x_{k+1}}\mathcal{F}_{k+1}(\pi_{k+1\to k}(x_{k+1})) \subseteq \ker d\mathsf{Sh}f_k(x_{k+1}).$$

Proof. If v is tangent to the fiber, then $d\pi_{k+1\to k}(x_{k+1})[v] = 0$ and thus

$$d\mathsf{Sh}f_k(x_{k+1})[v] = df_k(\pi_{k+1\to k}(x_{k+1}))\big[d\pi_{k+1\to k}(x_{k+1})[v]\big] = 0.$$

□

5 Shadow Entropy and Invariants

5.1 Shadow entropy

Definition 5.1 (Shadow entropy)**.** *Given μ_{k+1} on $\mathcal{L}_{k+1}$, define its* shadow entropy at level k *as*

$$\mathsf{Sh}H_k(\mu_{k+1}) := H(\mathsf{Sh}_k(\mu_{k+1})),$$

the entropy of the shadow measure.

Definition 5.2 (Identity entropy)**.** *The* identity entropy between layers $k+1$ and k *is*

$$H_{\mathrm{id}}(\mu_{k+1} \mid k) := H(\mu_{k+1}) - \mathsf{Sh}H_k(\mu_{k+1}) = H_{\mathrm{fiber}}(\mu_{k+1}, \pi_{k+1\to k}).$$

Remark 5.3. H_{id} is the amount of entropy that lives purely in the fibers, i.e., in distinctions between higher–layer selves that collapse to the same shadow. It formalizes the idea of "hidden identity variation" invisible to the lower layer.

5.2 Shadow invariants

Definition 5.4 (Shadow invariant observable)**.** *An observable $I : \bigsqcup_{k\geq 0} \mathcal{L}_k \to \mathbb{R}$ is a* shadow invariant *if for any x, y in the same identity class,*

$$x \sim y \implies I(x) = I(y).$$

Proposition 5.5 (Invariants descend to all layers)**.** *If I is a shadow invariant, then for each k there exists a unique function $I_k : \mathcal{L}_k \to \mathbb{R}$ such that $I(x_k) = I_k(x_k)$ for all $x_k \in \mathcal{L}_k$, and for any x_{k+1},*

$$I_{k+1}(x_{k+1}) = I_k(\pi_{k+1\to k}(x_{k+1})).$$

Proof. Because I is constant on identity classes and each identity class intersects each layer in at most one point on a given projection chain, I is determined by its restriction to any one layer, and the restriction to layer $k+1$ must factor through the projection to layer k. □

Remark 5.6. Shadow invariants are the quantities that are truly "the same self" seen from all dimensional perspectives. In physical or informational models, these are the candidates for conserved quantities across projection and coarse–graining.

6 Chains of Shadows and Shadow Calculus

6.1 Shadow chains

Definition 6.1 (Shadow chain)**.** *A* shadow chain *is a sequence*

$$\mu_0, \mu_1, \mu_2, \ldots$$

with μ_k a probability measure on $\mathcal{L}_k$, compatible with the projections:

$$\mu_k = \mathsf{Sh}_k(\mu_{k+1}) \quad \forall k \geq 0.$$

Remark 6.2. A shadow chain is what a single identity distribution looks like when viewed from all layers simultaneously. It encodes how much of the higher–dimensional structure survives at each lower level.

6.2 Shadow derivative along a chain

Consider a parameter t describing evolution in time in a fixed layer, and a family of measures $\mu_{k+1}(t)$ with shadows $\mu_k(t)$.

Definition 6.3 (Shadow derivative)**.** *Assume $\mu_{k+1}(t)$ and $\mu_k(t)$ are differentiable in a suitable sense (e.g., in the space of measures). The* shadow derivative *at level k is*

$$\frac{d}{dt_{\mathsf{Sh}}}\mu_k(t) := \frac{d}{dt}\mathsf{Sh}_k(\mu_{k+1}(t)).$$

Proposition 6.4 (Shadow derivative is projection of higher derivative)**.** *Formally,*

$$\frac{d}{dt_{\mathsf{Sh}}}\mu_k(t) = \mathsf{Sh}_k\left(\frac{d}{dt}\mu_{k+1}(t)\right).$$

Proof. This is just linearity and continuity of the pushforward operation, when differentiation is well defined. □

Remark 6.5. The shadow derivative is how the lower–dimensional observer perceives change that originates in the higher layer. It may be a "smoothed" or coarse version of the true change when fiber dynamics cancel out in the projection.

7 Narrative Summary and Further Directions

We summarize the core picture:

- The world is represented as a tower of layers $\mathcal{L}_k$, each a different dimensional or informational resolution of identity.
- The maps $\pi_{k+1\to k}$ cast shadows down the tower, losing fiber–level detail and creating irreversibility.
- Identity is an equivalence class across layers: different points in higher and lower dimensions can belong to the same self via shared shadows.
- Measures, entropy, and channels on higher layers induce shadow measures, shadow entropy, and shadow dynamics on lower layers.
- Shadow integrals and shadow differentials formalize how observables that live in a layer "pull back" to higher layers while ignoring fiber–level variation.
- Shadow invariants are quantities that survive every projection and thus encode what it means for a self to remain the same across dimensions.

Future pieces of the Shadow Math calculus

This foundational layer can be extended in many directions:

- **Shadow channels and data–processing inequalities:** studying how information–theoretic contractions (e.g. strong data processing) behave along shadow chains.

- **Shadow Laplacians:** defining diffusion operators that act along identity fibers vs. across layers, separating “internal” vs. “projected” dynamics.
- **Shadow geometry:** giving each layer a metric or Riemannian structure and studying curvature relationships under projection and coarse–graining.
- **Shadow learning:** treating training dynamics of AI models as movement in a higher–dimensional parameter space, and shadowing them into lower–dimensional observables (loss, accuracy, emergent behaviors).

These extensions give a full *calculus of identity across dimensions*, where the lower–dimensional shadow is always a partial but consistent view of a richer self in a higher layer.