

Containers Before Contents

Boundary Dynamics, Selective Permeability, and the Mathematics of Inside

Flyxion

Abstract

Most accounts of identity and persistence begin with distinctions: a system is identified with a partition between what belongs to it and what does not, and questions of survival are framed as questions about whether that partition is preserved. This paper argues that the partition is not the right primitive. A boundary is not a distinction with crossing rules; it is the ongoing activity by which a distinction continues to be produced. From this claim — the Boundary Primacy Theorem — the paper develops a formal account of containment as a maintained process rather than a static structure, anchored by a demarcation criterion that prevents the framework from degenerating into a relabeling of persistence itself. Containment failure is shown to admit at least four structurally distinct modes: absence of any identifiable maintenance process, maintenance that is identifiable but causally inert, maintenance that is causally effective but insufficient, and maintenance that becomes, in part, a source of the destabilization it was meant to resist. A further case — typified by the relationship between a tumor and its host — shows that containment cannot in general be evaluated as a property of a single system, forcing a relational extension in which containment status depends on a system's effect on the maintenance dynamics of whatever it is embedded within. The resulting picture treats containers not as isolated objects but as positions within graphs of mutual maintenance, destabilization, and dependency, with autonomy emerging as a limiting regime rather than a category. The framework is applied uniformly to cells, languages, states, and corporations, using a single repeated diagnostic question rather than four independent case studies.

The paper's central formal move is a distinction between two relations that nested systems are ordinarily treated as one: a constitutive relation, specifying what a system is made of, and an interaction relation, specifying what it merely affects or is affected by. Separating these resolves a case that would otherwise stand as a counterexample to the whole framework — a tumor satisfies every criterion for containment in isolation while functioning, relationally, as an attack on its host — and yields a refinement-invariant dependency ratio, a criterion for when reciprocally dependent systems should be treated as fusing into a single higher-order container, and a diagnostic procedure that locates collapse by elimination across the full taxonomy of failure modes. The paper's deepest claim, established only once this machinery is in place, is that containment is not ultimately a property of isolated systems at all, but a property of positions within ecologies of mutual maintenance, destabilization, and dependency.

Contents

1. The Boundary Primacy Theorem	2
2. The Demarcation Criterion	3
2.5 Boundaries Without Geometry	5
3. Efficacy and Sufficiency	6
3.5 Measuring Efficacy	7
4. Reflexivity	9
4.5 The Reflexivity Threshold	9
5. The Tumor Problem	11
6. Containment Ecologies and Dependency Regimes	14
7. Four Domains, One Diagnostic	16
7.5 Diagnostic Procedure	18
8. Conclusion	19
Appendix A: Relation to Autopoietic Theory	20

1. The Boundary Primacy Theorem

Most theories of structure begin with distinctions. A distinction partitions the world into two classes:

$$D = (X, \neg X).$$

Once the partition is specified, the distinction simply exists. It may be useful, arbitrary, correct, or incorrect, but nothing in the concept of a distinction requires ongoing work to keep the partition intact. A distinction is stated, not sustained.

Containment appears, at first glance, to be a special case of this. Every container seems to induce a distinction between inside and outside. This similarity has encouraged a natural reduction:

A boundary is merely a distinction together with a rule governing passage between its two sides.

This paper rejects that reduction.

The error is subtle. It assumes that the partition already exists and asks only how crossing occurs. But many of the systems we care about most are not characterized by a preexisting partition that crossing rules then regulate. They are characterized by the continuous *production* of one.

A cell membrane does not merely separate inside from outside. Ion pumps, repair pathways, trafficking systems, and metabolic activity continuously recreate the conditions

under which “inside” and “outside” remain meaningful biological categories. A state does not merely possess borders. Courts, taxation, administration, logistics, and enforcement continuously reproduce the distinction between member and nonmember. A language does not merely possess a vocabulary. Intergenerational transmission — schools, households, liturgy, media — continuously reproduces the distinction between utterances that belong to the language and utterances that do not.

In each case, the distinction is not primitive. The maintenance process is primitive. A boundary is not a distinction with additional properties. A boundary is the ongoing activity by which a distinction remains separately inhabitable over time.

Objection. Let $D_t = (X_t, \neg X_t)$. Why is a separate maintenance primitive needed? Time-dependence alone seems to resolve the difficulty: the partition simply updates at each instant.

Reply. The sequence $\{D_t\}$ specifies states, not the processes that make later states depend upon earlier ones. Nothing in the concept of a time-indexed distinction explains why D_{t+1} continues D_t rather than being replaced by an unrelated, coincidentally similar partition at the next instant. The persistence of the partition remains unexplained by the indexing itself. The primitive is therefore not temporal succession but *maintained succession*: the existence of processes that causally constrain which future partitions can emerge from present ones. A cell whose membrane chemistry stopped at this instant would not reliably generate the next biologically meaningful partition from the present one. The distinction would cease to be maintained, and lysis would follow. The next partition is not given. It is earned.

This gives the central claim of the paper.

Boundary Primacy Theorem. A distinction partitions a space. A boundary generates and maintains the conditions under which a partition continues to exist. Therefore boundaries cannot, in general, be reduced to distinctions plus directional crossing rules; the relation between successive states of a maintained partition is not itself expressible as a distinction.

The rest of the paper follows from this claim with a kind of inevitability. Once maintenance is primitive, a criterion is needed for identifying maintenance activity independently of the outcomes it is invoked to explain. Once maintenance is identifiable, a causal account is needed of how it generates the boundary work that sustains a partition. Once maintenance can fail, a taxonomy of failure is needed — failure to exist, failure to matter, failure to suffice, and failure through the maintenance process turning against the very interior it protects. And once failure can be generated *within* one system by activity that simultaneously constitutes another system’s maintenance — the case of the tumor, the insurgent faction, the parasitic subsidiary — the concept of containment is forced to become relational rather than unary: not “is B a container,” but “is B a container relative to a host H .”

Section 2 introduces the formal apparatus needed to state these claims precisely. But the apparatus is downstream of the argument made here in prose: a maintained partition and a static partition are different kinds of objects, and the difference is not a matter of degree.

2. The Demarcation Criterion

Section 1 established that a boundary is the activity that sustains a partition, not the partition itself. This raises an immediate danger. If anything that persists can be retroactively credited with “maintenance,” the concept does no work. A skeptical reader is entitled to ask: how would we know maintenance had occurred, other than by observing that the partition survived?

This is not a minor technical worry. It is the difference between a theory and a description. If maintenance is inferred only from the outcome it is meant to explain, the theorem of Section 1 collapses into a tautology — *things that persist were, by definition, maintained* — and the entire apparatus becomes bookkeeping dressed as explanation. The theory must therefore specify, at the outset, what would count as evidence of maintenance independent of whether containment ultimately succeeds.

A container, at time t , consists minimally of an interior I_t , an exterior O_t , a boundary ∂B_t separating them, and a maintenance process M_t responsible for sustaining that separation against decay. Further quantities are introduced later, individually, at the point each first does work: a selective permeability Π_t — a function over types of crossing, not a scalar — and a leakage rate L_t in Section 3, where the rate equation requires them, and a cost of separation C_t , also in Section 3, where it enters the failure inequality. No single composite object collects all of these at once. The quantities that matter are the ones that appear in an inequality or a rate equation somewhere in what follows; nothing is carried for its own sake.

The role of M_t is to name, independently, the thing Section 1 argued must exist. This is only useful if M_t is constrained by a criterion that prevents it from being read backward off the outcome.

Maintenance Identifiability Criterion. A candidate maintenance process M_t is admissible only if it can be specified, observed, or measured independently of whether containment ultimately succeeds or fails. Equivalently, M_t must not be a post hoc reconstruction from the containment outcome:

$$M_t \not\equiv f(\text{containment outcome}) \quad \text{for any reconstruction } f.$$

In the cell, M_t is ion-pump activity, membrane repair pathways, osmotic regulation, and protein trafficking — mechanisms a biologist can point to, measure, and characterize whether or not the cell happens to survive. In the state, M_t is taxation, courts, administration, and enforcement — institutions that can be inventoried and assessed in their own right, independent of whether the state in question collapses next year. In the language, M_t is schools, household transmission, literacy practice, and media production — observable social activity that can decline, persist, or strengthen independently of the language's ultimate fate. In each case the criterion is satisfiable: there exists something a competent observer could point to *before* knowing the outcome, and that something is not merely a redescription of survival.

This criterion does real work because it can fail. Not every persistent or boundary-like system satisfies it, and the cases where it fails are as informative as the cases where it succeeds.

The rock. A rock has an interior and an exterior, in the trivial sense that “inside the rock” and “outside the rock” are well-defined regions of space. A rock may persist for millennia. But there is no independently specifiable activity whose function is to preserve that partition against dissolution. The rock's continued existence is not maintained; it is merely *not yet eroded*. Nothing internal to the rock is doing work to keep “rock” and “not-rock” apart. The identifiability criterion fails — not because the rock lacks persistence, but because persistence here has no candidate cause that is not simply the absence, so far, of sufficient destructive force. Container(rock) is therefore false in the technical sense developed here, whatever one wants to say about the rock informally.

The fad. A fad — a slang term, a dance, a fashion — exhibits something that looks superficially like a boundary: there are utterances or behaviors that count as “in” and others that count as “out.” But there is no process whose specifiable role is to keep that

partition intact. A fad's adoption and abandonment are driven by the same diffusion dynamics that drive its emergence; nothing is recruited specifically to defend the in/out line once it is crossed too often. This is the relevant contrast with the language case above: a language has institutions — schools, dictionaries, editorial standards, corrective social pressure on “incorrect” usage — whose specifiable function is the defense of the in/out line. A fad has no such institutions. When a fad fades, nothing failed, because nothing was trying. Container(fad) is false for the same structural reason as the rock, despite the two cases looking, on the surface, like nothing alike.

These negative cases motivate a second theorem, which restricts the scope of everything that follows.

Demarcation Principle. Not every system exhibiting an inside/outside distinction is a container. Container(B) holds only if there exists M_t satisfying the Maintenance Identifiability Criterion. The converse does not hold: identifiable maintenance activity does not guarantee that containment succeeds, only that the system is the right *kind* of object for containment to be a meaningful question about.

The remainder of the paper is concerned with systems that pass this filter, and with the conditions under which their maintenance activity succeeds, fails outright, or — in the more troubling cases taken up in Section 4 — turns against the very interior it was meant to protect.

2.5 Boundaries Without Geometry

The examples used so far divide unevenly. A cell's membrane is a physical surface, locatable in space, and its rupture is a literal event. A state's border is partly the same — a line on the ground, a fence, a checkpoint — and partly something else, since a state's interiority is at least as much a matter of jurisdiction as of geography. A language and a corporation have no spatial boundary in any of these senses at all. Nothing in the theory developed so far has explained whether this is a difference of degree or a sign that the formalism has been silently smuggling in an assumption — spatial enclosure — that only some of its intended cases actually satisfy.

The formalism does not require that assumption, but the paper owes the reader an explicit statement of why. The relevant primitive, restated without reference to space, is this: a boundary is whatever mechanism sorts candidate continuations of a system into those that belong to it and those that do not.

Definition (Admissibility Boundary). A boundary ∂B for a system B is a mechanism that partitions the space of candidate next states, or candidate trajectories, into those admissible as continuations of B and those that are not.

For a cell, admissible continuations are states reachable through normal membrane transport and metabolic activity; a state reached by uncontrolled rupture is inadmissible, and the boundary that sorts these is, among other things, a physical surface. For a language, admissible continuations are utterances a competent speaker, an editor, a school, or a community of usage would recognize as belonging to the language; an utterance outside that admissible set is simply not part of the language, regardless of where it is spoken or written, and the boundary that performs this sorting is normative rather than spatial. For a corporation, admissible continuations are actions that fall within its legal structure, governance authority, and decision rights; an action outside that set — an employee's purchase made without authority, a contract signed by someone without standing to sign it — does not belong to the corporation's trajectory even though no spatial line was crossed. For a state, admissibility is typically a composite of both: geographic boundary together

with jurisdictional and administrative reach, which is why states are the domain in which the spatial and the normative pictures of containment are hardest to fully separate.

Under this definition, the maintenance activity of Section 2 has the same role in every domain regardless of whether the boundary it sustains happens to be spatial. Ion pumps maintain a spatial admissibility boundary by regulating which molecular states are reachable from the cell's interior. Schools and households maintain a normative admissibility boundary by regulating which utterances remain reachable, recognizable continuations of the language. Governance structures maintain an institutional admissibility boundary by regulating which actions remain reachable continuations of the corporation's authority. The mechanism differs. The function — sorting candidate continuations into admissible and inadmissible — does not.

Boundary Generalization Principle. Containment requires selective continuation, not spatial enclosure. A system is a candidate for containment whenever some independently identifiable maintenance process sorts its possible continuations into admissible and inadmissible classes, whether or not that sorting has any spatial signature.

This resolves the earlier worry by dissolving rather than answering it. The question “is a language's boundary metaphorical?” presupposes that the spatial case is literal and every other case is a figure of speech borrowed from it. The admissibility boundary defined here reverses the relation: the spatial case is one instantiation of a more general mechanism, not its template. A cell's membrane is not what containment “really” is, with languages and corporations merely behaving *as if* they had something similar. All four domains examined in Section 7 satisfy the same definition; only the substrate of the sorting mechanism — biochemical, social, legal — differs from one to the next.

3. Efficacy and Sufficiency

Identifiability settles whether a candidate maintenance process exists as a distinct object of inquiry. It does not settle whether that process does anything. A system can possess elaborate, well-documented, entirely genuine maintenance activity that nonetheless has no bearing on whether containment is preserved. A state may stage ceremonial demonstrations of sovereignty — parades, declarations, monuments — that are real, costly, intentional, and fully observable, while the actual work of holding the state together happens, or fails to happen, somewhere else entirely. Identifiability alone cannot distinguish a load-bearing process from a decorative one.

This calls for a second, stronger condition, stated in terms of the boundary work the maintenance process is supposed to generate. Let W_t denote the boundary work available at time t — the capacity to sustain Π_t against L_t — and let M_t produce W_t through some function g :

$$W_t = g(M_t).$$

Efficacy Criterion. A candidate maintenance process is admissible only if it is causally connected to boundary work over the relevant operating range:

$$\frac{\partial W_t}{\partial M_t} > 0 \quad \text{for } M_t \text{ in the system's typical operating region.}$$

A ritual demonstration of sovereignty that does not in fact influence tax compliance, court legitimacy, or administrative reach fails this criterion: M_t is identifiable, but $\partial W_t / \partial M_t \approx 0$. This is a distinct failure mode from the absence of any candidate process at all. The rock

fails identifiability — there is no M_t to evaluate. The ritual state fails efficacy — M_t exists, and is real, but does not connect to W_t .

3.5 Measuring Efficacy

The efficacy criterion as stated is a claim about a partial derivative, $\partial W_t / \partial M_t > 0$, and a partial derivative is not, by itself, something an observer in the field can read off a system directly. The criterion needs an operational counterpart: some indication of what would count as evidence that a candidate maintenance process actually moves boundary work, as opposed to merely accompanying it.

The natural form of that evidence is interventionist rather than observational. Correlation between M_t and containment outcomes is exactly what the demarcation criterion in Section 2 already warned against treating as sufficient, since correlation alone cannot distinguish a load-bearing process from a process that merely happens alongside survival. What distinguishes efficacy from a coincidental correlate is whether containment outcomes change under a change to M_t that does not also change anything else relevant.

Observational Criterion for Efficacy. A maintenance process M_t is efficacious if interventions on M_t systematically alter containment outcomes:

$$P(\text{containment} \mid \text{do}(M_t + \epsilon)) \neq P(\text{containment} \mid M_t)$$

for some perturbation ϵ within the system's operating range, where $\text{do}(\cdot)$ denotes an intervention that sets M_t to a new value rather than merely observing it at that value. This is the same distinction ordinary experimental reasoning already relies on: an intervention isolates the effect of M_t from whatever else happens to be correlated with it, in a way that passive observation of naturally occurring variation in M_t cannot.

Real cases satisfy this criterion to differing degrees of directness, and the differences are themselves informative. In the cell, the criterion is close to literal: a pump knockout — genetically or pharmacologically disabling a specific transport mechanism — is an intervention on M_t in the strict sense, and the resulting change, or absence of change, in osmotic stability is direct evidence of efficacy or its lack. In the state, school closures, the suspension of tax collection, or the withdrawal of court funding function as natural or deliberate interventions whose effects on administrative reach and compliance can be tracked, even where a controlled experiment is not available; historical episodes in which one component of M_t was removed while others remained constant — a tax revolt with courts and administration otherwise intact, for instance — approximate the intervention the criterion calls for. In the corporation, removal or replacement of a governance function — a board reorganization, the dissolution of a compliance department — serves a similar role, and the frequency with which such interventions are followed by measurable changes in operational outcomes is itself a rough empirical test of whether the governance structure in question was efficacious or merely ceremonial. The language case is the least direct: nothing closely resembling a controlled withdrawal of household transmission typically occurs, and efficacy there is usually inferred from naturally occurring variation — comparing communities or generations in which transmission practices differed — rather than from anything resembling a clean intervention.

This unevenness is not a defect particular to the present theory. It tracks a genuine difference in how directly maintenance processes can be manipulated across domains, the same difference that makes randomized experiments routine in cell biology and all but impossible in macro-institutional settings. The criterion does not require uniform experimental access across domains to be meaningful. It requires only that, in each domain,

some principled distinction can be drawn between a maintenance process whose removal or alteration would change containment outcomes and one whose removal would not — a distinction available in every case examined here, even where the available evidence falls well short of a controlled trial.

Even a maintenance process that satisfies both identifiability and efficacy can still fail to preserve containment, simply by being outmatched. This is the most intuitive of the failure modes, and the one ordinary language already tracks: a system can be trying, and trying effectively, and still lose.

Sufficiency Criterion. Containment is preserved at time t only if the boundary work generated by maintenance activity meets or exceeds the cost of sustaining separation:

$$g(M_t) = W_t \geq C_t(\partial B_t, \Pi_t, I_t, O_t).$$

Failure occurs when this inequality reverses:

$$C_t > g(M_t).$$

A state with a fully functioning tax apparatus, an effective judiciary, and competent administration — that is, with M_t satisfying both identifiability and efficacy — can still be overrun by invasion, fractured by debt, or fragmented by simultaneous internal crises that drive C_t above what any feasible $g(M_t)$ could supply. Nothing was performative about its maintenance. It was simply insufficient. This is the structural difference between a ritual state, which never had real W_t to begin with, and an overwhelmed state, which had real W_t and lost anyway. Conflating these two cases — both of which present, from the outside, as “the state collapsed” — is one of the more common errors in informal accounts of institutional failure. The formalism keeps them separate by construction.

These two criteria, efficacy and sufficiency, give content to the rate dynamics implicit in Section 1’s claim that boundaries are processes rather than states. The interior mass or content I_t changes according to

$$\frac{dI_t}{dt} = \Phi_{\text{in}}(\Pi_t(M_t), O_t) - \Phi_{\text{out}}(\Pi_t(M_t), I_t) - L_t(M_t),$$

where Φ_{in} and Φ_{out} are flows across the boundary regulated by the maintenance-dependent permeability $\Pi_t(M_t)$, and $L_t(M_t)$ is leakage — flow the maintenance process fails to regulate. Containment fails when accumulated cost exceeds available boundary work:

$$C_t > g(M_t).$$

Three failure modes have now been distinguished, in order of increasing subtlety. A system can fail to be a container at all, for lack of any independently identifiable maintenance process (identifiability failure: the rock). A system can possess identifiable maintenance that is nonetheless causally inert with respect to boundary work (efficacy failure: the ritual state). And a system can possess identifiable, causally effective maintenance that is simply outmatched by the cost of separation (sufficiency failure: the overwhelmed state).

A fourth mode remains, and it is the most consequential, because it cannot be diagnosed by asking whether maintenance is present, effective, or strong enough. It arises when maintenance activity itself becomes the source of the destabilization it was meant to resist — when g and C_t are no longer independent of one another because M_t feeds both. Section 4 takes up this case.

4. Reflexivity

Sections 2 and 3 treated cost, C_t , as though it arrived from outside the system — pressure, decay, leakage, invasion, demographic loss. This is adequate for many cases, but it conceals an asymmetry the theory should not assume away. Maintenance was given a careful, independently grounded treatment; the forces opposing it were folded into a single cost term with no comparable scrutiny. This section corrects the imbalance by promoting destabilization to a first-class process, D_t , subject to the same demands placed on M_t in Section 2: it must be independently identifiable, not inferred backward from the fact of collapse.

For the cell, D_t includes osmotic gradients, toxins, and viral intrusion. For the state, warfare, insurgency, and fragmentation. For the language, competing prestige languages, demographic loss, and institutional replacement. In each case D_t is observable on its own terms, exactly as M_t was required to be.

With D_t established as a genuine process rather than a residual cost, the boundary failure condition of Section 3 generalizes to

$$g(M_t) > C_t + h(D_t),$$

where h converts destabilization activity into its effect on the cost of separation, by analogy with g converting maintenance activity into boundary work.

Both g and h were implicitly assumed, in Section 3, to be monotonic — more maintenance always helps, more destabilization always hurts. Neither assumption survives scrutiny. Moderate external pressure on a state typically increases internal fragmentation, consistent with $\partial h/\partial D_t > 0$. But extreme external pressure sometimes produces the opposite effect: invasion or existential threat can generate solidarity, mobilization, and coordination that a moderate version of the same pressure would not — a region in which $\partial h/\partial D_t < 0$. The destabilizing process becomes, within that region, partially stabilizing. D_t and M_t are therefore better understood as families of processes with operating regimes, not as scalar forces with fixed sign.

The more troubling non-monotonicity, however, runs through M_t itself. There is no guarantee that $\partial W_t/\partial M_t$ remains positive across the full range of maintenance intensity. Beyond some threshold M_t^* , maintenance activity can begin to generate the very destabilization it was introduced to suppress. This deserves to be stated as its own result, since it is the clearest way the present theory departs from purely functionalist accounts of persistence, in which failure is always a matter of insufficient effort.

Reflexivity Failure Theorem. A containment system may fail not because maintenance activity is absent, nor because it is insufficient, but because maintenance activity becomes, in part, a source of the destabilization it opposes. Writing

$$D_t = D_{\text{ext}}(t) + D_{\text{int}}(M_t), \quad \frac{\partial D_{\text{int}}}{\partial M_t} > 0 \text{ over some region,}$$

the failure condition $g(M_t) > C_t + h(D_t)$ can fail to hold even as M_t increases, because increasing M_t simultaneously increases D_{int} .

4.5 The Reflexivity Threshold

The theorem as stated asserts that $\partial D_{\text{int}}/\partial M_t > 0$ holds “over some region,” which leaves open a question the rest of the theory has been careful not to leave open elsewhere: where that region begins, and what distinguishes it from the region in which maintenance behaves

as expected. Without an answer, reflexivity risks being read as a possibility rather than a structural feature — something that might happen to a given system, rather than something the theory predicts will happen under specifiable conditions.

The relevant claim is that g itself is not globally monotonic in M_t . At low to moderate intensity, additional maintenance activity increases boundary work in the ordinary way the efficacy criterion of Section 3 describes. Beyond some threshold, however, the same increase in M_t begins to increase $D_{\text{int}}(M_t)$ faster than it increases $g(M_t)$, so that net containment capacity — $g(M_t)$ less the cost imposed by D_{int} — turns downward even as M_t continues to rise.

Reflexivity Threshold. There exists M_t^* such that

$$\frac{\partial W_t}{\partial M_t} > 0 \quad \text{for } M_t < M_t^*, \quad \frac{\partial D_{\text{int}}}{\partial M_t} > 0 \quad \text{for } M_t > M_t^*,$$

with the second inequality eventually dominating the first. The region below M_t^* is the ordinary operating regime already covered by the efficacy and sufficiency criteria. The region above it is the **overmaintenance regime**: the condition under which increasing M_t increases destabilization faster than it increases boundary work, so that further maintenance activity becomes net-negative for containment even though it remains, by every local measure, identifiable and nominally efficacious.

This gives the Reflexivity Failure Theorem a more specific empirical content than “maintenance can sometimes become harmful.” It predicts a turning point, and the existence of a turning point is what distinguishes reflexivity from ordinary insufficiency in a way that can in principle be checked against a trajectory of M_t over time, rather than only against a single before-and-after comparison. The immune system again supplies the clearest instance: graded immune response is protective at low and moderate levels and increasingly self-damaging at high levels, with autoimmune pathology characteristically associated with sustained activation past whatever threshold separates ordinary defense from sustained attack on host tissue. Policing, auditing regimes, and content moderation systems show the same qualitative shape in the institutional cases examined in Section 7 — protective and stabilizing up to some level of intensity, counterproductive beyond it — though the threshold in these cases is set by social and institutional dynamics rather than by anything as cleanly measurable as immune titer.

The threshold M_t^* need not be fixed for a given system. It can itself shift over time, as a function of the surrounding environment, the state of I_t , or prior history — a possibility this paper does not pursue further, but one the formalism does not foreclose. What the threshold does establish, at minimum, is that reflexivity is not an exception requiring special pleading each time it appears. It is what the efficacy relation $\partial W_t / \partial M_t > 0$ looks like once it is asked to hold globally rather than locally, and found not to.

Stated this generally, “reflexivity” risks becoming a catch-all into which any case of maintenance-correlated harm is dropped without further analysis — which would undermine the very precision the theory has been built to provide. The examples that motivate the theorem in fact divide into two mechanistically distinct channels, and treating them as one would obscure exactly the structure the formalism exists to reveal.

Type I reflexivity (interior misidentification). Maintenance activity acts directly on the protected interior rather than on the boundary: $M_t \rightarrow I_t$ rather than $M_t \rightarrow \partial B_t$. Autoimmunity is the paradigm case — the maintenance system loses the capacity to distinguish boundary from interior and attacks the latter directly. Over-aggressive quality control that destroys the productive capacity it was meant to safeguard, or institutional self-censorship severe enough to extinguish the creativity it was meant to protect, follow the same pattern. This is an error of target selection: the maintenance process is still trying

to defend something, but has lost track of what.

Type II reflexivity (environmental provocation). Maintenance activity remains correctly directed at the boundary, but alters the surrounding environment in ways that amplify D_{ext} : $M_t \rightarrow D_t \rightarrow h(D_t)$. Policing that generates the unrest it is meant to suppress, censorship that generates the resistance it is meant to prevent, and tariffs that generate the retaliation they are meant to forestall, all follow this pattern. Nothing about the maintenance process is malformed in the way Type I requires; the boundary is correctly targeted. The failure is a feedback effect external to the system's own interior.

Decomposing destabilization accordingly,

$$D_t = D_{\text{ext}}(t) + D_{\text{prov}}(M_t, D_t) + D_{\text{mis}}(M_t, I_t),$$

reflexivity is no longer a single undifferentiated failure category but two distinguishable pathways, each with its own diagnostic signature: Type I shows up as damage to I_t correlated with M_t 's own intensity; Type II shows up as a rise in D_{ext} correlated with M_t 's own intensity, with I_t otherwise undamaged.

The theory now has four failure modes — identifiability, efficacy, sufficiency, and reflexivity, the last divided into two channels — rather than the single, intuitive notion of “not enough maintenance” with which the inquiry began. This is a substantively richer claim than a functionalist account of persistence can make. A theory in which the only failure mode is insufficient maintenance treats every collapse as a shortfall to be corrected by more of the same activity. A theory that can also locate failure in maintenance turned pathological — against its own interior, or against its own environment — has a correspondingly richer geometry, and a corresponding obligation to show that this richness is not bought at the price of the precision established in Section 2. The rock and the fad, having failed identifiability outright, cannot exhibit either type of reflexivity: a rock has no maintenance process capable of misidentifying an interior, nor one capable of provoking an environment. The negative cases grow stronger, not weaker, as the positive machinery grows around them — which is the relevant test of whether that machinery has been earned.

The four failure modes developed so far all presuppose that a container can be evaluated on its own terms, as a single system with an interior, a boundary, and a maintenance regime to be assessed in isolation. Section 5 shows why this assumption cannot be sustained, by way of a case in which a system's maintenance activity is, simultaneously, another system's destabilization.

5. The Tumor Problem

Every criterion developed so far has been stated as a property of a single system: does B have identifiable M_t , is M_t efficacious, is $g(M_t)$ sufficient. Nothing in Sections 2 through 4 referenced any system other than B itself. This was a simplification, and it is now time to show that it cannot be maintained.

Consider a tumor. It possesses a boundary, often a literal capsule. It possesses selective permeability, regulating what crosses in and out. It possesses maintenance activity in the full technical sense of Section 2: angiogenesis recruits a dedicated blood supply, and immune evasion mechanisms are observable, specifiable, and causally efficacious independent of whether the tumor survives any particular immune response. That maintenance activity is sufficient, by hypothesis, to sustain the tumor's own interior against the cost of separation. Every criterion established so far is satisfied. The theory, as stated, is committed to:

Container(T).

This is the wrong answer, or at least an incomplete one, and the question is why.

The temptation is to patch the demarcation criterion — to add some further condition that tumors fail and legitimate containers pass. This temptation should be resisted, because it does not survive contact with the theory's own positive cases. Any condition strong enough to exclude the tumor on the grounds that its maintenance activity is somehow illegitimate, parasitic, or directed at the wrong scale will also exclude organs, which draw resources from a host body exactly as a tumor does; symbiotic organisms, which are nested inside a host by definition; corporate subsidiaries, which depend on a parent for capital and legal standing; and immigrant enclaves, which maintain internal cultural transmission while drawing on a surrounding economy. The problem cannot be solved by tightening the criterion for Container(B), because the criterion is not actually false. It is incomplete. The tumor is not failing to be a container. It is succeeding at being a container of the wrong *kind*, and "kind," here, is not a property the tumor has on its own. It is a property of the tumor's relationship to something else.

This usage departs from ordinary language, in which a tumor would more naturally be described as a pathological exception to containment, or as a failure of the host's containment rather than an instance of containment in its own right. The present theory makes a different choice deliberately: it treats parasitic and pathological cases as genuine containers, evaluated relationally rather than excluded by definition. This is not a terminological accident to be smoothed over. It is the substantive claim of this section — that the ordinary-language instinct to exclude tumors from the category of "container" is tracking something real, namely a relational property (negative effect on a host), but mislocating it as though it were a property of the tumor in isolation. The theory's extension of the term is what makes that relational property expressible at all.

This is the moment at which Container(B), understood as a predicate of a single system, has to be abandoned in favor of a relational object. A tumor is not merely B . It is B_2 , nested within a host B_1 :

$$B_2 \subset B_1.$$

Its maintenance activity draws boundary work from the host's interior — W_{B_2} is funded, in part, by resources extracted from I_{B_1} — and from the host's point of view, that same activity functions as an internal destabilizing process:

$$M_{B_2} = D_{\text{mis}}(B_1).$$

This identity is the precise sense in which the tumor is not a counterexample to the theory but a confirmation of a structure the theory had not yet made explicit. What one system experiences as maintenance, another system, in which it is nested, can experience as the interior-misidentification channel of reflexivity introduced in Section 4. The two descriptions are not in tension. They are the same activity, evaluated from two different positions in a containment ecology.

This licenses a general criterion for classifying nested containment, stated not in the qualitative vocabulary of "symbiotic," "neutral," and "parasitic," but as a continuous quantity: the effect of the nested system's maintenance on the host's available boundary work.

$$\frac{\partial W_{B_1}}{\partial M_{B_2}} > 0 \Rightarrow \text{symbiotic nesting (an organ; a healthy subsidiary)}$$

$$\frac{\partial W_{B_1}}{\partial M_{B_2}} = 0 \Rightarrow \text{neutral nesting (a benign, non-extractive enclave)}$$

$$\frac{\partial W_{B_1}}{\partial M_{B_2}} < 0 \Rightarrow \text{parasitic nesting (a tumor; an embezzling subsidiary)}$$

These are not three discrete categories but three regions of a continuous derivative, which has the immediate advantage of admitting transitions without requiring a categorical jump: a startup nested within a corporation may begin with $\partial W_{B_1} / \partial M_{B_2} > 0$, drift toward ≈ 0 as it becomes administratively costly to sustain, and end strictly negative without ever crossing a boundary the theory has to treat as a discontinuity. The sign of a derivative, not membership in a category, is doing the classificatory work.

This also resolves, retroactively, the question left open at the end of Section 4: whether D_{ext} , D_{prov} , and D_{mis} require distinct weights in $h(D_t)$, posited independently of one another. They do not need to be posited, because the relational structure of Section 5 generates the relevant distinctions from containment topology rather than from labels assigned in advance. External pressure is destabilization arriving from a system with no containment relation to B_1 at all — some B_3 acting on B_1 without being nested within it. Provocation is destabilization that leaves B_1 , propagates through the environment, and returns: $B_1 \rightarrow B_3 \rightarrow B_1$. Interior misidentification is destabilization generated by a system nested within B_1 itself: $B_2 \subset B_1$. The three channels of Section 4 are not three independent kinds of harm. They are three positions a destabilizing process can occupy in the containment ecology relative to the system it destabilizes, and h can remain a single unweighted function precisely because the weighting was never missing — it was implicit in which of these three structural relations obtained.

What this section shows is that $\text{Container}(B)$ was never quite the right object. The right object is:

$$\text{Container}(B \mid H),$$

containment relative to a host, environment, or containing context H , where H may be another specific system, as with the tumor and its host, or may be diffuse and aggregate, as with a state's relationship to a surrounding trade network or a cell's relationship to the organism it composes.

Relational Containment Principle. For nested systems, containment status cannot in general be determined from the properties of a system alone. It depends upon the effect of that system's maintenance activity on the maintenance dynamics of the host within which it is embedded. $\text{Container}(B)$, understood as a unary predicate, is at best an approximation, valid only in the limiting case where a host's reaction to B 's maintenance activity is negligible.

The rock and the fad remain outside this discussion entirely, and it is worth pausing to note why, now that the relational machinery has grown considerably since Section 2. Neither possesses an identifiable maintenance process capable of entering the nesting relations developed here. There is no M_{B_2} to evaluate as symbiotic, neutral, or parasitic, because there is no independently identifiable maintenance activity in the first place. The relational turn extends what counts as a container; it does not loosen the criterion that excluded the rock and the fad to begin with. They fail at the first step, and everything built since Section 2 — efficacy, sufficiency, reflexivity, nesting — simply never becomes a question that can be asked of them.

That limiting case is itself worth naming, since nothing in the theory so far has been shown to be free-standing in any stronger sense, and the apparent autonomy of states,

organisms, and corporations needs to be accounted for rather than quietly assumed. Section 6 takes up this limit, along with the dependency structure it leaves behind.

6. Containment Ecologies and Dependency Regimes

Section 5 replaced the unary predicate $\text{Container}(B)$ with the relational $\text{Container}(B | H)$, and treated H as if it were a single, identifiable host. Real systems rarely have one. A state's maintenance draws on trade networks, military alliances, financial systems, and ecological inputs simultaneously; a language's maintenance draws on speaker communities, media systems, and educational institutions that have no single point of unification. Treating H as a unique host, as the tumor case did for expository convenience, is a special case of something more general: a system embedded not in one container, but in a graph of relations to many others.

This requires distinguishing two relations that the theory has so far treated as one. The constitutive relation, $B_2 \subset B_1$, answers the question of what a system *is*: which sub-containers belong to the same containment unit, recursively, down to whatever grain of description is being used. The interaction relation, $B_i \rightarrow B_j$, answers a different question entirely: which containers affect which others' maintenance dynamics, regardless of whether either contains the other. The constitutive relation is hierarchical by construction. The interaction relation is not, and need not be — it is a directed graph, and nothing in the theory developed so far requires it to be acyclic. A trade relationship, a military alliance, a symbiosis, or any other case of reciprocal maintenance is naturally cyclic: $B_i \rightarrow B_j \rightarrow B_k \rightarrow B_i$, with each system contributing resources that, through the cycle, ultimately support itself.

These two relations were conflated in Section 5, where nesting was made to do double duty: it specified both what the tumor was made of and how the tumor related to its host. Separating them resolves a problem that the conflation otherwise produces. The dependency ratio introduced informally in earlier discussion,

$$\Delta(B) = \frac{W_{\text{external} \rightarrow B}}{W_B^{\text{total}}},$$

is only well-defined once “external” is given a fixed reference frame, and the constitutive relation supplies exactly that frame. Boundary work supplied by anything constitutively nested within B — an organ's contribution to the organism, a department's contribution to a university — is internal to B by definition and belongs to W_B^{total} , not to the numerator. Only work crossing the constitutive boundary, supplied by systems that are not part of B 's own nested structure, counts as external:

$$\Delta(B) = \frac{W_{\text{cross}}(B)}{W_{\text{total}}(B)}, \quad W_{\text{cross}}(B) = \sum_{B_i \rightarrow B, B_i \not\subset B} W_{B_i \rightarrow B}.$$

This definition is invariant under refinement of B 's internal structure. Subdividing a liver into hepatocytes, or a state into agencies, or a language community into households, adds detail entirely within the constitutive closure of B ; none of it crosses the boundary the ratio is computed against, so $\Delta(B)$ does not change merely because the model became more granular. This is the property the earlier, single-host version of Δ lacked, and it is the property that makes the quantity a feature of the system rather than an artifact of how finely it happens to be described.

The practical measurement of $\Delta(B)$ is necessarily domain-dependent, and the theory does not require a single operationalization to hold across every case it is applied to. In a cell or a state, externally supplied boundary work has a relatively direct empirical

signature — metabolic resources crossing a membrane, trade and financial flows crossing a border — and Δ can in principle be estimated from such flows. In a language or a corporation, the corresponding quantities are less directly measurable: what counts as externally supplied boundary work for a language is a matter of some interpretation, whether construed as loanwords, external prestige effects, or transmission occurring outside the core speaker community. The theoretical role of $\Delta(B)$ in this paper does not depend on a uniform measurement procedure across domains. It depends only on the conceptual distinction between boundary work generated within a system’s own constitutive closure and boundary work supplied across it — a distinction each domain can satisfy on its own terms, even where the corresponding numbers are harder to pin down.

The tumor case, reconsidered in this vocabulary, shows the two relations operating simultaneously without conflict. The tumor’s own cells are constitutively nested within it: $\text{cells} \subset \text{tumor}$. The tumor’s relationship to its host is an interaction edge, not a constitutive one: $\text{host} \rightarrow \text{tumor}$, supplying the overwhelming majority of the tumor’s boundary work and yielding $\Delta(\text{tumor}) \approx 1$. The hierarchy fixes what the tumor is. The graph describes how it survives. A metastasis that establishes itself in distant tissue is, in this vocabulary, a node whose interaction edges have shifted to a different region of the graph, with a correspondingly different, though not necessarily lower, dependency ratio.

The same vocabulary gives the nested-to-autonomous transition a continuous description it previously lacked. A startup leaving an incubator, a colony becoming independent, a university department becoming its own institution: none of these require a discrete jump from “nested” to “autonomous.” Each is describable as $\Delta(B)$ decreasing over time as interaction edges that once supplied most of B ’s boundary work are replaced by boundary work B generates or sources for itself. Autonomy, in this account, is not a distinct category a system either occupies or does not. It is the limiting regime

$$\Delta(B) \rightarrow 0,$$

reached asymptotically as external contribution becomes negligible relative to total maintenance — not because external relations vanish, which they generally do not, but because their contribution to containment becomes small enough that local analysis can treat B as self-sustaining. This is the reservoir analogy stated correctly: a cell depends on the organism, an organism on an ecosystem, a state on a trade network, and none of these dependencies disappear as $\Delta \rightarrow 0$. What changes is only whether the dependency is large enough to matter for the question being asked.

Ground-Level Approximation. One might ask whether the relational formulation ultimately requires a privileged lowest-level container, relative to which all other containment is defined — some final B that is not itself nested within anything, or some outermost H that contains everything without itself being contained. Nothing in the theory developed here requires such a ground floor. Containment relations may continue indefinitely across scales, with every system embedded within broader containment ecologies and composed, in turn, of narrower ones. The practical stopping point is not metaphysical but analytical: local analysis becomes adequate whenever external contributions are sufficiently small relative to total maintenance, which is precisely the regime $\Delta(B) \rightarrow 0$ already introduced above. The theory therefore does not identify a final container, in the way a foundationalist ontology might seek a basic unit beneath which nothing further needs explaining. It identifies the conditions under which further containment relations cease to matter for the inquiry at hand — a claim about when analysis may stop, not a claim about where reality does.

The cyclic case raises a final possibility worth making explicit, because it points toward a

phenomenon the theory can describe but has not yet named. A tightly reciprocal interaction cycle — $B_i \rightarrow B_j \rightarrow B_k \rightarrow B_i$, with high mutual dependency throughout — can constitute a stable, robust ecology as a whole even though every individual node within it shows a high Δ . Two city-states bound by sufficiently dense, sufficiently stable mutual support; two firms whose reciprocal dependency becomes total; two organisms whose symbiosis becomes obligate on both sides — in each case, no single node in the cycle is autonomous, yet the cycle as a unit may be. This is precisely the condition under which the constitutive boundary itself should be redrawn: when interaction edges within a cycle become dense and stable enough, the principled move is to apply Fuse, treating the cycle as a single higher-order constitutive unit rather than a set of mutually dependent ones. The fused system's own Δ , computed against the remainder of the graph, may then be low even though no constituent's Δ ever was. This gives the theory a formal account of a real and otherwise hard-to-describe phenomenon — separate systems coalescing into one, whether through political unification, corporate merger, or evolutionary symbiosis becoming permanent — as a consequence of machinery already in place, rather than as an addition to it.

The rock and the fad, once more, admit no meaningful dependency ratio at all. There is no maintenance ecology to map for either: no interaction edges to trace to a host or a neighborhood, no constitutive closure whose internal versus external maintenance could be compared, and consequently no Δ to compute. This is not a degenerate or trivial value of $\Delta(B)$ — not $\Delta = 0$ and not $\Delta = 1$ — but the absence of a well-defined quantity, exactly as the absence of M_i made $\text{Container}(B)$ ill-posed for them in Section 2. The graph picture developed across this section sharpens, rather than dissolves, the boundary the demarcation criterion first drew.

What Section 6 has shown, in sum, is that the question “is B autonomous” was never quite well-posed. The better question is where B sits in a containment ecology at a given moment: how its constitutive boundary is drawn, how dense and reciprocal its interaction edges are, and how large $\Delta(B)$ is relative to that boundary. Autonomy is a position in that space, approached but never strictly reached, rather than a property a system simply has.

7. Four Domains, One Diagnostic

The preceding sections developed the theory's machinery by moving freely among cells, languages, states, and corporations as needed. This section reverses the order. It holds the machinery fixed and asks the same question of each domain in turn:

When containment in this domain fails, which specific maintenance operation ceased, and through which of the failure modes established in Sections 2 through 5 did that cessation produce collapse?

The point of this discipline is to resist the pull of analogy. Four domains drawn from four separate literatures risk becoming four separate essays loosely organized under a shared vocabulary, each illustrating the theory rather than testing it. Holding the question fixed and varying only the domain is what distinguishes a single theory applied four times from four anecdotes wearing the same notation.

Cell. Maintenance, M_t , is ion-pump activity, membrane repair pathways, osmotic regulation, and protein trafficking — mechanisms identifiable by direct biochemical assay, independent of whether the cell in question happens to be dying. Identifiability is rarely the binding constraint here; cells either have these mechanisms or they do not, and their presence is easy to confirm. Efficacy failure looks like a pump that is present but non-functional — a mutated ion channel that is transcribed, translated, and embedded in

the membrane, yet does not move ions. Sufficiency failure looks like a fully functional membrane apparatus overwhelmed by an osmotic gradient or toxin load beyond what any feasible $g(M_t)$ could counter: $C_t > g(M_t)$ with M_t entirely intact. Reflexivity is where the cell case is most instructive, because it supplies the cleanest instance of Type I failure in the entire paper: autoimmune attack is M_t acting directly on I_t , the maintenance system having lost the capacity to distinguish self from non-self at the molecular level that the whole apparatus exists to police. When the binding fails at last, the result is not a degraded distinction between inside and outside. It is lysis — the abrupt, total cessation of the partition the maintenance activity had been continuously regenerating, exactly as Section 1 anticipated.

Language. Maintenance is schools, household transmission, literacy practice, religious use, and media production — institutions and practices identifiable independently of whether the language in question survives another generation, since they can be inventoried, funded, or defunded on their own terms. A language can fail at efficacy: a school system nominally teaching the language while producing no fluent speakers, ceremonial instruction with no functional transmission, is M_t present but $\partial W_t / \partial M_t \approx 0$. A language can fail at sufficiency: vigorous, effective transmission within a community too small, too dispersed, or too economically marginal to generate enough W_t to outweigh the prestige and utility advantages of a competing language, C_t in this case being driven largely by D_{ext} from the dominant language rather than by anything internal. Reflexivity appears here as Type II rather than Type I: prescriptivist defense of a language — corrective pressure, purism, rigid standardization — can provoke exactly the resistance and abandonment among younger or marginalized speakers that the defense was meant to prevent, $M_t \rightarrow D_t$ feeding back through the speaker community itself. A dead language, in this vocabulary, is not one whose words have vanished — vocabulary frequently survives in loanwords, place names, and liturgical fragments — but one in which the maintenance operation of intergenerational transmission has ceased, leaving words without the activity that reproduced the distinction between licit and illicit utterance from one generation to the next.

State. Maintenance is taxation, courts, administration, and enforcement, each independently auditable regardless of the state's ultimate survival. The ritual state of Section 3 — sovereignty performed without corresponding institutional reach — is the clearest efficacy failure in the paper, M_t fully identifiable and entirely inert. The overwhelmed state — functioning courts and administration defeated by simultaneous invasion, debt, and internal fragmentation — is the clearest sufficiency failure, $g(M_t) > 0$ throughout but outpaced by C_t . States also supply the paradigm Type II reflexivity case introduced in Section 4: policing that generates the unrest it suppresses, a security apparatus whose own intensification becomes a primary source of the destabilization it exists to contain. And states are the domain in which Section 5's relational turn is least avoidable, since a state's Δ is rarely small — dependence on trade networks, alliances, and financial systems is the rule rather than the exception, and a state's collapse is frequently best described not as M_t failing internally but as a previously load-bearing interaction edge in the surrounding graph being withdrawn.

Corporation. Maintenance is legal structure, operational coordination, payroll, and decision rights — again independently identifiable through audit, governance records, and organizational charts, regardless of whether the firm in question survives its next fiscal year. Efficacy failure looks like governance that exists on paper — a board, a compliance department, a stated chain of command — that does not in practice constrain operational decisions; the structure is identifiable but $\partial W_t / \partial M_t \approx 0$. Sufficiency failure looks like genuinely functioning governance overwhelmed by a market shift, a liquidity crisis, or competitive pressure no feasible internal reorganization could counter in time.

The corporation is also the domain that most directly recapitulates Section 5's tumor case without metaphor: a subsidiary or division can satisfy every criterion for Container(B) in isolation while $\partial W_{\text{parent}}/\partial M_{\text{subsidiary}} < 0$ — a unit that is, by every internal measure, well-maintained, while extracting resources from the parent faster than it returns value, the corporate analogue of $M_{B_2} = D_{\text{mis}}(B_1)$. When a corporation dissolves, the legal shell frequently persists for some interval after the maintenance that mattered — coordinated control over assets, liabilities, and personnel — has already ceased, which is the institutional version of the same lesson the cell offered first: containment can be gone before its outward signature is.

Run side by side, the four domains do not converge on a single failure mode; they converge on a single question. Identifiability is rarely what fails in any of the four — cells, languages, states, and corporations all generally possess inventoriable maintenance activity of some kind. What varies, case by case and instance by instance within each case, is whether that activity is efficacious, whether it is sufficient, and whether it has begun to work against the very interior or environment it was built to protect. The theory's contribution is not that it identifies these four domains as containers — that conclusion required none of the preceding machinery. Its contribution is that it gives a single, shared vocabulary in which the difference between a ritual failure, an overwhelmed failure, and a reflexive failure can be stated precisely, and tested, in any of the four.

7.5 Diagnostic Procedure

The preceding sections established a taxonomy of failure modes one at a time, in the order each was forced into existence by a problem with the one before it. Applied to an actual case, the taxonomy is more useful read as a sequence of questions than as a list of categories, since real collapses are diagnosed by elimination rather than by recognition. The following procedure makes that sequence explicit.

Given a system B believed to be failing or to have failed:

Step 1 — Identifiability. Does there exist a candidate maintenance process M_t specifiable independently of the outcome being explained? If no such process can be named — if every candidate offered turns out, on inspection, to be a redescription of the persistence or collapse itself — the system fails the Maintenance Identifiability Criterion of Section 2. B is not a container in the technical sense, and the remaining steps do not apply. This is the rock and the fad.

Step 2 — Efficacy. If M_t is identifiable, does it causally affect boundary work — does $\partial W_t/\partial M_t > 0$ over the system's relevant operating range, confirmable in principle by the interventionist test of Section 3.5? If M_t is real but inert, the diagnosis is efficacy failure. This is the ritual state: institutions present, audited, and genuine, but disconnected from the dynamics they were thought to govern.

Step 3 — Sufficiency. If M_t is both identifiable and efficacious, does the boundary work it generates meet or exceed the cost of separation — does $g(M_t) \geq C_t$ hold across the period in question? If maintenance is real and effective but outpaced, the diagnosis is sufficiency failure. This is the overwhelmed state: nothing was performative, the apparatus simply lost the race.

Step 4 — Reflexivity. If maintenance was identifiable, efficacious, and ostensibly adequate, did its own intensity nonetheless correlate with rising destabilization — either through damage to the protected interior itself ($M_t \rightarrow I_t$, Type I), or through provocation of the surrounding environment ($M_t \rightarrow D_{\text{ext}}$, Type II)? If so, the diagnosis is reflexivity failure, and the Type I/Type II distinction of Section 4 indicates whether the failure orig-

inated inside the system's own target selection or in its effect on what surrounds it. A useful supplementary check, given the threshold structure of Section 4.5, is whether M_t 's trajectory shows a turning point — a level past which further increases in maintenance intensity coincide with worsening rather than improving outcomes — which is the clearest available signature that an overmaintenance regime has been entered.

Step 5 — Relational failure. If none of the first four steps locates the failure within B 's own maintenance dynamics, the destabilization may not originate within B at all. Does B have a host or sit within a denser interaction neighborhood, per Sections 5 and 6, whose own maintenance dynamics were withdrawn, redirected, or were never sufficient to begin with — that is, did $\Delta(B)$ play the decisive role rather than M_t ? If so, the failure is relational: a property of B 's position in a containment ecology rather than of B considered alone, and the diagnosis properly belongs to the host or neighborhood, not to B 's own maintenance activity.

This procedure is diagnostic by elimination rather than by direct classification, and that order is not incidental. Identifiability is checked first because, if it fails, no further question about the system is well-posed. Efficacy and sufficiency are checked next because they are the failure modes a functionalist account already anticipates, and ruling them out first isolates the cases the present theory was built to add: reflexivity, where the very presence of effective, sufficient maintenance is no guarantee against collapse, and relational failure, where the cause of collapse is not locatable within the failing system at all. A theory whose diagnostic procedure could stop at Step 3 would be a more conventional theory of insufficient effort. That the procedure regularly needs Steps 4 and 5 is the paper's central empirical wager, to be tested case by case rather than assumed.

8. Conclusion

The four domains of Section 7 were chosen, in part, because identifiability is rarely their binding constraint. It is worth recalling, by contrast, the two negative cases that opened the paper's positive machinery in Section 2. The rock and the fad were not included as illustrations; they were included because the demarcation criterion needs cases where it actually does work, and a theory that only ever confirms its examples has not yet been tested by them. Across the rest of the paper, identifiability failure does not recur — not because it ceased to matter, but because every domain subsequently examined was chosen for being the right *kind* of object for the remaining three failure modes to be meaningfully asked about. The rock and the fad mark the boundary of the theory's own domain of applicability, and that boundary, fittingly, turned out to be one more thing that had to be maintained rather than simply asserted.

The argument of this paper sits closest to a tradition it has deliberately held at arm's length until now: the theory of autopoietic systems developed by Maturana and Varela, in which an organism is characterized not by its material composition but by the self-producing, organizationally closed process that continuously regenerates its own components and the boundary distinguishing it from its environment. The affinity is real and substantial enough to deserve sustained treatment rather than a closing gesture; it is taken up in full in Appendix A, including the specific respects in which the present account's relational and ecological apparatus offers one possible formalization of a problem the autopoietic tradition left largely open.

A second affinity, narrower but worth stating, runs through the cost term C_t introduced in Section 3 and never fully developed. Sustaining a boundary against diffusion, leakage, and decay is dissipative work in the literal thermodynamic sense: a container is a region

in which some local measure of disorder is held below its surroundings at the continuous cost of exporting disorder elsewhere, which is the organizing idea behind dissipative structures in far-from-equilibrium thermodynamics. This paper has used that connection only as a bridge, not as a foundation, and deliberately so. Containment, in the account developed here, explains a class of local thermodynamic anomalies; it is not explained by thermodynamics in turn. A fully worked-out account of C_t in genuinely physical terms — what it would mean to measure containment cost in entropy units for an institution rather than a cell — is left for separate treatment, and the present paper has tried not to claim more for this connection than its current state of development supports.

What remains, after the demarcation criterion, the failure taxonomy, the relational turn, and the four domains, is a single argument stated in five steps, each forced by the one before it. A distinction, however finely time-indexed, cannot on its own explain why a system continues to be the same kind of thing from one moment to the next; something has to do the work of making successive states answerable to one another. That work — maintenance, in the technical sense developed from Section 2 onward — explains persistence where it explains anything. But maintenance is not a single fact a system either has or lacks: it fails in structurally distinct ways, several of which a purely functionalist notion of “trying hard enough” cannot represent, the sharpest being a system whose maintenance turns, past some threshold, against the very thing it protects. And once failure is taken this seriously, a further fact becomes unavoidable, first in the tumor case and then throughout Section 6: a system’s maintenance activity cannot in general be evaluated by looking at that system alone, because what counts as maintenance from one vantage can be exactly what counts as destabilization from another, and the difference is a fact about position, not about either system considered in isolation.

The deepest consequence of the argument, and the one the preceding sections were built to earn rather than merely assert, is that containment is not ultimately a property of isolated systems at all. It is a property of positions within ecologies of mutual maintenance, destabilization, and dependency — ecologies in which a tumor’s maintenance is its host’s reflexive injury, in which a startup’s dependency ratio falls continuously rather than crossing some threshold into autonomy, in which two systems with no individual self-sufficiency at all can together constitute something stable enough to be fused into one. The failure taxonomy of Sections 3 through 4.5 is what makes this claim precise rather than evocative: it is only because efficacy, sufficiency, and reflexivity were each given independent, checkable content that the relational turn could show, rather than merely suggest, that the same activity can occupy different categories of the taxonomy depending on which system’s boundary work is taken as the reference frame. A thing persists, in the end, not because its contents remain fixed and not merely because some boundary encloses it, but because somewhere in the ecology it belongs to — sometimes within it, sometimes at its edge, sometimes nowhere a single observer would think to look — maintenance is still being done, traceably, never guaranteed, and never quite by the system alone.

Appendix A: Relation to Autopoietic Theory

The affinity between the present account and the theory of autopoiesis, introduced by Maturana and Varela in the early 1970s and given its canonical statement in *Autopoiesis and Cognition: The Realization of the Living* (Maturana and Varela 1980; with earlier formulations dating to the early 1970s and a formal model presented in Varela, Maturana, and Uribe 1974 — publication-history details are given in the references below), is close enough that the comparison deserves more than a closing gesture. This appendix states the affinity

precisely, identifies where the two accounts genuinely diverge, and considers what the present paper's relational and ecological apparatus might offer relative to a question the autopoietic tradition raises but, in the texts examined here, does not fully resolve.

The core affinity. An autopoietic system, in Maturana and Varela's formulation, is a network of processes of production that (i) regenerate, through their own interactions, the network that produces them, and (ii) constitute the system as a distinguishable unity in the space in which it exists, by continuously specifying its own boundary. A cell is the paradigm case: its membrane is produced by metabolic processes occurring within the boundary the membrane itself defines, in a closed, self-referential loop. The relevant organizing concept here, *organizational closure*, is sometimes presented as though it originated with Maturana and Varela. According to Bich and Damiano's (2012) account of the broader organizational-approach tradition, it is more accurately traced to Piaget, and autopoiesis is properly understood as one specific model within a wider twentieth-century research program that also includes Rosen's metabolism-repair systems, Weiss's hierarchically organized systems, and Gánti's chemoton — each proposing a different concrete mechanism for the same underlying claim, that a living system's identity is carried by an invariant network of relations rather than by any particular material configuration. (The present paper has verified Bich and Damiano's text directly; the primary sources they in turn cite for Piaget, Rosen, Weiss, and Gánti have not been independently checked here, and the attribution rests on Bich and Damiano's reading of them.) This distinction between organization and structure is close to the one drawn in Section 1 of the present paper between a maintained partition and a static one: both accounts reject the idea that a system can be identified with an inventory of its components at a given instant, and both locate identity in an ongoing process rather than in any configuration that process happens to produce at a particular time. The Boundary Primacy Theorem and the concept of organizational closure are, in this respect, close cousins — though the present paper's primitive is closer to the general organizational-approach commitment than to autopoiesis specifically, since nothing in Sections 1–2 requires the self-production of material components that autopoiesis, as one instance of the broader tradition, demands.

A direct precedent for Section 2.5. The question Section 2.5 was written to settle — whether a boundary must be spatial — turns out to be live and unresolved within the organizational-approach literature itself, rather than a problem this paper had to raise from outside the tradition. Bich and Damiano (2012) describe self-distinction, the capability of a living system to mark itself off from its environment, as a property the tradition agrees living systems must have while disagreeing sharply about its form, ranging “from a purely functional to a topological distinction.” They further note that even the paradigmatic topological case — the cell membrane as a continuous physical barrier — is itself contested in the literature, citing work questioning whether the cellular membrane constitutes the kind of continuous enclosure the classical picture assumes (Pollack 2001). This is a genuine point of contact rather than a borrowed analogy: the admissibility boundary of Section 2.5, defined in terms of which continuations are sorted as licit rather than in terms of any spatial enclosure, is one way of resolving a tension the organizational-approach tradition had already identified in its own foundational case, the cell, before this paper ever reached the question of languages or corporations.

This point needs an immediate qualification, however, once Maturana and Varela's own definition is consulted directly rather than through the broader tradition's later debates about it. Bich and Damiano (2007), summarizing the orthodox formulation, present autopoietic organization as having two necessary components: a network of processes that regenerates itself, and the constitution of the system as “a concrete unity in the space in which they exist, by establishing its boundary and thus specifying the topological domain

of its realization.” Topological boundary production is not optional or one reading among several in this formulation — it is stated as a defining condition, and the same paper goes on to criticize computational models that achieve only “topological closeness” without organizational closure as satisfying just one half of the definition, not the whole of it. The internal debate Bich and Damiano (2012) describe is therefore better read as a debate within the wider organizational-approach tradition about whether Maturana and Varela’s specific requirement should be relaxed, not as evidence that the orthodox theory itself was already agnostic about spatiality. Section 2.5’s admissibility boundary is, on this more precise reading, a more substantial departure from autopoiesis as originally formulated than the previous paragraph suggested — closer to a deliberate generalization past a condition the original theory insisted on, than to a resolution already implicit in it.

Where the scope diverges. Autopoiesis was developed, and has primarily been defended, as a criterion for *life* specifically — operational closure was offered as close to a necessary and sufficient condition for biological autonomy, and the theory’s strongest empirical and conceptual commitments are at the cellular level. Extending the framework beyond biology has always required additional argument that the original theory does not itself supply. Niklas Luhmann’s adaptation of autopoiesis to social systems (*Social Systems*, 1984; English translation 1995) is the relevant precedent here, and it is worth citing accurately rather than implying, as an earlier draft of this paper did, that Maturana and Varela’s own framework simply extends to institutions. Luhmann’s move was to identify communications, rather than biochemical components, as the self-reproducing elements of a social system, with society itself construed as operationally closed at the level of communication rather than of organisms. The present paper’s extension to languages, states, and corporations is closer in spirit to Luhmann’s move than to anything in Maturana and Varela’s own writing, though it differs from Luhmann’s in a respect worth stating plainly: this paper does not require that the elements of a system’s maintenance be of any particular kind — communicative, biochemical, administrative — only that they be independently identifiable and causally efficacious in the sense of Sections 2 and 3.5. The admissibility boundary of Section 2.5 is, in effect, an attempt to state the relevant generalization directly, rather than by analogy to a specific prior extension of autopoiesis to a specific further domain.

Where the present account adds genuine machinery: the nesting problem. An earlier draft of this conclusion claimed that the relational turn of Section 5 “has no clean analogue” in the autopoietic tradition, citing a claim about second- and third-order autopoietic systems that could not be verified against a primary source and has since been withdrawn from this appendix. What can be verified, directly, is more modest and arguably more useful. Weiss’s model of hierarchically organized living systems, as presented by Bich and Damiano (2012), addresses a relation of a similar shape to the one Section 6 develops: a whole constrains the degrees of freedom of its parts, and a deviation in the parts triggers a compensatory reaction at the level of the whole, so long as the destabilization remains below some stability threshold of the system as a unity. This is recognizably the same structure as Section 4’s reflexivity and Section 3’s sufficiency criterion, restated for the relation between a system and its own constituent parts.

What this model does not extend to, in the texts examined here, is the relation Section 5 was built to represent: not a part’s relation to the whole it constitutively belongs to, but the relation between two systems each capable, in principle, of standing on its own — a tumor and a host, a subsidiary and a parent, an organ and an organism considered as two autonomy claims rather than one. Weiss’s hierarchical model and its relatives in the organizational-approach tradition formalize the intra-system part-whole relation in some detail; nothing in the material reviewed here extends that formalization to the inter-system

case, where one nominally autonomous unity is nested within another and the question is not how the whole regulates its parts but how one whole's self-production bears on a second whole's. This is a narrower and more defensible claim than the one it replaces, limited to the specific texts engaged with rather than asserting a gap in the secondary literature on autopoiesis as a whole, which this appendix has not attempted to survey exhaustively. Section 5's continuous derivative classification of nested containment, the constitutive/interaction distinction of Section 6, and the Fusion criterion are offered as one way of extending hierarchical, part-whole thinking of the kind Weiss's model already supplies to the harder case of relations between systems that are not, in the first instance, parts of one another at all.

Where the present account differs in kind: the failure taxonomy. Autopoietic theory treats organizational closure largely as a condition a system either maintains or loses; loss of closure, for a living system, is death, and the theory's primary diagnostic resource is the closure/non-closure distinction itself. The present paper's taxonomy — identifiability, efficacy, sufficiency, and reflexivity, with reflexivity further divided into interior misidentification and environmental provocation and given a threshold structure in Section 4.5 — has no real counterpart in the autopoietic literature, which was not built to ask which of several structurally distinct things went wrong when closure failed, only whether it did. This is not a deficiency of autopoietic theory relative to its own aims; the theory was constructed to answer a different question, namely what distinguishes the living from the non-living, not to provide a diagnostic procedure for institutional or ecological collapse. But it does mean that Section 7.5's diagnostic procedure — and in particular its capacity to distinguish a system whose maintenance was merely insufficient from one whose maintenance became actively self-destructive — is genuinely new relative to the tradition it is closest to, rather than a relabeling of something the tradition already provided.

A documented limitation worth acknowledging directly. Autopoietic theory has, since its earliest reception, faced a charge of circularity: that the concept of an autopoietic unity is partly defined by reference to the very domain of description it claims to be independent of, since the boundary that constitutes the unity is specified by an observer using categories the theory does not itself ground (see, for instance, the contemporaneous review in *The Review of Metaphysics*, 1981, which raises this objection directly against the 1980 text). The Maintenance Identifiability Criterion of Section 2 was developed for an unrelated reason — to prevent the present theory's own central concept from being read backward off the outcome it explains — but it is worth noting that it also forecloses a version of the same circularity charge, since it requires that M_t be specifiable independently of containment outcomes, by a procedure that does not presuppose the very partition under investigation. Whether this fully answers the objection as leveled against autopoiesis itself is a question this paper does not attempt to resolve; the point made here is narrower, namely that the present framework's demarcation criterion was built, for its own purposes, in a way that happens to address a long-standing vulnerability in its nearest theoretical relative.

Summary comparison.

	Autopoietic theory / organizational approach	The present account
Primitive	Organizational closure (Piaget; instantiated by autopoiesis and related models)	Maintenance (identifiable, efficacious, sufficient activity)

	Autopoietic theory / organizational approach	The present account
Domain of original application	Living systems, paradigmatically the cell	No domain restriction; applied uniformly to cells, languages, states, corporations
Extension beyond biology	Via Luhmann, to communications in social systems	Built into the core definition via the admissibility boundary (Section 2.5)
Self-distinction / boundary	Topological for Maturana and Varela's own formulation; disputed within the wider organizational-approach tradition (Bich and Damiano 2007, 2012)	Generalized via the admissibility boundary: sorting of continuations, spatial enclosure as one instance among several
Part-whole and composite systems	Weiss's hierarchical model: parts constrained by, and regulated through, the whole they belong to	Constitutive/interaction distinction, extending part-whole regulation to relations between otherwise self-standing systems (Sections 5–6)
Failure	Binary: closure maintained or lost (death)	Four structurally distinct modes, with a threshold-governed reflexive regime (Sections 2–4.5)
Diagnostic procedure	Not a stated aim of the theory	Explicit, stepwise, by elimination (Section 7.5)
Vulnerability to circularity	Long-standing critique (observer-dependence of boundary specification)	Addressed structurally by the identifiability criterion, for independent reasons

The two accounts are not competitors so much as siblings with different parents. Autopoiesis arrived at the centrality of self-maintaining process from within theoretical biology, in service of a definition of life. The present account arrives at a structurally similar primitive from a different direction — the demand that a theory of boundaries explain its own negative cases — and, having arrived there, finds itself needing further machinery, a failure taxonomy and a treatment of relations between otherwise self-standing systems, that extends rather than contradicts the part-whole thinking already present in the organizational-approach tradition, applied here to a case that tradition's own primary texts, so far as they have been checked in preparing this appendix, do not appear to have worked out in comparable detail.

References (Appendix A). Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and Cognition: The Realization of the Living*. Dordrecht: D. Reidel. (Underlying formulations developed earlier: Varela, F., & Maturana, H. (1972). Mechanism and biological explanation. *Philosophy of Science*, 39(3), 378–382; Maturana, H., & Varela, F. (1973). *De Máquinas y Seres Vivos*. Santiago: Editorial Universitaria — cited here with the 1973 date given in Bich and Damiano's own reference lists, which differs by a year from some secondary sources consulted in preparing this appendix; the discrepancy has not been independently resolved. Formal model: Varela, F. G., Maturana, H. R., & Uribe, R. (1974). Autopoiesis: The organization of living systems, its characterization and a model. *BioSystems*, 5, 187–196.) Luhmann, N. (1984). *Soziale Systeme*. Frankfurt: Suhrkamp. (English translation: *Social Systems*, Stanford University Press, 1995.) Bich, L., & Damiano, L. (2007). Theoretical and artificial construction of the living: redefining the approach from an autopoietic point of

view. *Origins of Life and Evolution of Biospheres*, 37, 459–464. Bich, L., & Damiano, L. (2012). Life, autonomy and cognition: an organizational approach to the definition of the universal properties of life. *Origins of Life and Evolution of Biospheres*. Both Bich and Damiano papers were directly consulted and verified for this appendix. Citations to Piaget (1967), Rosen (1958, 1991), Weiss (1968), Gánti (2003), and Pollack (2001) above are drawn from these two papers' discussion of those sources and have not been independently checked against the primary texts themselves.