Complexity of Deep Computations via Topology of Function Spaces

Eduardo Dueñez[1]
J. Iovino[1], T. Matos-Wiederhold[2], L. Salvetti[2], F. D. Tall[2]

[1]University of Texas at San Antonio

[2]University of Toronto

ISQGD Special Session on Ergodic Theory and Topological Dynamics

22 March 2026

Stone-Weierstrass: A computational theorem?!

- **Computations** involving real numbers hinge on evaluating certain basic **continuous** operations (e.g., arithmetic) to within tolerances specified in advance.
- Basic operations are implemented by a *floating-point library*.
  - **Caveat!** Any floating point library implements a (*digital!*) algorithm which *only* approximates real-valued arithmetic on *bounded* intervals: each library only works on a specific **compact** subset of $\mathbb{R}$.

Theorem (Stone-Weierstrass —a computational view)

*Given: (**1**) a continuous map $f : V \to \mathbb{R}$ on a finite-dimensional real vector space V, (**2**) a compact subset $K \subseteq V$ and (**3**) $\varepsilon > 0$ arbitrary, there is an algorithm $\mathbf{A} = \mathbf{A}_{f,K,\varepsilon}$ which ε-uniformly approximates f on K in the sense that $f(x)$ is ε-approximately evaluated relying only on composition of arithmetic operations, applied to the coordinates of x (in some basis of V), as implemented by some fixed-precision floating-point library.*

Note: The *Universal Approximation Theorem* for neural networks is a strengthening of Stone-Weierstrass narrowing the operations the algorithm $\mathbf{A}$ is allowed to compose.

What does a choice of algorithm depend on?

Take-homes from last slide:

- An algorithm to approximate a function on a vector space V depends on the choices of:
  - a precision $\varepsilon > 0$;
  - the compactum $K \subseteq V$ on which the function is to be ε-approximated;
  - the coordinatization of V.

There are many circumstances where a function is naturally defined **only** on some subset $L \subseteq V$ of a real vector space V.

- E.g., log is defined only on $(0, +\infty) \subseteq \mathbb{R}$.
  - We wish to be able to compute such functions!

Lead-in to real-valued structures

The finite-dimensionality of V is rather a red-herring:

- Even if V is infinite-dimensional, any algorithm **A** on compacta $K \subseteq V$ –or perhaps $K \subseteq L$ for a fixed $L \subseteq V$– can only use finitely many real-valued "features" of points $x \in V$.
- Such "real features" might perhaps be coordinates —but not necessarily so:
  - **Example:** V may carry a norm $\|\cdot\|$. An algorithm (on some subset of V) may depend on the value of $\|x\|$; thus, the algorithm might take the norm as an implicitly known feature of its input argument rather than as an intermediate quantity to be computed.

We take purely relational real-valued structures from model theory as the class of domains L on which real-valued computations "make sense".

Real-valued structures and computation states

Definition (Computation States Structure (CSS))

A *Computation states structure with countably many predicates* is of the form

$$\underline{\mathbf{L}} = \langle L, P_i \rangle_{i \in \mathbb{N}},$$

where

- L is a nonempty topological space;
- P_i is a continuous real function on L (called a *structural predicate*) for each $i \in \mathbb{N}$.,

and such that the following **faithful embedding axiom** is satisfied:

*The **type map*** $\mathrm{tp} : L \to \mathbb{R}^{\mathbb{N}} : x \mapsto \big(P_i(x)\big)_{i \in \mathbb{N}}$ *is a homeomorphic embedding into the product space* $\mathbb{R}^{\mathbb{N}}$.

One may identify L with its type image $\mathrm{tp}[L] \subseteq R^{\mathbb{N}}$.

- The predicate P_i becomes the projection $x \mapsto x_i$.

Remarks on computation states structures

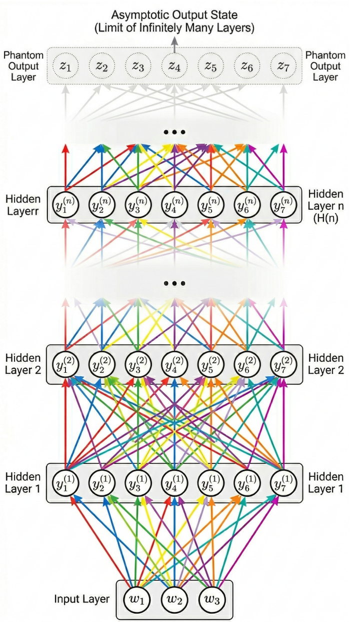
A CSS $\underline{\mathbf{L}}$ carries the topology of coordinate-wise convergence.

- A sequence/net in L, or filter on L, converges to $x \in L$ iff the pushforward of the sequence/net/filter by P_i converges to $P_i(x)$, for all i.

The model-theoretic (structural) perspective shines light on:

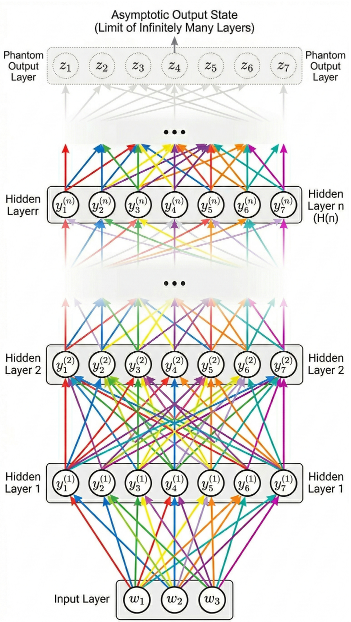
1. Theories: Relationships between predicate values,
   - such theories may be at the heart of specific algorithms;
2. Explicit definability of specific real functions on the states space,
   - definability is the model-theoretic notion capturing the existence of algorithms.

Deep Equilibria of Neural Networks



- A deep (but finite) neural network is an example of a CSS with finitely many predicates.
- An "infinitely deep" neural network is a CSS with countably many predicates.
  - Such infinitely deep networks are obtained by **"tying weights"**, i.e., by iterating a fixed map (single-layer transition) *ad infinitum*.
  - A dynamic viewpoint leads to the notion of Compositional Computations Structure (CCS).
    - *Simplest CCS:* $(\underline{\mathbf{L}}, T)$ where $\underline{\mathbf{L}}$ is a CSS and $T : L \to L$ is continuous.

Deep Equilibria of Neural Networks



- A "deep state" of $(\underline{\mathbf{L}}, T)$ is any function $\tau : L \to L$ arising as pointwise (ultra)limit of maps T^n as $n \to \infty$.
  - Some such deep state $\tau \in L^L$ often arise (only) as *ultra*limits, i.e., as *accumulation points* of $\{T^n\}_{n\in\mathbb{N}}$ in the product topology, called ultracomputations.

Theorem (Alva, D, Iovino, Walton '24)

If $\underline{\mathbf{L}}$ is realcompact and, for each fixed i, the iterates set $\{P_i \circ T^n\}_{n\in\mathbb{N}}$ is bounded on compacta $K \subseteq L$, then $(\underline{\mathbf{L}}, T)$ has deep equilibria, i.e., ***idempotent*** *ultracomputations $\widetilde{T} : L \to L$.*

Proof: Ellis-Numakura Lemma.

CCS Example: Newton-Raphson Dynamics

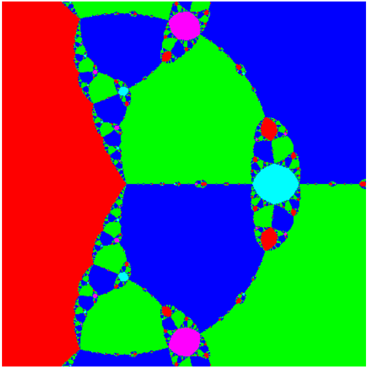
Task: Find roots of a complex polynomial $p(z)$.

- **States**: $L = \hat{\mathbb{C}} \cong S^2$ (compact).
- **Predicates**: Inverse stereographic coordinates (P_1, P_2, P_3).
- **Evolution**: Iterates of the Newton map $N_p(z) = z - \frac{p(z)}{p'(z)}$.

Deep Computations: *not* Baire-1!

A deep computation $\widetilde{T}$ "chooses" one specific **sub-sequential limit** $\widetilde{T}(z_0)$ of $\left(N_p^{(n)}(z_0)\right)_n$ for each z_0. The limit for **every** z_0 is **not** realizable sequentially globally:

- such $\widetilde{T}$ is not the pointwise limit of any subsequence of $(N_p^{(n)})_n$.



Newton Fractal for $z^3 - 2z + 2$

Visualizing Newton-Raphson Dynamics

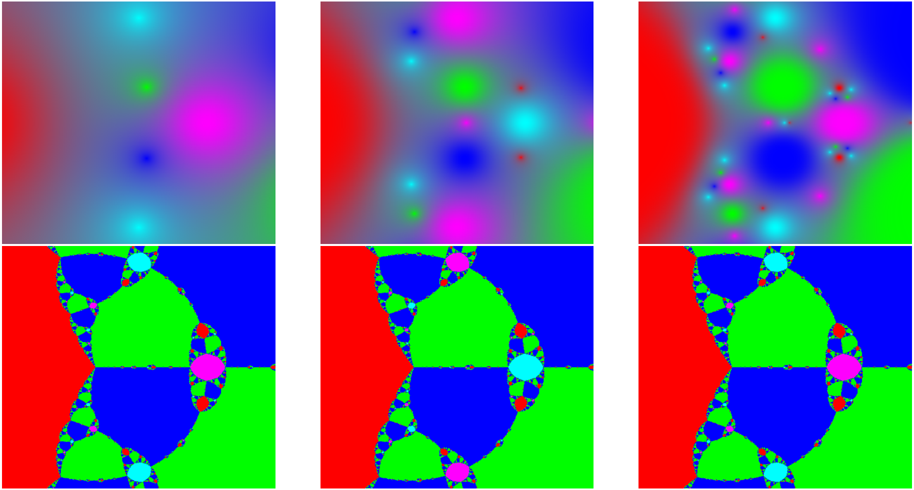


Figure: Newton iterations ($\mathbf{n = 1, 2, 3, 99, 100, 101}$). Red/Green/Blue: Convergence to roots. Magenta/Cyan: Limit 2-cycle $\{0, 1\}$. Boundaries: Julia (fractal) set.

Baire Class 1

$B_1(X, Y)$: Baire Class 1 functions $X \to Y$

Set of ("Baire-1") functions: pointwise limits of continuous $X \to Y$.

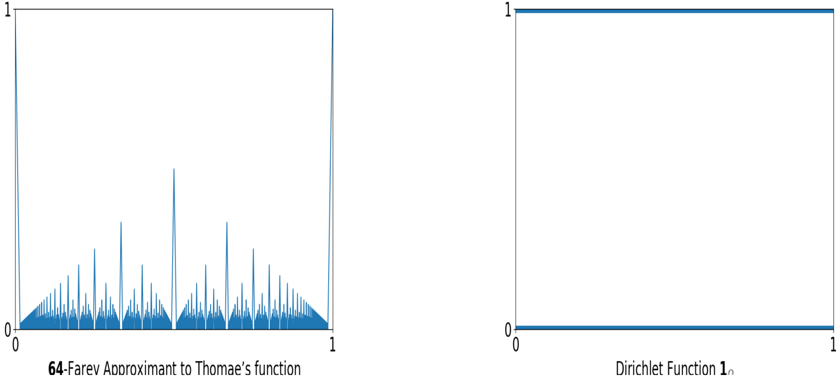


Figure: **Left:** Thomae's ("popcorn") function $x \mapsto [\![x = \frac{m}{n} \in \mathbb{Q}]\!] \cdot \frac{1}{n}$ is Baire class 1.
Right: Dirichlet's indicator function $\mathbf{1}_{\mathbb{Q}}$ is not Baire-1 (but it is Baire-2).

Baire-1 functions & NIP

Deep computations arise as limits. But how wild are these limits?

Theorem (Fundamental Dichotomy (BFT / NIP))

Let $\Delta \subseteq \mathrm{C_p}(L, L)$ be a set of continuous transformations of a realcompact CSS $\underline{\mathbf{L}}$. The following are equivalent:

1. *$\overline{\Delta}$ is relatively compact in $\mathrm{B}_1(L, L)$ (**tameness**).*
2. *Δ has the No Independence Property (**NIP**).*

If the above conditions hold, then every deep computation $\tau \in \overline{\Delta}$ is the pointwise limit of some sequence in $\overline{\Delta}$ over arbitrary compacta $K \subseteq L$.

NIP (intuition): Δ cannot not "shatter" arbitrarily large finite sets of inputs.
Discrete analogue: Finiteness of Vapnis-Chervonenkis (VC) dimension.

Ultracomputations: From tame to wild

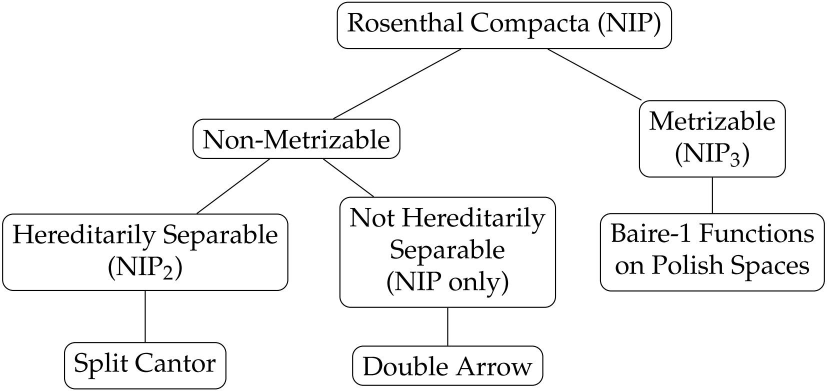
If a system satisfies NIP, it is tame. But there are *levels of tameness*.

- **Todorčević's Trichotomy** for Rosenthal Compacta (compact subsets of $B_1(X)$) implies complexity classes of ultracomputations.
  - Philosophically at least, ultracomputations *not* Baire-1 are wild: they cannot be approximated by any algorithm relying solely on floating-point arithmetic over any compactum $K \subseteq L$ witnessing the failure of NIP. —i.e., wild/non-Baire-1 ultracomputations are largely incomputable over the entire states space L in any reasonable sense.
  - A Julia set is an example of a "bad" such compactum.

Remark: Glasner-Megrelishvili's results parallel some of ours. Important technical differences include:

- Our framework applies to functions on the class of states spaces that are Tychonoff; our strongest results assume only realcompactness (rather than compact metrizability).
- We emphasize connections with the theory of real-valued computing.

The Hierarchy of Tameness



- NIP$_3$: Metrizable (most tame). Limits are unique and well-behaved.
- NIP$_2$: Hereditarily Separable but not Metrizable.
- NIP$_1$: First Countable.

The Seven Prototypes (ADK Heptachotomy)

Argyros, Dodos, and Kanellopoulos identified *7 minimal building blocks* for Rosenthal compacta. Any wilder behavior must contain one of these prototypes.

1. D_1: Alexandrov compactification (Discrete)
2. D_2: One-point compactification
3. D_3: Split Cantor
4. ...
5. D_7: Extended split Cantor

Implication for Computing

If a deep computation system satisfies NIP, its asymptotic behavior is effectively approximated by a **countable discrete set** of "archetypal" limits.

Summary

We introduce a framework to classify the complexity of asymptotic/limit real-valued computations.

Complexity	**Rosenthal Topology**	**NIP Class**
Most Tame	Metrizable	NIP_3
Intermediate	Hereditarily Separable	NIP_2
Tame	First Countable	NIP_1
Wild	Independence Property	IP

Essential Takeaway:
NIP is the dividing line between “tame” (Baire-1) transformations/functions that may be evaluated (up to a small error) computationally feasibly, and unapproximable/wild ones.

Thank You!

References

- S. Alva, E. Dueñez, J. Iovino, C. Walton. *Approximability of deep equilibria. arXiv preprint arXiv:2409.06064*, 2024. To appear in *Mathematical Structures in Computer Science*.
- Spiros A. Argyros, Pandelis Dodos, and Vassilis Kanellopoulos. A classification of separable Rosenthal compacta and its applications. *Dissertationes Mathematicae*, 449:1–55, 2008.
- E. Dueñez, J. Iovino, T. Matos-Wiederhold, L. Salvetti, F. D. Tall. Complexity of deep computations via topology of function spaces. *arXiv preprint arXiv:2601.00528*, 2026.
- Eli Glasner and Michael Megrelishvili. Todorčević' trichotomy and a hierarchy in the class of tame dynamical systems. *Transactions of the American Mathematical Society*, 375(7):4513–4548, 2022.
- Stevo Todorčević. Compact subsets of the first Baire class. *Journal of the American Mathematical Society*, 12(4):1179–1212, 1999.