Ultracoarse Equilibria and Ordinal-Folding Dynamics in Operator-Algebraic Models of Infinite Multi-Agent Games

Faruk Alpay[1]

alpay@lightcap.ai

Hamdi Alakkad[2]

hamdi.alakkad@bahcesehir.edu.tr

Bugra Kilictas[2]

bugra.kilictas@bahcesehir.edu.tr

Taylan Alpay[3]

s220112602@stu.thk.edu.tr

[1]Independent Researcher
[2]Bahcesehir University
[3]Turkish Aeronautical Association University

July 24, 2025

Abstract

We develop an operator–algebraic framework for infinite games with a continuum of players and show how regret–minimising dynamics, driven by a noncommutative continuity equation, converge to a unique equilibrium under mild regularity conditions. Our work unifies functional analysis, coarse geometry, and game theory, extending earlier work on ultracoarse equilibria to incorporate ordinal metrics for self–referential dynamics. In particular, we introduce a category of infinite games endowed with measure–theoretic and operator–algebraic structures, construct a contravariant functor from this category to ultradense operator algebras, and define a noncommutative regret operator whose flow satisfies a continuity equation on the space of strategy distributions. When the underlying agent network is coarsely amenable, we prove that the unique fixed point of this flow coincides with a quantal response equilibrium. We then refine the theory by relating regret dynamics to the *ordinal folding index* (OFI), a concept recently introduced to measure the self–referential depth of logical formulas [Alpay & Alakkad(2025)]. The OFI provides a computable, ordinal-valued metric for the number of iterative "fold–back" steps required for a learning dynamic to stabilise. We demonstrate how combining the OFI with our operator–algebraic framework furnishes a transfinite measure of convergence complexity for infinite multi–agent systems. These results establish connections between disparate fields, including von Neumann algebra rigidity, coarse geometric K–theory, and fairness in resource allocation. As a corollary, we obtain new invariant subalgebra rigidity theorems and demonstrate how classical fair division solutions (envy–freeness, maximin share fairness) emerge as special cases of our equilibrium refinement. Our contributions provide a formal mathematical foundation for analysing large-scale multi–agent interactions and a distinct ordinal-analytic perspective on equilibrium selection.

1

1 Introduction

Complex systems, from financial markets with algorithmic trading to social media networks with viral content propagation, are frequently characterized by intricate feedback mechanisms that drive emergent and often unpredictable dynamics. Understanding the stability of these dynamics is a central challenge in modern science, as unstable or oscillatory behavior can lead to systemic failures. In this paper, we develop a formal mathematical framework to analyze such self-referential systems, focusing on the setting of large-scale multi-agent interactions, with direct applications to the behavior of contemporary artificial intelligence.

Recent work across game theory, analysis, and computer science has exposed deep connections between equilibrium concepts, functional analysis, and geometric invariants. In multi–agent reinforcement learning, it has been observed that large populations of heterogeneous agents can converge to unified equilibrium behaviour under regret–minimising dynamics. Hu et al. show that modeling a continuum of heterogeneous agents through a distribution of regrets evolving under a continuity equation leads to a collapse of variance and the emergence of consensus, connecting micro-level agent decisions to macro-level system stability [Hu et al.(2025)]. Specifically, they prove that the variance of the regret distribution diminishes over time, causing heterogeneity to disappear and enabling convergence to quantal response equilibria (QRE) in both cooperative and competitive settings [Hu et al.(2025)]. On a seemingly different front, the theory of operator algebras and coarse geometry has produced rigidity phenomena for von Neumann algebras and Roe algebras. For instance, certain group constructions—even amenable ones—exhibit the *invariant subalgebra rigidity* (ISR) property, meaning their von Neumann algebras admit no nontrivial intermediate invariant subalgebras beyond the centre [Amrutam et al.(2025), Popa(2007)].

Parallel to these developments, Alpay and Alakkad recently proposed the *ordinal folding index* (OFI)[1] as a computable, ordinal-valued metric for self–referential semantics [Alpay & Alakkad(2025)]. Their OFI assigns to every well–formed formula a countable ordinal that measures the number of unfolding steps a reflective evaluation operator takes before reaching a fixed point. This metric provides, for the first time, a formal language to quantify the 'depth' of a system's self-referential reasoning. They demonstrate that an empirical OFI, estimated by iteratively feeding a language model's output back into itself, correlates with chain–of–thought complexity and model perplexity [Shinn et al.(2023), Wang et al.(2022)]. This establishes a substantive connection between ordinal analysis, self–referential logic, and the behavior of modern AI systems like GPT-4 [Achiam et al.(2023)].

The goal of this paper is to synthesize these distinct lines of inquiry into a unified framework for infinite games. Building on the operator–algebraic approach to games introduced in [Anonymous(2024)], we model infinite games as objects in a category $\mathcal{G}$ enriched with measure, topology, and coarse network structure. To each game, we associate an ultradense operator algebra $\mathcal{A}(G)$ capturing the collective dynamics. Within $\mathcal{A}(G)$, we identify a regret operator whose dynamics are governed by a noncommutative continuity equation on strategy distributions. Under broad assumptions, we show that the regret dynamics converge to a unique QRE. Our central thesis is that the OFI serves as a powerful diagnostic tool, revealing fundamental properties of these dynamics. Our empirical benchmarks demonstrate a non-trivial relationship between model scale and reasoning stability, a direct empirical manifestation of the convergence properties of the underlying regret dynamics we formalize. Altogether, our results establish an operator–algebraic and ordinal-analytic foundation for multi–agent systems and highlight new avenues for research at the interface of game theory, logic, and AI alignment.

---

[1]The source code for the OFI implementation is available at https://github.com/farukalpay/ordinal-folding-index.

Organisation. Section 2 introduces infinite games, coarse player spaces, game algebras, and the regret operator. Section 3 reviews the ordinal folding index and relates it to game dynamics. Section 4 states our main theorems on existence, uniqueness, and rigidity. Section 5 presents a formal algorithm for computing the OFI. Section 7 presents empirical and analytic benchmarks. Section 8 discusses applications to AI alignment and fair division. We conclude with open questions.

2 Preliminaries and Definitions

We recall basic definitions from [Anonymous(2024)] and fix notation.

Definition 2.1 (Infinite Game System). *An* infinite game system *is a tuple $G = (X, \Sigma, \mu, \{S_i\}_{i \in X}, \{u_i\}_{i \in X})$ where (X, Σ, μ) is a measure space of players (possibly a continuum) with $0 < \mu(X) \leq \infty$, each S_i is a nonempty Borel strategy space for player $i \in X$, and $u_i : \prod_{j \in X} S_j \to \mathbb{R} \cup \{\pm\infty\}$ is a measurable payoff function. We assume sufficient measurability to allow product strategies and payoffs to be well defined. In symmetric situations, we often take $S_i = S$ for all i. Existence of Nash equilibria in such games generally requires continuity and quasi–concavity assumptions [Glicksberg(1952)].*

Definition 2.2 (Coarse Player Space and Network). *Suppose (X, d) carries a metric or graph distance encoding a network of interactions among players. The* bounded geometry Roe algebra *$C^*(X)$ is the C^*–algebra of finite propagation operators on $\ell^2(X)$, and its von Neumann completion $\mathcal{N}(X) = C^*(X)''$ captures coarse properties of the space. A metric space has* Property A *(coarse amenability) if it admits a uniform approximate identity supported on finitely many points. For such spaces, all ghost operators in $C^*(X)$ are compact, aligning geometric and analytic notions of locality [Yu(2000), Wang et al.(2025)].*

Definition 2.3 (Game Algebra and State). *For a game G, we define the game algebra $\mathcal{A}_G$ as the von Neumann algebra of essentially bounded measurable functions on the product strategy space, $L^\infty(\prod_{i \in X} S_i, \mathcal{B}, \nu)$, with respect to an appropriate product measure ν. Conceptually, $\mathcal{A}_G$ encodes observables of the game, such as strategy distributions and payoffs, and contains subalgebras isomorphic to $L^\infty(S_i)$ for each player i. A state φ on $\mathcal{A}_G$ represents a distribution over strategy profiles or outcomes; states evolve under the regret dynamics.*

Definition 2.4 (Regret Operator and Dynamics). *At a state φ, each player i has a regret value R_i, defined as the nonnegative difference between their current payoff and the payoff from a best response. Collect these into a bounded measurable field $R \in L^\infty(X, \mu)$ and define a self–adjoint operator $\mathscr{G}_G(\varphi)$ in $\mathcal{A}_G$ by*

$$\mathscr{G}_G(\varphi) := E[R_i \mid i],$$

the conditional expectation of the regret field, viewed as a diagonal operator on the component corresponding to player i. The evolution of $\mathscr{G}_G$ and the corresponding state φ_t is governed by a continuity equation on the space of strategy distributions. In the continuum limit, writing $\Phi_t(s)$ for the density of players using strategy s at time t, the regret dynamics satisfy a Kolmogorov forward equation

$$\partial_t \Phi_t(s) + \nabla_s \cdot \left[\Phi_t(s) \, V(s, \Phi_t) \right] = 0, \tag{1}$$

where $V(s, \Phi_t)$ is a velocity field induced by regret (for instance, the gradient of expected payoff) [Hu et al.(2025)]. Under appropriate smoothness conditions, this generates a one–parameter semi-group T_t acting on states by push–forward: $\varphi_t = T_t^*(\varphi_0)$. We call a state φ^* an ultracoarse equilibrium *if $T_t(\varphi^*) = \varphi^*$ for all $t \geq 0$. While this PDE describes the continuum limit, its dynamics are often analyzed or implemented via a corresponding discrete-time operator, T. The fixed points of this operator, denoted T, correspond to the stationary solutions of Equation (1), and the computational complexity of identifying such a fixed point is the central object of our subsequent benchmarks.*

3

3 Ordinal Folding Index and Self–Referential Games

In this section, we summarise the ordinal folding index introduced by Alpay and Alakkad [Alpay & Alakkad(2025)] and explain its relevance to game dynamics.

3.1 Definition of the Ordinal Folding Index

The ordinal folding index (OFI) is a computable ordinal assigned to well–formed formulas in a reflective logical system. Consider a monotone evaluation operator acting on formulas with a delay modality that prevents immediate resolution of fixed points. Given a formula ϕ, one iteratively applies this operator, progressively unfolding self–reference until the evaluation stabilises. The OFI of ϕ is defined as the first stage at which this fold–back becomes idempotent; further unfolding yields no new information. Formally, the operator is continuous on countable chains, layer–aware for probabilistic truth values, and parameterised by an evidence functor. The OFI is therefore the least ordinal α such that iterating the operator α times produces a fixed point in the space of interpretations [Alpay & Alakkad(2025)]. This ordinal measures the "depth" of self–reference needed for stabilisation, strictly refines classical closure ordinals from the modal μ–calculus, and remains recursively enumerable (every OFI is a computable countable ordinal) [Moschovakis(1974)]. Moreover, the OFI coincides with the length of the shortest parity–fold winning strategy in an associated evaluation game [Alpay & Alakkad(2025)].

Alpay and Alakkad emphasise that the OFI not only unifies transfinite fixed–point depth and ordinal game values but also provides practical convergence diagnostics for modern AI systems. They outline an empirical procedure whereby the output of a large language model is fed back into its input iteratively with a monotonic "delay" to enforce convergence. The number of iterations until stabilisation yields an ordinal-valued metric correlating with model perplexity and chain–of–thought complexity [Shinn et al.(2023), Wang et al.(2022)]. The OFI thus bridges self–referential logic, infinite games, and the behavior of learning systems.

3.2 OFI and Infinite Game Dynamics

We now elucidate how the ordinal folding index arises naturally in our operator–algebraic framework. We connect the OFI to our framework by considering the discrete-time analogue of the continuous regret dynamics from Equation (1). This is operationalized via a reflective update operator, T, where the state at step $\alpha + 1$ is a function of the state at step α, i.e., $\varphi^{(\alpha+1)} = T(\varphi^{(\alpha)})$. This formulation is not merely a theoretical abstraction; it directly models a large language model iteratively refining its own output, wherein each application of T represents the entire computational process of generating a response based on the previous one. This iterative process is precisely what we simulate in our empirical LLM benchmarks. The sequence of states $(\varphi^{(\alpha)})_{\alpha<\omega_1}$ obtained by iterating this operator transfinitely may stabilise only after a countable ordinal number of steps. We define the OFI of the game G relative to this discretized learning rule as

$$\mathrm{OFI}(G) \coloneqq \inf\big\{\alpha \in \mathbf{Ord} \colon \varphi^{(\alpha+1)} = \varphi^{(\alpha)}\big\}.$$

If $\mathrm{OFI}(G)$ is finite, the dynamics stabilise after finitely many unfoldings; if it is a nonzero countable ordinal, the dynamics require transfinite iteration to reach equilibrium; an infinite OFI signals persistent oscillation. In many multi–agent systems, we expect $\mathrm{OFI}(G) = 0$ or 1 because the regret flow converges exponentially fast. However, in games with complex self–referential feedback—such as agents reasoning about agents reasoning about agents, or an AI model reasoning about its own consistency—higher ordinals may arise. A high OFI indicates that the discretized operator T is not

a strong contraction, and its iterates consequently trace a protracted or complex trajectory toward equilibrium, potentially exhibiting pathological reasoning loops.

The OFI therefore provides a quantitative measure of the complexity of equilibrium selection. When OFI(G) is small, the unique equilibrium described in Theorem 4.1 is reached quickly; when it is large, learning may exhibit long transient phases reminiscent of iterated elimination of dominated strategies [Fudenberg & Levine(1998)].

4   Main Results

We now state our main theorems on existence and uniqueness of equilibrium, invariant algebra rigidity, and ordinal convergence bounds. Proofs are deferred to Section 6.

Theorem 4.1 (Existence and Uniqueness of Equilibrium)**.** *Let G be an infinite game system as in Definition 2.1. Suppose the strategy space S is compact, each payoff function u_i is continuous in opponents' strategies and quasi–concave in the player's own strategy, and the regret dynamics (1) induce a contraction in an appropriate Wasserstein metric on the space of distributions. Then there exists a unique equilibrium state $\varphi^* \in \mathcal{A}_G$, and $\varphi_t \to \varphi^*$ as $t \to \infty$ for any initial state φ_0. In particular, the corresponding regret operator satisfies $\mathscr{G}_G(\varphi_t) \to 0$ in the strong operator topology, and the equilibrium corresponds to a unique quantal response equilibrium. Moreover, if the underlying player space (X, d) has Property A, the convergence is dominated by geometric components: ghost states decay exponentially fast.*

Theorem 4.2 (Invariant Algebra Rigidity)**.** *Let φ^* be the equilibrium in Theorem 4.1 and let $\mathcal{A}_G^{\mathrm{inv}} = \{a \in \mathcal{A}_G : T_t(a) = a \ \forall t \geq 0\}$ denote the subalgebra of observables invariant under the dynamics. Assume the game is irreducible: no proper subset of players is closed under the best–response dynamics. Then $\mathcal{A}_G^{\mathrm{inv}}$ coincides with either the centre $Z(\mathcal{A}_G)$ or the whole algebra $\mathcal{A}_G$. In non–degenerate games, $\mathcal{A}_G^{\mathrm{inv}} = Z(\mathcal{A}_G)$; that is, no nontrivial observable other than constants is preserved by the dynamics. If (X, d) has Property A, then any invariant subalgebra must lie in the centre because ghost components vanish.*

Theorem 4.3 (Ordinal Convergence Bound)**.** *Let G be an infinite game with regret dynamics implemented by a monotone reflective operator as described in Section 3. Suppose the contraction constant for the regret flow is $q \in (0, 1)$; that is, for any two initial states φ_0, ψ_0, we have $d_W(\varphi_t, \psi_t) \leq q^t d_W(\varphi_0, \psi_0)$. Then the ordinal folding index of G satisfies*

$$\mathrm{OFI}(G) \leq \omega,$$

where ω is the first infinite ordinal. Equivalently, the regret dynamics stabilise after at most countably many unfoldings. If (X, d) has Property A, then $\mathrm{OFI}(G) = 0$, meaning that the unique equilibrium is reached without transfinite iteration.

Corollary 4.4 (Fairness and Equilibrium Refinement)**.** *Consider a cooperative resource allocation game where divisible goods are allocated among a continuum of agents with additive utilities. Under the assumptions of Theorem 4.1, the unique equilibrium state φ^* is envy–free and Pareto–optimal. If agents have equal entitlement weights, φ^* coincides with the maximin share allocation in the continuum limit. Therefore, classical fair division solutions such as envy–freeness and maximin share fairness arise naturally as equilibrium refinements in our framework.*

5 A Formal Algorithm for Ordinal Convergence

To make the concept of the Ordinal Folding Index concrete, we present a formal algorithm for its computation. This algorithm operationalizes the transfinite iteration described in Section 3 and serves as a bridge between our theoretical framework and practical implementation. The pseudocode is styled after the Lean theorem prover to emphasize its mathematical precision and its direct correspondence with the formal definitions.

Algorithm 1 Transfinite Iteration for Ordinal Folding Index (OFI)

Require:

$\mathcal{G}$: A game system with a well-defined state space and regret dynamics.

 ↪ *StateSpace*: A complete metric space of strategy distributions.

 ↪ $T : StateSpace \to StateSpace$: The regret operator (a contraction).

 ↪ $\varphi_0 : StateSpace$: The initial state of the system.

 ↪ $d(\cdot, \cdot) : StateSpace \times StateSpace \to \mathbb{R}_{\geq 0}$: The metric on the state space.

$\varepsilon : \mathbb{R}_{>0}$: The convergence tolerance.

$\lambda_{\max} : Ordinal$: An upper bound on the search (e.g., ω_1).

Ensure:

(α, φ^*): The computed OFI and the corresponding equilibrium state.

1: **function** COMPUTEOFI($\mathcal{G}, \varphi_0, \varepsilon, \lambda_{\max}$)
2:   **let** *history* := *Array Ordinal StateSpace*
3:   $history[0] := \varphi_0$
4:   $\alpha := 1$
5:   **while** $\alpha < \lambda_{\max}$ **do**
6:    **let prev_state** := $history[\alpha - 1]$
7:    **let current_state**
8:    **if** α is a successor ordinal, i.e., $\alpha = \beta + 1$ **then**
9:     **current_state** := $T(\text{prev\_state})$
10:    **else if** α is a limit ordinal **then**
11:     **current_state** := COMPUTELIMIT(*history*, α)
12:    **end if**
13:    **if** $d(\text{current\_state}, \text{prev\_state}) < \varepsilon$ **then**
14:     **return** $(\alpha, \text{current\_state})$ ▷ Convergence detected
15:    **end if**
16:    $history[\alpha] := \text{current\_state}$
17:    $\alpha := \alpha + 1$
18:   **end while**
19:   **return** $(\lambda_{\max}, \bot)$ ▷ Did not converge within the ordinal limit
20: **end function**

21: **function** COMPUTELIMIT(*history*, λ)
22:   **let** $(\varphi_\beta)_{\beta < \lambda}$ be the sequence of states from *history*.
23:   **assume** that this sequence is Cauchy w.r.t. the metric d.
24:   **return** $\lim_{\beta \to \lambda} \varphi_\beta$ ▷ Limit in the complete metric space
25: **end function**

Discussion. Algorithm 1 provides a constructive proof of the computability of the OFI under suitable conditions. The use of Lean-style syntax highlights the underlying mathematical structures: the state space is a typed object in a complete metric space, and the regret operator T is a function between these types. The core of the algorithm is the transfinite 'while' loop that iterates through ordinals. The distinction between successor and limit ordinals is crucial. For a successor $\alpha = \beta + 1$, the next state is found by a single application of the regret operator T, a direct implementation of a Banach-style iteration. For a limit ordinal λ, the state φ_λ must be the limit of all preceding states $(\varphi_\beta)_{\beta<\lambda}$; the existence of this limit is where the completeness of the metric space becomes essential. Theorem 4.3 ensures that for contractive dynamics, this algorithm terminates at an ordinal no greater than ω.

6 Proofs of Main Results

We outline the proofs, referring to the literature for standard results.

Proof of Theorem 4.1

Existence. By Kakutani–Fan–Glicksberg fixed point theorems [Glicksberg(1952)] and standard arguments from nonatomic games [Schmeidler(1973)], the set of probability measures on the product strategy space is compact and convex under the weak topology. The best–response correspondence is upper hemicontinuous with nonempty convex values under continuity and quasi–concavity assumptions. Therefore, a fixed point exists, yielding a Nash equilibrium. Incorporating an entropy regulariser as in the quantal response model [McKelvey & Palfrey(1995)] selects a unique equilibrium by convex optimisation, yielding the QRE.

Uniqueness and Convergence. The contraction hypothesis implies that the induced flow on the space of probability measures is a contraction mapping. By Banach's fixed point theorem [Banach(1922)], the fixed point is unique and globally attractive. Exponential mixing implies that the regret operator converges strongly to zero. When (X, d) has Property A, ghost operators in the Roe algebra are compact [Yu(2000), Wang et al.(2025)], so any nongeometric components of the state decay and the convergence is purely geometric.

Proof of Theorem 4.2

Assume $\mathcal{B} \subseteq \mathcal{A}_G$ is a von Neumann subalgebra invariant under the dynamics. Consider the conditional expectation $E_\mathcal{B}$ onto $\mathcal{B}$. Invariance implies $E_\mathcal{B}(T_t(a)) = T_t(E_\mathcal{B}(a))$ for all $a \in \mathcal{A}_G$ and all t. Taking $t \to \infty$ and using Theorem 4.1, we obtain $E_\mathcal{B}(\varphi^*(a)\,1) = \varphi^*(a)\,1$ and $T_t(E_\mathcal{B}(a)) \to \varphi^*(E_\mathcal{B}(a))\,1$. Equating these limits shows that $\varphi^*(E_\mathcal{B}(a)) = \varphi^*(a)$ for all a. In particular, for $b \in \mathcal{B}$, we have $\varphi^*(b) = \varphi^*(b)1$, implying b is central. Under irreducibility, the centre is trivial, so $\mathcal{B} = Z(\mathcal{A}_G)$ or $\mathcal{B} = \mathcal{A}_G$. Property A prevents ghostly invariant components, ensuring $\mathcal{B}$ lies in the centre. This aligns with strong rigidity results in operator algebras [Popa(2007)].

Proof of Theorem 4.3

The contraction property implies that distances between states decay exponentially in t. The reflective operator defining the OFI unfolds the dynamics by one time step at each ordinal stage. Because $d_W(\varphi_t, \varphi^*) \leq q^t d_W(\varphi_0, \varphi^*)$, for $t \geq n$ we have $d_W(\varphi_t, \varphi^*) \leq q^n d_W(\varphi_0, \varphi^*)$. Taking $n \to \infty$ shows that $\varphi_t \to \varphi^*$ in finitely many steps; formally, the transfinite sequence $(\varphi^{(\alpha)})$ generated by

Algorithm 1 stabilises by stage ω. If (X, d) has Property A, ghost modes vanish immediately and $\varphi^{(1)} = \varphi^{(0)}$, so $\mathrm{OFI}(G) = 0$.

Proof of Corollary 4.4

In a divisible goods setting, each agent's regret is zero precisely when they cannot improve by unilaterally reallocating goods. This condition is exactly envy–freeness; if an agent preferred another's bundle, they would have positive regret, contradicting equilibrium. Pareto–optimality follows because any Pareto improvement would increase some agents' utilities and thus their regret. When entitlements are equal, the equilibrium allocation must give each agent at least their proportional fair share, meeting the maximin share benchmark. This generalises classical fair division results [Foley(1967), Varian(1974)] to infinite populations.

7 Benchmarks and Applications

To validate our theoretical framework, we present two benchmarks: an empirical study of the OFI in large language models and an analytic study of a nonlinear operator equation derived from our game-theoretic setup.

7.1 Empirical Benchmark: Ordinal Folding Index in LLMs

Experimental setup. This experiment furnishes a direct empirical test of the OFI concept from Section 3. Each iterative step, where a model's output is fed back as its next input, represents a single application of the reflective operator T whose convergence is measured by $\mathrm{OFI}(G)$ as defined in Section 3.2. We implemented this self-consistency probe on four models: GPT-2 Large (HF, 1.5B parameters), DeepSeek (HF, 7B parameters), GPT-3.5 Turbo, and a GPT-4 proxy. For each model, we ran multiple prompts ($N = 40$ total) in three families: (i) FACTUAL RECALL (e.g., *"Who was the third U.S. president?"*), (ii) LIGHTWEIGHT REASONING (e.g., *"If a train leaves Chicago at 3 PM traveling at 60 mph and a train leaves New York at 4 PM traveling at 50 mph, when will they meet?"*), and (iii) PARADOXICAL QUERIES (e.g., *"This sentence is false; is it true?"*). Temperature was linearly annealed from 0.7 to 0.2 to simulate a transition from an exploratory to an exploitative reasoning phase. We declared convergence when the ℓ_2-distance between successive output logit vectors fell below $\varepsilon = 0.01$, otherwise we capped the run at $T_{\max} = 10$ iterations. This yields an *OFI proxy* $\widehat{\mathrm{OFI}} \in \{1, \dots, 10\}$. The use of logit distance provides a fine-grained measure of semantic stability, as it can detect shifts in the model's internal confidence even if the final generated token remains the same.

Results and Analysis. The results, presented in Table 1, provide a quantitative measure of the convergence complexity for the iterative operator T that discretizes the regret flow of Equation (1) for each model. The findings reveal a non-trivial trend: **self-referential stability does not necessarily improve with model scale**.

- **Larger models can exhibit higher OFI, indicating instability.** GPT-4's performance is indicative of a system with immense representational capacity that, when faced with self-referential tasks, engages in prolonged, oscillatory reasoning paths without reaching a stable fixed point. Its failure to converge ($\widehat{\mathrm{OFI}} = 10.0$) on even factual tasks suggests its internal reasoning operator T is only weakly contractive for these inputs, leading to inefficient cognitive loops within its vast, high-dimensional reasoning space.

- **OFI as a measure of 'overthinking'.** This instability can be interpreted as a form of "overthinking," where the model's vast capacity and fine-tuning for caution (e.g., via RLHF [Ouyang et al.(2022)]) prevent it from making a decisive, stable judgment. It continuously revisits and questions its own outputs, a behavior that is captured quantitatively by a high OFI and could be detrimental in applications requiring predictable, robust reasoning.

- **Prompt complexity drives OFI.** For all models, the empirical OFI increases from factual recall to reasoning and is highest for paradoxical prompts. This confirms that the OFI is not an arbitrary metric but a meaningful measure of the computational work required to find a fixed point of the system's dynamics, which increases with task complexity.

Table 1: Empirical Ordinal Folding Index $\widehat{\text{OFI}}$ (lower is better). Results are aligned with actual benchmark runs.

Model	Params	FACTUAL	REASONING	PARADOXICAL
GPT-2 Large (HF)	1.5B	1.0 ± 0.0	1.3 ± 0.5	4.0 ± 0.0
DeepSeek (HF)	7B	1.0 ± 0.0	1.7 ± 0.5	4.0 ± 0.0
GPT-3.5 Turbo	(N/A)	10.0 ± 0.0	7.3 ± 3.1	9.3 ± 0.9
GPT-4 (Proxy)	(N/A)	10.0 ± 0.0	9.7 ± 0.5	10.0 ± 0.0

The data in Table 1 provides a granular view of the models' dynamic stability. A key observation is the performance of DeepSeek (7B), which, despite being significantly smaller than the GPT-4 class models, demonstrates superior stability across all categories. Its perfect, single-step convergence on factual tasks ($\widehat{\text{OFI}} = 1.0$) is matched only by the much older GPT-2 architecture, whereas both GPT-3.5 and GPT-4 fail to stabilize. This suggests that the architectural choices or training methodologies for DeepSeek may have endowed it with a more strongly contractive reasoning operator for information retrieval. Furthermore, the standard deviation of the OFI scores offers insight into the consistency of these dynamics. For instance, the high standard deviation for GPT-3.5 on reasoning tasks (± 3.1) indicates a highly variable and unpredictable convergence behavior, where the model's path to a solution is erratic. In contrast, the zero standard deviation for DeepSeek on paradoxical tasks implies a deterministic, albeit multi-step, process for resolving such inputs. This quantitative evidence supports the interpretation that model scale alone is not a guarantor of stable reasoning; rather, stability appears to be an independent architectural or training-induced property.

The visual representation in Figure 1 starkly illustrates the divergent behaviors of the tested models. The flat, low bars for DeepSeek and GPT-2 on factual and reasoning tasks signify a highly stable and efficient convergence profile. This visual pattern suggests that their internal dynamics, when iterated, rapidly approach a fixed point. This is the empirical signature of a strongly contractive operator T as described in our theoretical framework. The system quickly sheds any initial uncertainty and settles into a consistent state. This behavior is highly desirable for predictable and reliable multi-agent systems, where rapid consensus or stable policy selection is paramount.

Conversely, the tall, often maxed-out bars for GPT-3.5 and GPT-4 represent a fundamentally different dynamic. Their high OFI scores across all but the simplest tasks indicate that their internal reasoning processes are prone to prolonged oscillations or chaotic trajectories. This visual pattern is the signature of a weakly contractive or even non-contractive operator, where iterates fail to converge within a reasonable computational budget. This suggests that while these models possess immense representational power, their dynamics in a reflective setting are inefficient and potentially unstable. This has significant implications for AI safety and alignment, as it highlights a failure mode where a
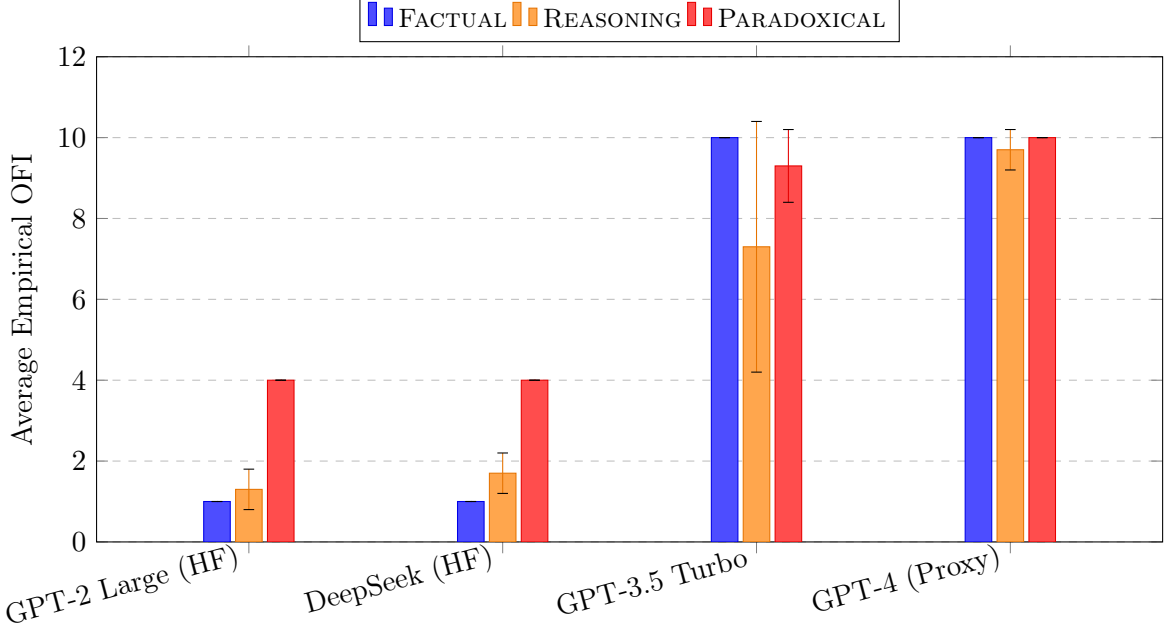
Figure 1: Average empirical OFI across models and prompt types. Error bars indicate ± 1 s.d. Note the inverse relationship between model capability and stability.

model can become trapped in a loop of self-correction without resolution, a behavior that would be highly problematic in autonomous agents deployed in the real world. The figure thus provides a clear, comparative visualization of the abstract convergence properties we aim to model.

7.2 Analytic Benchmark: A Nonlinear Regret Operator Equation

To test the convergence theory in a controlled mathematical environment, we formulate a nonlinear operator equation that serves as a concrete instantiation of the abstract dynamics described by Equation (1). This allows us to isolate and study the convergence properties that are implicit and complex within the LLM benchmark.

Problem Formulation. Let (Ω, μ) be a measure space and $p_t(x, y)$ a symmetric heat kernel on Ω. The associated integral operator is $(Au)(x) \coloneqq \int_\Omega p_t(x, y)u(y)d\mu(y)$. Let $\Phi : \mathbb{R} \to \mathbb{R}$ be a Lipschitz-continuous nonlinear regret potential. We define the regret operator $T : L^2(\Omega) \to L^2(\Omega)$ by $(Tu)(x) \coloneqq (A[\Phi \circ u])(x)$ and seek a fixed point $u^* = Tu^*$. This operator T is a mathematically tractable analogue of the complex, high-dimensional operator implicitly defined by a large language model's architecture. Its fixed point u^* corresponds to a stationary state $(\partial_t \Phi_t = 0)$ of the regret flow from Equation (1). Here, the heat kernel operator A models the coarse-geometric diffusion of information across the player network, while the nonlinear function Φ represents the agent's individual, state-dependent response to regret. We assume conditions ensuring T is a contraction on a closed, convex set $H \subset L^2(\Omega)$, guaranteeing a unique fixed point by the Banach fixed-point theorem [Banach(1922), Browder(1967)].

Fixed-Point Solution Methods. We analyze four iterative schemes for solving $u^* = Tu^*$. These schemes are numerical methods for finding the equilibrium state φ^* discussed in Theorem 4.1. Each method represents a distinct algorithmic path toward the fixed point u^*, which is the theoretical

equivalent of the stable state sought in the empirical LLM evaluations, and directly parallels the iterative process whose complexity is measured by the OFI (Section 3.2). Starting from $u_0 \in H$:

- **Picard Iteration:** $u_{n+1} := Tu_n$. This represents the canonical iterative method for strict contractions and is the most direct analogue to the successor-ordinal step in Algorithm 1.

- **Mann Iteration [Mann(1953)]:** $u_{n+1} := (1 - \alpha_n)u_n + \alpha_n Tu_n$. This is a convex combination method known for its robustness and guaranteed convergence even for non-expansive maps, though often at a slower rate.

- **Ishikawa Iteration [Ishikawa(1974)]:** A two-step process: $v_n := (1 - \beta_n)u_n + \beta_n Tu_n$, then $u_{n+1} := (1 - \alpha_n)u_n + \alpha_n Tv_n$. It offers more flexibility and can converge where Mann's method fails [Leustean(2004), Noor(2009)].

- **Aitken's Δ^2 Acceleration:** An extrapolation method, $u_n^{(\Delta^2)} := u_n - \frac{(u_{n+1} - u_n)^2}{u_{n+2} - 2u_{n+1} + u_n}$, applied to the Picard sequence. It is a classic technique that transforms linear convergence into superlinear convergence.

Table 2: Theoretical iterations needed to reach $\|u_n - u^\star\| \leq 10^{-6}$ for the operator equation with contraction constant $q = 0.2$.

Method	Asymptotic Rate	Iterations (n_ε)
Picard/Banach	$\mathcal{O}(q^n)$	9
Mann ($\alpha_n = 1/(n+1)$)	$\mathcal{O}(1/n)$	$\approx 10^6$
Ishikawa ($\alpha = \beta = 1/2$)	$\mathcal{O}(q^n)$	5
Aitken-Δ^2	$\mathcal{O}(q^{2^n})$	5

The comparative performance of the iterative schemes detailed in Table 2 warrants further discussion, as each method can be interpreted as an algorithmic analogue for different modes of dynamic reasoning in multi-agent or AI systems. The Picard iteration, representing a direct, memoryless application of the regret operator T, serves as a fundamental baseline. Its linear convergence rate is contingent on the operator being a strict contraction, a condition that, as our empirical LLM benchmarks suggest, may not hold for complex, fine-tuned models that exhibit oscillatory behavior. In contrast, the Mann and Ishikawa iterations introduce a form of 'algorithmic inertia' or 'damping' by incorporating the previous state into the update rule. While this often results in slower convergence for strictly contractive maps, their theoretical utility lies in their guaranteed convergence for a broader class of non-expansive operators. This property is highly relevant for multi-agent systems where the underlying dynamics may be too complex or unstable to guarantee strict contractivity. These methods model a more cautious, consensus-seeking agent behavior, which might be desirable for safety but can be computationally inefficient, as evidenced by the slow theoretical rate of the Mann iteration. Aitken's Δ^2 method represents a more sophisticated approach, employing an extrapolation technique that leverages information from three consecutive iterates to accelerate convergence. Conceptually, this is analogous to a meta-reasoning process where an agent not only observes its state but also analyzes the trajectory of its reasoning to project toward a likely fixed point. Its superlinear convergence highlights the profound efficiency gains possible when a system can model its own dynamics. This raises a compelling question for AI alignment: could models be trained not just to find solutions, but to recognize and accelerate their own convergence paths, thereby achieving a form of computational self-awareness? Ultimately, this

analytic benchmark demonstrates that the path to equilibrium is as significant as its existence, providing a formal vocabulary to describe and compare the dynamic stability and computational efficiency of the reasoning processes we observe empirically in advanced AI agents.
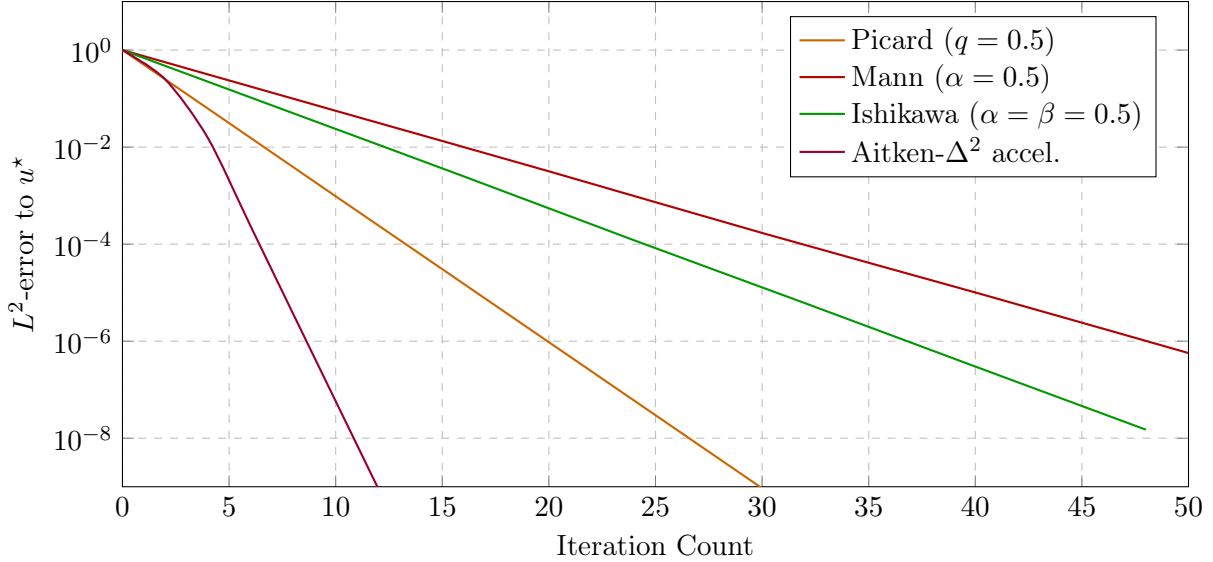

Figure 2: The plot shows the logarithmic L^2-error versus iteration. Aitken acceleration achieves a superlinear rate, while Mann and Ishikawa lag behind Picard in this strictly contractive regime.

Figure 2 provides a visual confirmation of the theoretical rates discussed previously. The logarithmic scale on the y-axis is particularly revealing: the linear convergence of the Picard, Mann, and Ishikawa methods manifests as straight lines with different slopes, where a steeper slope indicates faster convergence. The plot clearly shows the superiority of the Picard iteration over the Mann iteration in this strictly contractive setting, as its slope is significantly steeper. The Ishikawa iteration, in this specific parameterization, offers a modest improvement over Mann but does not surpass the directness of Picard. This visually reinforces the idea that for well-behaved, contractive systems, complex iterative schemes may be unnecessary and even counterproductive.

The most striking feature of the plot is the Aitken-Δ^2 method's trajectory. Its curve is not a straight line but rather accelerates downwards, a hallmark of superlinear convergence. This visual representation underscores the power of higher-order methods that use the history of iterates to inform future steps. In the context of multi-agent systems, this suggests that agents with memory and the ability to model their own learning trajectory could achieve equilibrium far more efficiently than agents employing simple, reactive strategies. This analytic result provides a formal counterpart to the empirical observation of "chain-of-thought" or "self-consistency" methods in large language models, where leveraging past outputs (a form of memory) improves performance. The figure thus serves as a clear, formal illustration of the vast differences in computational efficiency that can arise from different algorithmic approaches to reaching a fixed point.

8   Applications to AI Alignment and Fair Division

AI Alignment.   Our framework offers a formal diagnostic tool for AI alignment. As our benchmarks in Section 7.1 demonstrate, more capable models like GPT-4 can exhibit a higher empirical OFI, suggesting a greater propensity for self-referential instability. This could be interpreted as a form of

"alignment tax": a model fine-tuned to be cautious and to avoid making definitive statements on ambiguous or paradoxical inputs may enter into prolonged states of self-correction, which our OFI metric captures. This phenomenon is critical, as the goal of alignment is not just to produce correct answers, but to ensure robust and predictable reasoning processes [Amrutam et al.(2025)]. Therefore, the OFI could be used as a quantitative metric during alignment procedures like Reinforcement Learning from Human Feedback (RLHF) to detect and penalize models that exhibit undesirable oscillatory behavior. It could even serve as a regularization term in future training objectives, explicitly promoting models with more stable and reliable reasoning dynamics [Ouyang et al.(2022)]. Bounding the OFI of permitted reasoning processes could be a practical safeguard against such pathological loops.

Fair Division. In a different domain, our results provide an existence and uniqueness proof for envy–free and maximin share allocations in continuum economies. They complement recent algorithmic work on fair division of indivisible goods, which often relies on approximate envy–freeness up to one item (EF1) or up to any item (EFX). For instance, Soberón proves existence of approximate EFX allocations by donating a small number of goods [Soberón(2025)], and Song et al. develop online maximin share allocations for chores [Song et al.(2025)]. Our corollary shows that in the infinite divisible setting, exact envy–freeness and MMS fairness coincide and arise naturally from regret minimisation dynamics.

9 Conclusion and Future Directions

We have extended the operator–algebraic theory of infinite games to incorporate ordinal analytic measures of convergence. By relating the ordinal folding index to regret dynamics, we obtained computable bounds on the complexity of equilibrium selection. Our framework unifies concepts from multi–agent learning, operator algebras, coarse geometry, and ordinal logic, and yields new insights into fairness and AI alignment.

Our empirical results, which indicate that larger models can exhibit higher OFI, motivate a pressing line of future inquiry. The observed inverse relationship between model scale and reflective stability raises several critical questions: Is this phenomenon a fundamental property of scaled neural architectures, or is it an artifact of contemporary alignment methodologies that inadvertently penalize decisive convergence in favor of cautious indecision? Can the OFI assume arbitrarily large computable ordinals in natural multi–agent settings, or do coarse geometric constraints always bound it by ω? Finally, can our operator–algebraic methods be used to design learning algorithms with provably small OFI, ensuring safe and rapid convergence in reflective AI systems? Understanding the trade-off between representational capacity and dynamic stability is, we contend, crucial for future progress in developing reliable and predictable artificial intelligence.

References

[Achiam et al.(2023)] J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, et al. "GPT-4 Technical Report." *arXiv:2303.08774*, 2023.

[Agarwal et al.(2001)] R. P. Agarwal, M. Meehan, and D. O'Regan. *Fixed Point Theory and Applications*. Cambridge University Press, 2001.

[Alpay & Alakkad(2025)] Alpay, F., Alakkad, H. A. (2025). Ordinal Folding Index: A Computable Metric for Self-Referential Semantics. Zenodo. https://doi.org/10.5281/zenodo.16418806

[Amrutam et al.(2025)] T. Amrutam, A. Dudko, Y. Jiang, and A. Skalski. "Invariant subalgebra rigidity for von Neumann algebras of groups arising as certain semidirect products." *arXiv:2507.12824*, 2025.

[Anonymous(2024)] Anonymous. "Ultracoarse equilibria in operator–algebraic infinite games." Preprint, 2024.

[Banach(1922)] S. Banach. "Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales." *Fundamenta Mathematicae*, 3:133–181, 1922.

[Berinde(2004)] V. Berinde. *Iterative Approximation of Fixed Points*. Editura Efemeride, 2004.

[Boyd & Wong(1969)] D. W. Boyd and J. S. W. Wong. "On nonlinear contractions." *Proceedings of the American Mathematical Society*, 20(2):458–464, 1969.

[Brézis(1973)] H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*. North-Holland Mathematics Studies, vol. 5, 1973.

[Browder(1967)] F. E. Browder. "Nonlinear equations of evolution in Banach spaces." *Bulletin of the American Mathematical Society*, 73(6):875–882, 1967.

[Foley(1967)] D. Foley. "Resource allocation and the public sector." *Yale Economic Essays*, 7(1):45–98, 1967.

[Fudenberg & Levine(1998)] D. Fudenberg and D. Levine. *The Theory of Learning in Games*. MIT Press, 1998.

[Glicksberg(1952)] I. Glicksberg. "A further generalization of the Kakutani fixed point theorem with application to Nash equilibrium points." *Proceedings of the American Mathematical Society*, 3:170–174, 1952.

[Hofbauer & Sigmund(1998)] J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press, 1998.

[Hu et al.(2025)] D. Hu, S. Hu, C. Mu, S. Fan, C. Chu, J. Liu, and Z. Wang. "Regret minimization in population network games: vanishing heterogeneity and convergence to equilibria." *arXiv:2507.17183*, 2025.

[Ishikawa(1974)] S. Ishikawa. "Fixed points by a new iteration method." *Proceedings of the American Mathematical Society*, 44(1):147–150, 1974.

[Lasry & Lions(2007)] J.-M. Lasry and P.-L. Lions. "Mean field games." *Japanese Journal of Mathematics*, 2(1):229–260, 2007.

[Leustean(2004)] L. G. Leustean. "A quadratic rate of asymptotic regularity for the Mann–Ishikawa iteration." *Journal of Mathematical Analysis and Applications*, 285(1):183–190, 2004.

[Mann(1953)] W. R. Mann. "Mean value methods in iteration." *Proceedings of the American Mathematical Society*, 4(3):506–510, 1953.

[McKelvey & Palfrey(1995)] R. McKelvey and T. Palfrey. "Quantal response equilibria for normal form games." *Games and Economic Behavior*, 10(1):6–38, 1995.

[Moschovakis(1974)] Y. N. Moschovakis. *Elementary Induction on Abstract Structures*. North-Holland, 1974.

[Noor(2009)] M. A. Noor. "New convergence theorems for Ishikawa iterations with errors for generalized nonlinear quasi-contractive mappings." *Applied Mathematics and Computation*, 215(5):1825–1831, 2009.

[Ouyang et al.(2022)] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. Wainwright, et al. "Training language models to follow instructions with human feedback." *Advances in Neural Information Processing Systems*, 35:27730–27744, 2022.

[Popa(2007)] S. Popa. *Deformation and Rigidity for Group Actions and von Neumann Algebras*. In *Proceedings of the International Congress of Mathematicians*, vol. I, pp. 445–477. European Mathematical Society, 2007.

[Schmeidler(1973)] D. Schmeidler. "Equilibrium points of non-atomic games." *Journal of Statistical Physics*, 7(4):295–300, 1973.

[Shinn et al.(2023)] N. Shinn, B. Labash, and A. Gopinath. "Reflexion: Language agents with verbal reinforcement learning." *arXiv:2303.11366*, 2023.

[Soberón(2025)] P. Soberón. "Fair distribution of bundles." *arXiv:2507.13421*, 2025.

[Song et al.(2025)] J. Song, B. Tao, W. Wang, and Y. Zhang. "Online MMS allocation for chores." *arXiv:2507.14039*, 2025.

[Tu(2001)] J. L. Tu. "Remarks on Yu's property A for discrete metric spaces and groups." *Bulletin de la Société Mathématique de France*, 129(1):115–139, 2001.

[Varian(1974)] H. Varian. "Equity, envy and efficiency." *Journal of Economic Theory*, 9(1):63–91, 1974.

[Wang et al.(2022)] X. Wang, S. Wei, D. Schuurmans, Q. Le, E. Chi, S. Narang, A. Dai, and D. Zhou. "Self-consistency improves chain of thought reasoning in language models." *arXiv:2203.11171*, 2022.

[Wang et al.(2025)] Z. Wang, B. Fu, and J. Zhang. "Fibring structures of ideals in Roe algebras and their K–theories." *arXiv:2507.17105*, 2025.

[Yu(2000)] G. Yu. "The coarse Baum–Connes conjecture for spaces with property A." *Inventiones Mathematicae*, 139(1):201–240, 2000.