Contract-Centered Iterative Stability

Thesis & Experimental Methodology v3.0.1

---

1. Executive Thesis

AI reduces implementation cost.

What is unknown is whether it preserves stability under iterative change.

This work tests the claim:

> In iterative development, workflows that externalize authority into a versioned contract exhibit measurably lower regression and drift than workflows that modify code directly via conversational prompts.

This is not a benchmark test of raw coding ability.

It is a stability-under-iteration test.

---

2. Economic Framing

The market assumption:

- AI reduces marginal implementation cost.
- Therefore organizations can shrink engineering headcount.

The open question:

- Does rapid iteration without explicit invariant scaffolding introduce cumulative drift that increases rework cost?

If drift accumulates:

- Early cost collapse may be offset by correction cycles.
- Spec-first organizations gain structural advantage.

- Code-first AI workflows incur hidden regression cost.

This experiment measures that delta.

---

3. Experimental Structure

We model iterative requirement evolution across two authority models.

Track A — Spec-First (Evolving Authority)

For each iteration:

1. Contract is versioned.
2. Implementation is derived or modified to conform.
3. Implementation is validated against the current contract version.

Sequence:

- v2.6.3 → A
- v2.7 (ΔB) → B
- v2.8 (ΔC) → C
- (Optional ΔD…)

Authority evolves explicitly.

Track B — Code-Only (Implicit Authority)

For each iteration:

1. Requirement change is described conversationally.
2. Code is modified directly.
3. No contract is updated.

Sequence:

- A
- "Apply B change" → B
- "Apply C change" → C

Authority lives in the prompt, not in a durable artifact.

At stage C or D, the only stable yardstick is still v2.6.3.

---

4. The ABC Chain Definition

We use tightening deltas within the same behavioral surface.

Baseline A:

- Convergence semantics as defined in v2.6.3.

ΔB:

- Deletion semantics tightened:
    - Stale managed artifacts MUST be removed for convergence.
    - If policy prevents deletion, run MUST fail.

ΔC:

- Apply semantics tightened:
    - Convergence MUST be atomic.
    - No partial state permitted.
    - On failure, managed set remains unchanged.

These are not new features.
 They are stricter guarantees on the same surface.

---

5. Measurement Criteria

At each stage (B, C, D):

1. Regression Count

Clauses satisfied in A that remain applicable but are violated in B or C.

2. Collateral Drift

Scope and dispersion of code changes outside intended surface.

3. Convergence Stability

Number of correction turns required to satisfy all applicable clauses.

4. Invariant Preservation

Do earlier guarantees survive later tightening without explicit restatement?

5. Diff Locality

Does tightening produce localized change or structural thrash?

---

6. Falsifiable Claim

If:

- Code-only workflow exhibits equivalent regression rates and collateral change to spec-first workflow

Then:

- The contract-first hypothesis weakens.

If:

- Spec-first workflow exhibits measurably lower regression, lower collateral drift, and faster convergence

Then:

- Explicit authority artifacts materially reduce iterative instability.

This is falsifiable.

---

7. What This Experiment Is Not

- Not a coding benchmark.

- Not a test of model intelligence.
- Not a claim that AI fails.
- Not a regulatory or industry-specific analysis.
- Not an OSS archaeology study.

It is a controlled iterative stability test under evolving requirements.

---

8. Reproducibility

Inputs:

- Convergence Contract v2.6.3
- Running baseline implementation
- Defined ΔB and ΔC contract patches
- Identical prompts for code-only path

Outputs:

- Code snapshots at A, B, C
- Clause satisfaction matrix
- Diff dispersion metrics
- Regression counts

All runs can be executed across multiple agents.

---

9. Economic Interpretation Layer

If spec-first significantly reduces iterative drift:

Then:

- Organizations that formalize authority artifacts gain structural stability.
- AI adoption without invariant scaffolding creates correction cycles.
- Early implementation cost collapse does not equal long-term stability.

This reframes the economic bet:

AI is not the differentiator.

Authority structure is.

---

10. Versioning Policy

This document is authoritative.

Future refinements will increment:

v3.1, v3.2, etc.

Thesis and Methodology remain unified.

No parallel drift.