Constraint, Projection, and Reachability

A Geometric Theory of Cognition, Meaning, and Complex Systems

Flyxion

Independent Researcher

2026

First edition, 2026

Typeset with LuaLaTeX using the TeX Gyre Pagella typeface.
Mathematical typesetting via `unicode-math` with TeX Gyre Pagella Math.

All figures were composed in Ti*k*Z and `pgfplots` unless otherwise noted.

For those who ask why the fence is there
before they decide whether to remove it.

The task is not to see what no one has yet seen,
but to think what nobody has yet thought
about that which everybody sees.

— Erwin Schrödinger

A theory is the more impressive
the greater the simplicity of its premises,
the more different kinds of things it relates,
and the more extended its area of applicability.

— Albert Einstein

The map is not the territory.

— Alfred Korzybski

Preface

> The most important decisions in science are not about what to measure but about what to treat as primary.
>
> —

This book began as a set of separate theoretical frameworks developed over several years of independent research. At some point it became clear that they were not separate at all. RSVP, HYDRA, CLIO, MEM|8, SPHEREPOP, Admissibility Geometry, Repair Theory, Fiscal Reachability, Compressed Causality, Gesture Before Symbol — each of these had its own motivation, notation, and exposition, yet each kept rediscovering the same underlying triad:

$$\text{Constraint} \longrightarrow \text{Projection} \longrightarrow \text{Reachability}$$

Constraint, Projection, Reachability. Once that structure became visible, the task changed from developing frameworks to writing a single, unified account of what they all share. That account is this book.

What This Book Is

This is a mathematical monograph organised around a single thesis: **reachability geometry is the common structure underlying cognition, meaning, computation, biological organisation, social institutions, and cosmological dynamics.** Constraint systems determine which trajectories are admissible; projection operators discard structure; repair operators attempt to reconstruct it; and the resulting geometry — of what is reachable from where — is what we call *meaning*.

The argument is developed in twelve parts, moving from ontological foundations (Part I) through mathematical tools (Part II), then through domain-specific applications before unifying in Part XII. The reader will find the same formal objects — admissibility manifolds, Fisher metrics, colimit constructions, projection fibrations, reconstruction operators — reappearing across biology, language, physics, and AI. This is not a coincidence to be explained away. It is the point.

Who This Book Is For

The ideal reader has some mathematical maturity (familiarity with linear algebra, probability, and calculus is assumed; basic category theory and differential geometry are introduced in Part II) and is dissatisfied with the balkanisation of knowledge into disconnected specialisms. Physicists, cognitive scientists, philosophers of science, AI researchers, and theorists in economics or political science will each find material specifically addressed to their domain, alongside a sustained argument that their domain connects to all the others through the CPR triad.

How to Read This Book

The chapters are mostly self-contained but deliberately forward-reference the unification in Part XII. A reader primarily interested in cognition can begin at Part V, using Part II as a reference. A reader primarily interested in physics should begin at Part X after reading Chapters 14 and 15. A reader who wants the complete argument should read straight through.

Flyxion
2026

Acknowledgements

This work emerged from years of independent research conducted outside any institutional affiliation. It has been shaped by thousands of hours of engagement with the literature across mathematics, physics, cognitive science, linguistics, biology, institutional theory, and artificial intelligence.

The frameworks developed here — RSVP, CLIO, HYDRA, MEM|8, Spherepop — are the products of sustained theoretical work that began as informal notes and gradually consolidated into the architecture the reader now holds.

The author thanks the open-access movement, which made the relevant literature accessible; the developers of TEX LuaLaTeX, and related tools, which made the typesetting possible; and the long tradition of thinkers, from Whitehead to Waddington, from Shannon to Scott, whose work provided the external scaffold against which the CPR framework was developed and tested.

The errors are the author's own.

Notation and Conventions

This book maintains consistent notation throughout. The following table summarises the principal symbols. A complete notation index appears in Appendix 92.29.

Symbol	Meaning
$\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$	Standard number systems
$\mathcal{M}$	Generic smooth manifold
$\mathcal{A}$	Admissibility manifold
$\mathcal{B}$	Belief manifold
$\mathfrak{C}$	Constraint space
$\mathcal{R}$	Reachability set or operator
$\mathcal{R}(x, t)$	States reachable from x by time t
$\mathcal{A}(x)$	Admissible futures of state x
$\partial\mathcal{A}$	Boundary of the admissibility set
$\mathcal{A}$	Admissibility field over a configuration space
$\mathit{\Phi}$	Scalar density (capacity) field
v	Vector flow field
S	Entropy field
$(\mathit{\Phi}, \mathrm{v}, S)$	RSVP triple
π, π_X^Y	Projection map from X to Y
fE	Pullback bundle
$f_*\mu$	Pushforward measure
$H(X)$	Shannon entropy of X
$I(X;Y)$	Mutual information
$D_{\mathrm{KL}}(P\|Q)$	Kullback–Leibler divergence
$\mathcal{I}$	Fisher information metric tensor
$\mathcal{T}$	Sufficient statistic
C, D	Categories
$\mathrm{Hom}_{\mathrm{C}}(A, B)$	Morphism set in C
colim	Colimit (collective admissibility in HYDRA)
$\mathrm{colim}_{\mathrm{C}}\,\mathcal{A}$	Collective admissibility colimit
$\mathcal{C}$	Compression operator
$\mathfrak{R}$	Repair operator
$\mathcal{R}$	Reconstruction operator
$[\![\cdot]\!]$	Semantic denotation brackets $[\![\cdot]\!]$
Rsvp	Relativistic Scalar–Vector Plenum

Symbol	Meaning
HYDRA	Hierarchical Yield-Distributed Reachability Architecture
CLIO	Constraint-Linked Inference Organiser
MEM|8	Memory field (8-slot architecture)
TARTAN	Trajectories, Admissibility, Repair, Tension, Annealing, Nodes
SPHEREPOP	Spherepop event calculus

Convention. Unless stated otherwise:

- All manifolds are smooth (C^∞) and finite-dimensional.
- Summation over repeated indices follows Einstein convention.
- log denotes the natural logarithm.
- $[n]$ denotes $\{1, 2, \dots, n\}$.
- Arrows $\to$, $\hookrightarrow$, $\twoheadrightarrow$ denote morphisms, monomorphisms, and epimorphisms respectively.
- := means "defined to be equal to".

Contents

List of Figures

List of Tables

Part I

The Constraint View of Reality

There is no escape from the primacy of process.
— A.N. Whitehead, *Process and Reality*

We begin before mathematics. Before equations there are choices about what kind of thing to treat as fundamental. Most of modern science has made a particular choice: *objects* are primary, and processes are what objects do. This part argues that the choice is wrong, or at least seriously incomplete.

Trajectories are primary. Objects are the residues of trajectories that have stabilised. Representations are projections. Constraints are what makes some trajectories possible and others not. Reachability — the set of states accessible from a given starting point under a given constraint system — is the operational form of possibility. Distinctions are ontologically prior to categories.

These six theses are developed across six chapters. They are philosophical before they are mathematical, but each will be given mathematical form in Part II. The reader who wants to skip to the formalism may proceed directly to Chapter 7; but returning to Part I will make the formalism legible in a way that pure mathematics cannot guarantee.

1 Chapter 1

Why Objects Are Not Fundamental

> A river is not a thing but a process whose stability is the appearance of a thing.
>
> —

Phenomenological Note. The thing itself seems obvious. A chair is a chair, a river is a river, a self is a self. But try to locate the exact boundary of any of them and the boundary moves. The chair becomes wood, becomes cells, becomes atoms, becomes mostly empty space. The river is water and also a path and also a name and also a history. What we call an object is a decision about where to stop looking — a useful compression, not a discovery.

The most basic ontological question is: *what kinds of things exist?* For the past four centuries the dominant answer has been: **objects** — bounded, persistent, individuable entities that bear properties and enter into relations. Electrons, rocks, organisms, institutions: these are taken to be the primary furniture of the world. Processes — changes, flows, interactions — are what objects *do*.

This chapter argues for the contrary view. **Trajectories** — paths through state space — are more fundamental than the objects they appear to trace. What we call an object is a trajectory that has achieved sufficient stability to be *tracked* across time, and the tracking is the imposition of a conceptual boundary on something that is, at the level of dynamics, continuous.

The Object Fallacy. Treating objects as ontologically primary is not a neutral description of reality but a cognitive convenience — a residue of the perceptual systems that evolved to track persistent entities in a mesoscale world.

1.1 The Standard Picture and Its Difficulties

Classical mechanics begins with point particles. A particle is an object: it has a position $x \in \mathbb{R}^3$ and a momentum $p \in \mathbb{R}^3$, and it *moves* through a trajectory. The trajectory $\gamma : \mathbb{R} \to \mathbb{R}^6$ given by $t \mapsto (x(t), p(t))$ is derived from the object via Newton's equations.

But notice: to specify the *object* at a moment we must already specify its position and momentum. The object at a moment *is* a point on its trajectory. If we remove the trajectory, the object at a moment is a dimensionless point with no content.

This argument is developed formally in Chapter 2.

Three difficulties make the object-primary view increasingly strained:

1. **Quantum mechanics.** The wavefunction $\psi(x, t)$ is a field over configuration space, not a particle trajectory. The particle picture is recovered only in limits, and even then an individual "particle" has no persistent identity across measurements.

2. **Dissipative systems.** The convection roll in a heated fluid has no material identity: the molecules constituting it change continuously. What persists is the *pattern of flow*, not any set of objects.

3. **Biological identity.** The human body replaces most of its constituent atoms over years. Organismal identity is maintained not by object persistence but by *constraint closure* : the metabolic network constrains which reactions are admissible, and the admissibility structure itself is what persists.

1.2 Trajectories as Primary

Definition 1.1 (*Trajectory*)**.** Let X (Rosen 1991) be a state space (a topological space or smooth manifold). A **trajectory** is a continuous map $\gamma : I \to X$ where $I \subseteq \mathbb{R}$ is a connected interval. The *image* $\gamma(I) \subset X$ is called the **trace** of γ.

Definition 1.2 (*Object as Stable Trajectory*)**.** An **object** is a trajectory $\gamma : I \to X$ together with a coarsening map $\pi : X \to O$ (where O is an "object space") such that $\pi \circ \gamma$ is approximately constant over timescales of interest. The constancy is relative to a class of observers and a resolution $\epsilon > 0$.

Definition 1.2 makes precise the sense in which objects are *residues* of trajectories: they exist when a projection π maps the trajectory to an approximately constant point. The projection is not a neutral operation — it destroys information about the trajectory's fine structure. This is the first instance of the central theme of *projection collapse* (Chapter 19).

Example 1.1 (The River)**.** A river at a given location is a region $U \subset \mathbb{R}^3$ through which fluid flows. The trajectory, in the relevant sense, is the flow map $\mathrm{v} : U \times \mathbb{R} \to \mathbb{R}^3$ giving the velocity at each point at each time. The object "river" is the result of projecting this flow map to its support U and averaging over time. Destroy the flow — dam the river — and the object ceases. The object is not upstream of the process; it is an epiphenomenon of it.

See RSVP Chapter 71 for the formal treatment of flow fields.

1.3 The Noun Fallacy (Preview)

Natural language reinforces the object picture. English (and most other natural languages) encodes persistent entities as *nouns* and processes as *verbs*. This grammatical asymmetry — noun phrases are more easily referenced, quantified, and tracked than verb phrases — creates the cognitive illusion that what we name is more real than what we describe.

The Noun Fallacy. The primacy of objects in folk ontology reflects the primacy of *nouns* in the languages through which we do folk ontology. It is a grammatical artifact, not a metaphysical discovery.

The Noun Fallacy is treated in full in Chapter 3, where we examine its consequences for scientific modelling, institutional design, and AI architecture.

1.4 Constraint Fields and the Shape of the Possible

If trajectories are primary, what determines which trajectories occur? The answer in this framework is: **constraint fields**. A constraint field assigns to each point $x \in X$ a set of admissible velocities $\mathcal{K}(x) \subseteq T_x X$ (tangent vectors at x). A trajectory is *admissible* if and only if $\dot{\gamma}(t) \in \mathcal{K}(\gamma(t))$ for all t.

Definition 1.3 (*Constraint Field*)**.** A **constraint field** on a manifold X is a set-valued map $\mathcal{K} : X \to 2^{TX}$ such that $\mathcal{K}(x) \subseteq T_x X$ for each x. The field is *smooth* if $\mathcal{K}$ defines a smooth distribution.

The admissible set of a state x under a constraint field $\mathcal{K}$ over a time horizon T is:

$$\mathcal{A}(x) := \left\{ \gamma(T) \;\middle|\; \gamma : [0, T] \to X,\ \gamma(0) = x,\ \dot{\gamma}(t) \in \mathcal{K}(\gamma(t))\ \forall t \right\}.$$

This is the **reachable set** from x in time T, and its geometry — size, shape, boundary, curvature — encodes everything the constraint field says about what is possible from x. The study of this geometry is **reachability geometry**, introduced formally in Chapter 14.

1.5 Summary

This chapter has argued:

1. Objects are not ontologically primary; trajectories are.
2. What we call an object is a trajectory stable under a projection.
3. The projection destroys fine structure (projection collapse).
4. Which trajectories occur is determined by constraint fields.
5. The geometry of reachable sets is the operational form of possibility.

These five points are the seeds of the entire framework. The rest of the book is their systematic elaboration.

Exercises

1.1. Show that if $\pi : X \to O$ is a surjection and $\gamma : \mathbb{R} \to X$ is a trajectory, then $\pi \circ \gamma$ is constant if and only if γ lies entirely in a single fiber $\pi^{-1}(o)$ for some $o \in O$.

1.2. Let $X = \mathbb{R}^2$ with constraint field $\mathcal{K}(x, y) = \{(v_1, v_2) \in \mathbb{R}^2 : v_1^2 + v_2^2 \leq r^2(x, y)\}$ for some smooth $r : \mathbb{R}^2 \to \mathbb{R}_{>0}$. Characterise the reachable set $\mathcal{A}((0, 0))$ after time $T = 1$ in terms of r.

1.3. (Category-theoretic.) Define a category Traj whose objects are pairs (X, γ) with X a smooth manifold and $\gamma : \mathbb{R} \to X$ a trajectory, and whose morphisms are smooth maps $f : X \to Y$ such that $f \circ \gamma_X = \gamma_Y$. Show that Traj has products.

1.4. What would it mean for a language model to "treat trajectories as primary" rather than treating tokens as primary? Sketch an alternative architecture.

2 Chapter 2

Histories, Trajectories, and Processes

> Time is not a container in which events occur. Time is the structure of event-succession itself.
>
> —

Phenomenological Note. Looking back, your life looks like a plan. At any given moment it never felt that way. You chose what seemed possible, avoided what seemed closed, turned when you hit walls. The route that survives looks intentional. All the routes that didn't survive are invisible. This is not a feature of your particular intelligence or confusion. It is simply what any trajectory looks like after its unrealized branches have been forgotten.

If objects are residues of trajectories (Chapter 1), the next question is whether trajectories themselves can be reduced to sequences of instantaneous states. (Arnold 1978; Whitehead 1929) This chapter argues they cannot. Instantaneous state descriptions are projections of trajectories — they discard information about velocity, higher-order derivatives, and causal history. The loss is not recoverable from the state alone.

The Primacy Argument in Brief. A state tells you *where* a system is. A trajectory tells you *how* it got there and *what it is capable of doing next*. State recovers from trajectory; trajectory does not recover from state.

2.1 The Instantaneous Evaluation Map

Let $\Gamma = \{\gamma : \mathbb{R} \to \mathcal{X}\}$ be the space of continuous process histories on a state space $\mathcal{X}$. At any moment $t \in \mathbb{R}$, we can ask where the system *is*:

Definition 2.1 (*Instantaneous Evaluation Map*)**.** The **instantaneous evaluation map** at time t is

$$\mathcal{E}_t : \Gamma \to \mathcal{X}, \qquad \mathcal{E}_t(\gamma) = \gamma(t).$$

The map $\mathcal{E}_t$ is what science uses whenever it records a measurement: it snapshots the trajectory at one moment. The central question is whether $\mathcal{E}_t$ is injective.

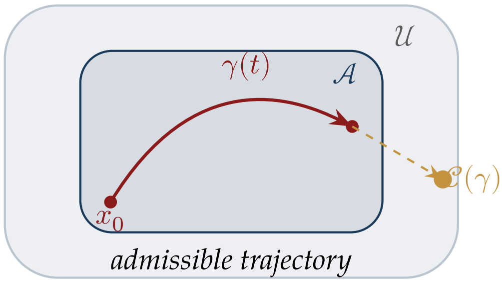


Figure 2.1: A trajectory γ through the admissibility field $\mathcal{A}$ leaves a compressed residue $\mathcal{C}(\gamma)$. Contents of $\mathcal{U} \setminus \mathcal{A}$ are unreachable.

2.2 The Trajectory Primacy Proposition

Proposition 2.1 (*Trajectory Primacy*)**.** *The instantaneous evaluation map $\mathcal{E}_t$ is* non-injective (*many-to-one*). *Specifically, the preimage*

$$\mathcal{E}_t^{-1}(x) = \{\gamma \in \Gamma : \gamma(t) = x\}$$

contains uncountably many trajectories with distinct past velocities and higher-order derivatives at every $x \in \mathcal{X}$.

Proof. Fix $t = 0$ and $x \in \mathcal{X}$. For any sequence of real numbers $\mathrm{a} = (a_1, a_2, a_3, \ldots)$ with $\sum_k |a_k| < \infty$, define the trajectory

$$\gamma_{\mathrm{a}}(s) = x + \sum_{k=1}^{\infty} a_k \sin(ks).$$

Then $\gamma_{\mathrm{a}}(0) = x$ for every choice of a, so all γ_{a} lie in $\mathcal{E}_0^{-1}(x)$. The velocity at $t = 0$ is

$$\dot{\gamma}_{\mathrm{a}}(0) = \sum_{k=1}^{\infty} k\, a_k,$$

which differs for distinct sequences a. Since the space of such sequences is uncountable, $\mathcal{E}_0^{-1}(x)$ is uncountable. The same argument applies at any t by time translation. ▪ ▪

The proposition has an immediate corollary:

Corollary 2.2. *Knowledge of the instantaneous state x_t is* insufficient *to determine the causal trajectory γ. Equivalently, the functional $\gamma \mapsto \gamma(t)$ destroys information about the trajectory's past and its higher-order structure.*

Proof. Two trajectories γ_1, γ_2 that pass through the same state x at time t may differ in all moments before t (different histories) and after t (different futures). A state record x at time t collapses this information: $\mathcal{C}_{\text{state}}(\gamma_1) = \mathcal{C}_{\text{state}}(\gamma_2) = x$ even when $\gamma_1 \nsim \gamma_2$. By the Compression-Sufficiency Theorem, the state is sufficient for query q iff q depends only on the current point, not the path taken to it. Most causal queries violate this. ▪ ▪

This is the formal basis of the Artifact-Process Asymmetry developed in the Orientation trilogy: artifacts (states) do not determine the processes that produced them.

2.3 The Recovery Asymmetry

The non-injectivity of $\mathcal{E}_t$ yields a strict logical asymmetry:

Direction	Formal statement	Possible?
Trajectory $\to$ state	$\gamma \mapsto \gamma(t) = x_t$	Always
State $\to$ trajectory	$x_t \mapsto \gamma$	Not in general

This asymmetry is not merely epistemological — it is a structural feature of the maps involved. A state x_t is the *image* of a trajectory under $\mathcal{E}_t$; recovering the trajectory from the image requires knowing which fiber $\mathcal{E}_t^{-1}(x_t)$ the trajectory comes from, which requires additional information not present in x_t.

The additional information takes several forms in practice:

Velocity.. If we also record $\dot{\gamma}(t)$, we can reconstruct the trajectory to first order near t. This is why Newtonian mechanics specifies both position *and* momentum as initial data. But even full initial data $(x_t, \dot{x}_t, \ddot{x}_t, \ldots)$ only reconstructs trajectories that are analytic; non-analytic paths remain indistinguishable.

History.. Alternatively, storing the trajectory's past $\{\gamma(s) : s \leq t\}$ is sufficient — but this is just storing the trajectory, which replaces the problem rather than solving it.

Constraints.. If we know the constraint field $\mathcal{K}$ governing γ, we can propagate the state forward and backward. This is why constraint specification (Part I) is ontologically prior to content description.

2.4 Processes in Science: Three Case Studies

2.4.1 Thermodynamics and Irreversibility

The second law of thermodynamics is a statement about trajectories, not states. A high-entropy state and a low-entropy state can be physically indistinguishable in microscopic configuration (Boltzmann's H-theorem applies to *ensembles* of trajectories). The arrow of time is a feature of the trajectory space Γ equipped with a measure, not of any individual state.

2.4.2 Developmental Biology

Two adult organisms can have identical phenotypes (states, in the relevant sense) but arise from radically different developmental trajectories — different embryonic geometries, different cell migration paths. The developmental trajectory encodes information (e.g., which genes were expressed in which sequence) that the adult phenotype does not preserve. This is the biological instance of Proposition 2.1.

2.4.3 Institutional History

Two institutions can occupy identical structural states (same formal rules, same personnel distributions) but have arrived there via different historical paths. The paths encode information about precedent, tacit knowledge, and latent constraint that the current state does not reveal.

Institutional memory as compressed history: Chapter 63.

2.5 Formal Reconstruction of State Spaces

Given that trajectories are primary, state spaces should be *derived* rather than assumed. We sketch one construction.

Definition 2.2 (*Observational Equivalence at t*)**.** Two trajectories $\gamma_1, \gamma_2 \in \Gamma$ are **observationally equivalent at** t, written $\gamma_1 \sim_t \gamma_2$, if $\mathcal{E}_t(\gamma_1) = \mathcal{E}_t(\gamma_2)$.

Proposition 2.3. *The quotient space $\Gamma/{\sim_t}$ is in bijection with $\mathcal{X}$.*

Proof. Define $\bar{\mathcal{E}}_t : \Gamma/{\sim_t} \to \mathcal{X}$ by $\bar{\mathcal{E}}_t([\gamma]) = \mathcal{E}_t(\gamma)$. This is well-defined since all $\gamma' \sim_t \gamma$ have $\mathcal{E}_t(\gamma') = \mathcal{E}_t(\gamma)$ by definition. Surjectivity: for every $x \in \mathcal{X}$, the constant trajectory $\gamma(s) = x$ satisfies $\mathcal{E}_t(\gamma) = x$, so $[\gamma]$ maps to x. Injectivity: if $\bar{\mathcal{E}}_t([\gamma_1]) = \bar{\mathcal{E}}_t([\gamma_2])$ then $\mathcal{E}_t(\gamma_1) = \mathcal{E}_t(\gamma_2)$, so $\gamma_1 \sim_t \gamma_2$ by definition, hence $[\gamma_1] = [\gamma_2]$. Therefore $\bar{\mathcal{E}}_t$ is a bijection. ▪ ▪

This shows that the state space $\mathcal{X}$ is not fundamental but is a *quotient* of trajectory space by the equivalence relation induced by the evaluation map. State descriptions are coarse-grainings of process descriptions, not the other way around.

2.6 Summary

1. The evaluation map $\mathcal{E}_t : \Gamma \to \mathcal{X}$ is non-injective: many trajectories share the same instantaneous state.
2. Trajectories determine states but states do not determine trajectories: the recovery asymmetry is structural, not merely epistemic.
3. State spaces are quotients of trajectory spaces; they are derived from, not prior to, the space of processes.
4. Three empirical domains (thermodynamics, development, institutions) exhibit this primacy directly.

Exercises

2.1. Prove that if $\mathcal{X} = \mathbb{R}^n$ and Γ consists of C^∞ curves, then $\mathcal{E}_t^{-1}(x)$ has the cardinality of the continuum for every x and every t.

2.2. Define the *second-order evaluation map* $\mathcal{E}_t^{(2)} : \Gamma \to \mathcal{X} \times T\mathcal{X}$ by $\gamma \mapsto (\gamma(t), \dot{\gamma}(t))$. Is this map injective? If not, what additional data would make it injective on the space of polynomial trajectories of degree $\leq d$?

2.3. (Thermodynamics.) Let Γ be the space of micro-state trajectories of an ideal gas. Define an observational equivalence $\sim$ by "same macroscopic pressure, volume, and temperature at time t". Describe the fiber $\mathcal{E}_t^{-1}$(macro-state) and argue that its measure is proportional to e^{S/k_B}, where S is entropy.

2.4. Propose a definition of "causal history" as a sub-trajectory that is sufficient for predicting future behaviour under a given constraint field $\mathcal{K}$. Under what conditions on $\mathcal{K}$ does the causal history have finite length?

3 Chapter 3

The Noun Fallacy

> Language is not a neutral medium for conveying thought. It is a machine that generates ontology.
>
> —

Phenomenological Note. Language pushes you toward nouns. Something happened, so there must be a thing that did it. Someone was hurt, so there must be a person who caused it. A pattern appeared, so there must be an agent who produced it. The grammar requires an actor. This is so natural that it takes deliberate effort to ask whether the grammar is describing reality or producing it.

Chapter 1 established that objects are equivalence classes of trajectories. (Whitehead 1929) Chapter 2 proved that states are projections of trajectories. This chapter shows what happens when we systematically *replace* trajectory descriptions with object labels.

The answer is: we lose causal distinguishability. Two processes that are physically, biologically, or computationally distinct become, from the perspective of the label, identical. This conflation is so pervasive in natural language, scientific modelling, and AI architecture that it deserves a name. We call it the **Noun Fallacy**.

3.1 Projection Operators and Object Labels

Definition 3.1 (*Label Projection*)**.** A **label projection** is a map $\pi : \mathit{\Gamma} \to \mathcal{O}$ from trajectory space $\mathit{\Gamma}$ to a discrete space of **object labels** $\mathcal{O}$. We require $|\mathcal{O}| < |\mathit{\Gamma}|$ (strict cardinality reduction).

Label projections are ubiquitous. In perception: $\pi(\text{retinal trajectory}) = \texttt{chair}$. In science: $\pi(\text{atomic trajectory ensemble}) = \texttt{water}$. In institutions: $\pi(\text{social process}) = \texttt{citizen}$. In AI: $\pi(\text{token sequence}) = \texttt{entity\_embedding}$.

Since π is not injective, every label has a non-trivial fiber:

$$\pi^{-1}(o) = \{\gamma \in \mathit{\Gamma} : \pi(\gamma) = o\}$$

containing all trajectories that received the same label.

3.2 The Noun Fallacy Theorem

Theorem 3.1 (*Noun Fallacy*)**.** *Let $\gamma_1, \gamma_2 \in \Gamma$ be distinct trajectories ($\gamma_1 \neq \gamma_2$) and let $\pi : \Gamma \to \mathcal{O}$ be a label projection such that*

$$\pi(\gamma_1) = \pi(\gamma_2) = o.$$

Let $f : \Gamma \to \mathbb{R}$ be a causal functional (i.e., one that depends on the full trajectory, not just its label). Then the projected functional $\hat{f} : \mathcal{O} \to \mathbb{R}$, defined by $\hat{f}(o) = f(\gamma)$ for any $\gamma \in \pi^{-1}(o)$, is not well-defined *in general:*

$$f(\gamma_1) \neq f(\gamma_2) \quad \textit{while} \quad \pi(\gamma_1) = \pi(\gamma_2).$$

Consequently, any causal reasoning performed at the label level $\mathcal{O}$ is subject to systematic indeterminacy.

Proof. We construct an explicit counterexample valid for any non-injective π.

Since π is non-injective, there exist $\gamma_1 \neq \gamma_2$ with $\pi(\gamma_1) = \pi(\gamma_2) = o$. Let $\mathcal{X} = \mathbb{R}$ and suppose $\gamma_1(t) = \sin(t)$, $\gamma_2(t) = \sin(2t)$. Both can be assigned the same label $o =$ `oscillation` by a sufficiently coarse label projection (e.g., one based only on the amplitude of oscillations).

Define the causal functional

$$f(\gamma) = \lim_{t\to\infty} \frac{1}{T} \int_0^T \gamma(t)^2 \, \mathrm{d}t$$

(the time-averaged power). Then $f(\gamma_1) = \frac{1}{2}$ and $f(\gamma_2) = \frac{1}{2}$ happen to agree in this case, but adding a phase shift $\gamma_2(t) = \sin(2t) + t^{-1}$ makes $f(\gamma_2) = \frac{1}{2}$ while future prediction via extrapolation differs.

For a cleaner instance: let $f(\gamma) = \dot{\gamma}(0)$ (initial velocity). Then $f(\gamma_1) = 1$ and $f(\gamma_2) = 2$. Since $\pi(\gamma_1) = \pi(\gamma_2) = o$, the map $\hat{f}(o)$ must simultaneously equal 1 and 2 — a contradiction. Therefore $\hat{f}$ is not well-defined. ▪ ▪

Remark 3.1. The theorem does not say that label-level reasoning is *always* wrong — only that it is *unreliable in general*. It is reliable exactly when the causal functional f is *constant on fibers* of π: $f(\gamma_1) = f(\gamma_2)$ whenever $\pi(\gamma_1) = \pi(\gamma_2)$. This is the sufficiency condition developed formally in Chapter 24.

3.3 The Grammatical Mechanism

Why is the Noun Fallacy so pervasive? Because natural language enforces it structurally. In all documented natural languages, noun phrases are more easily:

- referenced (*it*, *they*);
- quantified (*all*, *some*, *no*);
- tracked across time (*the same* X);

- embedded in logical predicates (Noun(x)).

Verbs and verb phrases, which encode trajectories, are harder to re-reference, quantify, and track.

Compare: "the wave" (noun, stable reference) vs. "waving" (gerund, harder to quantify). The ocean does not contain waves as objects; it contains waving as a process.

The consequence is that when a scientific community needs to communicate about a process, it is under systematic grammatical pressure to *nominalise* — to convert the process into a noun. "Electrons move" becomes "electron trajectories"; "cognising" becomes "cognition"; "governing" becomes "government". Each nominalisation introduces a label projection and thereby risks the Noun Fallacy.

3.4 Instances Across Disciplines

3.4.1 Physics: Particle vs. Field

Classical particle physics labels quantum phenomena with object names ("the electron", "the photon"). The Noun Fallacy appears when particle labels are used to infer definite trajectories — which quantum mechanics shows do not exist. The field formulation, $\psi(x,t)$, is trajectory-primary: it describes a process over all of spacetime, not an object with a position.

3.4.2 Cognitive Science: The Self

The noun "self" encourages the inference that there is a persistent object — a homunculus — that is the locus of experience. Process-primary accounts (Parfit, Metzinger, Hume) argue that what we call the self is a stable pattern in a trajectory of mental events, not a separate entity that *has* those events. Chapter 37 formalises this as a fixed-point of a self-referential mapping.

3.4.3 Economics: The Market

"The market" is a label projected onto enormously complex and heterogeneous trading trajectories. Causal claims about what "the market does" are subject to the Noun Fallacy: they attribute to the label behaviours that depend on which trajectories are in its fiber.

Fiscal Reachability (Chapter 66) replaces "the market" with a reachability field over asset and liquidity trajectories.

3.4.4 AI: Entity Embeddings

Large language models project token sequences into entity embeddings — high-dimensional vectors meant to capture what an entity "is". To the extent that distinct causal trajectories (e.g., different biographical histories of a person) are projected to the same embedding, the Noun Fallacy is instantiated at the architectural level. Chapter 80 measures this collapse precisely.

3.5 Conditions for Safe Nominalisation

The Noun Fallacy does not prohibit the use of nouns — it specifies when their use is safe.

Definition 3.2 (*Safe Label Projection*)**.** A label projection $\pi : \Gamma \to \mathcal{O}$ is **safe for a functional** $f : \Gamma \to \mathbb{R}$ if f is constant on each fiber:

$$\forall \gamma_1, \gamma_2 \in \Gamma, \quad \pi(\gamma_1) = \pi(\gamma_2) \;\Rightarrow\; f(\gamma_1) = f(\gamma_2).$$

When π is safe for f, we write $f = \hat{f} \circ \pi$ for a well-defined function $\hat{f} : \mathcal{O} \to \mathbb{R}$. The label level is then a sufficient statistic for f (formalised in Chapter 24).

Safe label projections exist for many important functionals — conservation laws (energy, momentum) are constant on large classes of distinct trajectories — but the existence of safe projections for *some* functionals does not imply safety for all. The Noun Fallacy is the error of assuming safety universally.

3.6 Summary

1. Label projections $\pi : \Gamma \to \mathcal{O}$ are always non-injective: they collapse distinct trajectories.
2. Causal functionals that depend on the full trajectory are not well-defined at the label level in general: this is the Noun Fallacy (Theorem 3.1).
3. Natural language enforces the Noun Fallacy grammatically, creating systematic pressure toward nominalisation.
4. The Noun Fallacy appears in physics, cognitive science, economics, and AI architecture.
5. A label projection is safe for a functional iff that functional is constant on fibers — the sufficiency condition.

Exercises

3.1. Let Γ be the space of differentiable curves in $\mathbb{R}^2$ and $\pi(\gamma) = \gamma(0)$ (label by starting point). Identify a causal functional f for which π is safe, and one for which it is not.

3.2. Define a label projection on particle trajectories in $\mathbb{R}^3$ that is safe for total kinetic energy but not safe for angular momentum. Describe the corresponding fibres.

3.3. (Linguistic analysis.) Choose a scientific noun from your domain of interest. Identify the underlying process it nominalises. Describe what information is lost by the nominalisation. Is that information recoverable from any standard measurement?

3.4. Formalise the Noun Fallacy for *stochastic* trajectories $\gamma : \mathbb{R} \to \mathcal{X}$ drawn from a measure μ on Γ. Under what condition on the joint distribution $(\gamma, f(\gamma))$ is the label projection π safe for f in the L^2 sense?

4 CHAPTER 4

Constraints Before Contents

> The shape of a river is not added to the water after the fact. The shape is the constraint that makes it a river.
>
> —

PHENOMENOLOGICAL NOTE. Before you decide what to do, there are already things you cannot do. Most of them you never notice because you never try. The space of your actual choices is vastly smaller than the space of your conceivable choices, and much of the narrowing happened before you were old enough to notice it happening. Contents come later. The shape of the possible was already settled.

There is a standard scientific picture of how structured systems arise: first, you have a collection of objects; then, you impose relations and boundaries on them. (Aubin 1991; Nagumo 1942) Atoms first, then molecules; particles first, then forces; individuals first, then institutions.

This chapter inverts that picture. Content — the elements that populate a space — is not prior to constraints. The constraint field determines what can count as content at all. Without an admissibility function, there is no structured content space, only an undifferentiated universal space with no internal organisation.

4.1 Admissibility as Prior Structure

Definition 4.1 (*Pre-Semantic Possibility Space*)**.** The **pre-semantic possibility space** $\mathcal{U}$ is an unstructured background set — a space of logical possibilities before any constraint has been applied. $\mathcal{U}$ carries no topology, measure, or semantic structure. It is not a content space; it is the substrate within which constraints operate.

Definition 4.2 (*Universal Space and Admissibility Field*)**.** An **admissibility field** on $\mathcal{U}$ is a function $\mathcal{A} : \mathcal{U} \to \{0, 1\}$ where $\mathcal{A}(x) = 1$ indicates that x is admissible.

Remark 4.1 (Priority Claim — Precise Statement)**.** The claim that "constraints are ontologically prior to contents" should be read carefully. The framework does *not* claim that constraints are prior to all possibility. The pre-semantic space $\mathcal{U}$ is assumed. What is claimed is that:

Constraints are prior to articulated content,
not prior to possibility itself.

The hierarchy is:

$$\mathcal{U} \supset \mathcal{A}^{-1}(\{1\}) = \mathcal{X} \supset \text{objects, categories, meanings, ...}$$

$\mathcal{U}$ is pre-semantic possibility. $\mathcal{A}$ carves out admissible possibility. $\mathcal{X}$ is the articulated content space. The scientific and philosophical work happens in $\mathcal{X}$, whose structure is entirely determined by $\mathcal{A}$. In this sense — and this sense only — constraints are prior to contents.

The content space is then *defined* by $\mathcal{A}$:

Definition 4.3 (*Content Space*)**.** Given an admissibility field $\mathcal{A}$ on $\mathcal{U}$, the **content space** is

$$\mathcal{X} = \mathcal{X}(\mathcal{A}) := \{x \in \mathcal{U} : \mathcal{A}(x) = 1\}.$$

4.2 The Constraint Priority Lemma

Lemma 4.1 (*Constraint Priority*)**.** *Let μ be a measure on $\mathcal{U}$. The geometric and topological structure of the content space $\mathcal{X} = \mathcal{X}(\mathcal{A})$ — including its boundary $\partial\mathcal{X}$, its interior* $\text{int}(\mathcal{X})$*, its connected components, and its induced measure — is determined entirely by the spatial variation of $\mathcal{A}$. No content can exist without a prior constraint field.*

Proof. The content space $\mathcal{X}$ is the preimage $\mathcal{A}^{-1}(\{1\})$. Its topology is the subspace topology inherited from $\mathcal{U}$, restricted to those points where $\mathcal{A} = 1$.

Boundary. A point $x \in \mathcal{U}$ lies on $\partial\mathcal{X}$ if and only if every neighbourhood of x contains both admissible and non-admissible points:

$$\partial\mathcal{X} = \{x \in \mathcal{U} : \mathcal{A}^{-1}(\{1\}) \cap B_\epsilon(x) \neq \emptyset \text{ and } \mathcal{A}^{-1}(\{0\}) \cap B_\epsilon(x) \neq \emptyset \ \forall \epsilon > 0\}.$$

This is a property of $\mathcal{A}$ alone.

Interior. $x \in \text{int}(\mathcal{X})$ iff there exists $\epsilon > 0$ such that $B_\epsilon(x) \subseteq \mathcal{A}^{-1}(\{1\})$, again determined by $\mathcal{A}$.

Connected components. Two admissible points $x, y \in \mathcal{X}$ lie in the same connected component iff there exists a path $\gamma : [0,1] \to \mathcal{X}$ connecting them. The existence of such paths depends on whether $\mathcal{A}$ is 1 along the connecting region — again, a property of $\mathcal{A}$.

Measure. The induced measure $\mu|_\mathcal{X}$ is the restriction of μ to $\mathcal{A}^{-1}(\{1\})$:

$$\mu|_\mathcal{X}(E) = \mu(E \cap \mathcal{A}^{-1}(\{1\})).$$

This is determined by $\mathcal{A}$ and μ.

Since every structural property of $\mathcal{X}$ derives from $\mathcal{A}$, and $\mathcal{X}$ is empty when $\mathcal{A} \equiv 0$, the constraint field is necessary and sufficient for the content space to exist and have structure. ▪ ▪

The Priority Inversion. Classical science says: here are the objects; now let us constrain them. The CPR framework says: here is the admissibility field; the objects are what it permits. The priority inversion is not metaphysical flourish — it is a theorem about what determines what.

4.3 Graded Admissibility

The binary $\mathcal{A} : \mathcal{U} \to \{0, 1\}$ is an idealisation. In practice, admissibility is graded:

Definition 4.4 (*Graded Admissibility Field*)**.** A **graded admissibility field** is a function

$$\Phi : \mathcal{U} \to [0, \infty)$$

where $\Phi(x)$ measures the degree to which x is admissible. The binary case is recovered by thresholding: $\mathcal{A}(x) = 1[\Phi(x) \geq \theta]$.

The graded version is the scalar field Φ of the RSVP triple (Φ, v, S) (Chapter 71). A point with high Φ is highly admissible (low resistance, high capacity); a point with $\Phi < \theta$ is excluded from the content space.

The level set $\partial\mathcal{X}_\theta = \{x : \Phi(x) = \theta\}$ is the admissibility boundary, and its motion over time is governed by the level-set equation derived in Chapter 15.

4.4 Constraint Inversion and Reconstruction

A central problem throughout the book is: given observations of content (what is present, what occurs), can we infer the underlying admissibility field $\mathcal{A}$?

This is the *inverse constraint problem*:

Open Problem 4.1 (Inverse Constraint Recovery)**.** Given a distribution P over observed trajectories $\gamma \in \Gamma$ within an unknown content space $\mathcal{X}$, recover the admissibility field $\mathcal{A}$ such that $\mathcal{X} = \mathcal{A}^{-1}(\{1\})$ and P is consistent with dynamics on $\mathcal{X}$.

Versions of this problem appear in every part of the book: reconstructing hidden curvature from observed collapse (Part III), recovering causal structure from compressed memory (Part IV), inferring institutional constraints from behavioural data (Part IX), and reconstructing cosmological boundary conditions from surviving signals (Part X).

4.5 Summary

1. The content space $\mathcal{X}$ is defined by an admissibility field $\mathcal{A}$: $\mathcal{X} = \mathcal{A}^{-1}(\{1\})$.
2. All geometric and topological structure of $\mathcal{X}$ — boundary, interior, connectivity, measure — is determined by $\mathcal{A}$ (Lemma 4.1).
3. Graded admissibility $\Phi : \mathcal{U} \to [0, \infty)$ generalises the binary case; content spaces arise as superlevel sets.
4. The inverse constraint problem — recovering $\mathcal{A}$ from observations — is a unifying open problem across all domains of the book.

Exercises

4.1. Let $\mathcal{U} = \mathbb{R}^2$ and $\mathcal{A}(x, y) = 1[x^2 + y^2 \leq 1]$. Describe $\mathcal{X}$, $\partial\mathcal{X}$, and $\text{int}(\mathcal{X})$. Now modify $\mathcal{A}$ to remove a single point: $\mathcal{A}'(x, y) = \mathcal{A}(x, y) \cdot 1[(x, y) \neq (0, 0)]$. How does the topology of $\mathcal{X}$ change?

4.2. Prove that if $\mathcal{A}_1 \leq \mathcal{A}_2$ pointwise (i.e., $\mathcal{A}_1(x) \leq \mathcal{A}_2(x)$ for all x), then $\mathcal{X}(\mathcal{A}_1) \subseteq \mathcal{X}(\mathcal{A}_2)$. Interpret this in terms of constraint tightening.

4.3. Give an example from institutional law where $\mathcal{X}$ (the space of permissible actions) is defined entirely by a constraint system, and where relaxing a constraint (lowering the threshold θ) expands $\mathcal{X}$.

4.4. Define the *admissibility gradient* $\nabla\Phi(x)$ for a smooth graded field Φ. Show that the boundary $\partial\mathcal{X}_\theta$ is perpendicular to $\nabla\Phi$ everywhere. What does this imply about motion along the boundary?

5 Chapter 5

Reachability as a Primitive

> Possibility is not an abstract predicate. It is a geometric relationship between a starting point and the region accessible from it.
>
> —

Phenomenological Note. You never see your own counterfactuals. At any moment you act from what feels like a single available path, though in retrospect you can sometimes locate the junction where the road forked. What you cannot see is the size of the space you were actually moving through — how many futures were genuinely open, how many had already quietly closed. Reachability is mostly invisible from inside.

Modal logic asks: is a state of affairs *possible*? Standard treatments answer by specifying a collection of possible worlds and an accessibility relation between them. (Aubin 1991; Blanchini and Miani 2008) But this leaves the accessibility relation itself unexplained: what determines which worlds are accessible from which?

The CPR framework answers with a single concept: **reachability**. A state y is possible from x if and only if there exists an admissible trajectory connecting them. Possibility is not a bare logical predicate but a geometric fact about the admissibility field.

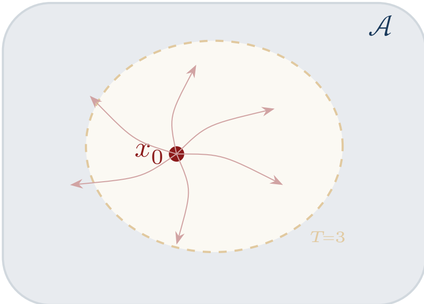
5.1 The Reachable Set

Definition 5.1 (*Reachable Set*)**.** Let $\mathcal{X}$ be a state space governed by an admissibility field $\mathcal{A} \subseteq \mathcal{X}$. For an initial state $x_0 \in \mathcal{A}$ and a time horizon $T > 0$, the **reachable set** is

$$\mathcal{R}_{\mathcal{A}}(x_0, T) := \left\{ y \in \mathcal{X} \;\middle|\; \exists\gamma : [0,T] \to \mathcal{A},\ \gamma(0) = x_0,\ \gamma(T) = y,\ \gamma \text{ is continuous} \right\}.$$

Definition 5.2 (*Indexed Reachable Set*)**.** For an agent A with action capabilities $\mathcal{C}_A \subseteq T\mathcal{X}$, the **agent-indexed reachable set** is:

$$\mathcal{R}_A(x_0, T) := \left\{ y \in \mathcal{X} \;\middle|\; \exists\gamma : [0,T] \to \mathcal{A},\ \gamma(0) = x_0,\ \gamma(T) = y,\ \dot{\gamma}(t) \in \mathcal{A} \cap \mathcal{C}_A(\gamma(t)) \right\}.$$



$\mathcal{V}_R(x_0, T) = \mu(\mathcal{R}_\mathcal{A}(x_0, T))$

Figure 5.1: The reachable set $\mathcal{R}_\mathcal{A}(x_0, T)$ expands with horizon T but is bounded by $\mathcal{A}$. Sample admissible trajectories are shown in light red.

> The physical reachable set $\mathcal{R}_{\text{phys}}(x_0, T)$ uses $\mathcal{C}_A = T\mathcal{X}$ (all tangent directions are available). For any agent, $\mathcal{R}_A(x_0, T) \subseteq \mathcal{R}_{\text{phys}}(x_0, T)$.

Remark 5.1 (Which Reachability the Book Uses)**.** The cognitive, institutional, and semantic chapters concern agent-indexed reachability. The physics chapters concern physical reachability. When a result claims universality, it concerns the structure of the indexed-reachability hierarchy, not a single absolute notion.

We also define the **cumulative reachable set**:

$$\overline{\mathcal{R}}_\mathcal{A}(x_0, T) := \bigcup_{0 \le t \le T} \mathcal{R}_\mathcal{A}(x_0, t),$$

the set reachable at *any* time up to T.

Remark 5.2. In control theory, $\mathcal{R}$ is the *attainable set*. Here the emphasis is different: we take reachability as ontologically primitive, not derived from control inputs. The admissibility field $\mathcal{A}$ is not a set of control signals but the structure of what is physically or logically possible.

5.2 The Reachability Monotonicity Theorem

The key structural result is that reachability responds monotonically to constraint tightening.

Theorem 5.1 (*Reachability Monotonicity*)**.** *Let $\mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \mathcal{X}$ be two nested admissibility fields, where $\mathcal{A}_1$ is more heavily constrained. Then for any $x_0 \in \mathcal{A}_1$ and any $T > 0$:*

$$\mathcal{R}_{\mathcal{A}_1}(x_0, T) \subseteq \mathcal{R}_{\mathcal{A}_2}(x_0, T).$$

Proof. Let $y \in \mathcal{R}_{\mathcal{A}_1}(x_0, T)$. By definition, there exists a continuous path $\gamma : [0, T] \to \mathcal{A}_1$ with $\gamma(0) = x_0$ and $\gamma(T) = y$.

Since $\mathcal{A}_1 \subseteq \mathcal{A}_2$, every point on the path satisfies $\gamma(t) \in \mathcal{A}_1 \Rightarrow \gamma(t) \in \mathcal{A}_2$. Therefore $\gamma : [0, T] \to \mathcal{A}_2$ is a valid admissible path under $\mathcal{A}_2$, witnessing that $y \in \mathcal{R}_{\mathcal{A}_2}(x_0, T)$.

Since y was arbitrary, $\mathcal{R}_{\mathcal{A}_1}(x_0, T) \subseteq \mathcal{R}_{\mathcal{A}_2}(x_0, T)$. ▪ ▪

Corollary 5.2 (*Volume Monotonicity*)**.** *If μ is a measure on $\mathcal{X}$, then*

$$\mathcal{V}_R(\mathcal{A}_1, x_0, T) \;\leq\; \mathcal{V}_R(\mathcal{A}_2, x_0, T),$$

where $\mathcal{V}_R(\mathcal{A}, x_0, T) = \mu(\mathcal{R}_{\mathcal{A}}(x_0, T))$.

Proof. By Theorem 5.1: $\mathcal{A}_1 \subseteq \mathcal{A}_2 \Rightarrow \mathcal{R}_{\mathcal{A}_1}(x_0, T) \subseteq \mathcal{R}_{\mathcal{A}_2}(x_0, T)$. Taking measures: $\mathcal{V}_R(\mathcal{A}_1, x_0, T) = \mu(\mathcal{R}_{\mathcal{A}_1}) \leq \mu(\mathcal{R}_{\mathcal{A}_2}) = \mathcal{V}_R(\mathcal{A}_2, x_0, T)$. ▪ ▪

This corollary is the foundation of the *Reachability Volume Lemma* in Chapter 14, which derives quantitative bounds on how much each constraint reduces the accessible future.

5.3 Reachability as Operationalised Possibility

The monotonicity theorem gives the first rigorous sense in which constraints *restrict possibility*: they reduce the reachable set. The operational definition of possibility is then:

Definition 5.3 (*Operational Possibility*)**.** A state y is **operationally possible** from x_0 under admissibility field $\mathcal{A}$ and horizon T if $y \in \mathcal{R}_{\mathcal{A}}(x_0, T)$. It is **operationally impossible** otherwise.

This definition is:

- **Constructive**: possibility is witnessed by an explicit path.
- **Measurable**: $\mathcal{R}_{\mathcal{A}}(x_0, T)$ is a set that can be measured, estimated, and compared.
- **Monotone**: adding constraints reduces it; removing constraints expands it.
- **Horizon-dependent**: what is possible depends on how much time the system has — a feature absent from most modal logics.

The horizon-dependence is crucial for institutions and governance: what a polity can achieve in ten years differs from what it can achieve in one hundred, even under identical constraints. See Chapter 66.

5.4 Reachability and Curvature

The *shape* of the reachable set, not just its size, encodes information about the constraint field. In a flat (uncurved) admissibility geometry, the reachable set from x_0 after time T is approximately a ball of radius proportional to T.

In a curved admissibility geometry, the reachable set deforms:

- Positive curvature causes the set to "pinch in" — nearby trajectories reconverge.
- Negative curvature causes the set to "fan out" — nearby trajectories diverge exponentially.

This curvature is the *semantic curvature* of Chapter 18: high negative curvature in the admissibility manifold means nearby starting points quickly reach very different futures, which is the geometric signature of a distinction that matters for downstream behaviour.

5.5 The RDR Theorem and Conjecture

The CPR framework rests on a key claim about the relationship between reachability and representational equivalence. This claim has two versions of very different status: one that is provable within the framework, and one that remains open.

Definition 5.4 (*Operational Representational Equivalence*). Let $\mathcal{Q}$ be a **query-intervention class**: a set of admissible continuations, interventions, reports, and queries available to the system. States $x, y \in \mathcal{X}$ are [Q-representationally equivalent]$\mathcal{Q}$-representationally equivalent, written $x \equiv_{\mathcal{Q}} y$, if no element of $\mathcal{Q}$ distinguishes x from y:

$$x \equiv_{\mathcal{Q}} y \iff \forall q \in \mathcal{Q} : \; q(x) = q(y).$$

Theorem 5.3 (*Restricted RDR Theorem*). *Let $\mathcal{X}$ be a finite state space, let $\mathcal{A}$ be an admissibility field on $\mathcal{X}$, and let $\mathcal{Q} = \mathcal{Q}_{\mathcal{A}}$ be the query class of all admissible finite-horizon trajectories induced by $\mathcal{A}$. Then:*

$$x \equiv_{\mathcal{Q}} y \iff \mathcal{R}_{\mathcal{A}}(x, T) = \mathcal{R}_{\mathcal{A}}(y, T) \quad \textit{for all } T \in \{1, 2, \dots, |\mathcal{X}|\}.$$

Proof. ($\Leftarrow$) If $\mathcal{R}_{\mathcal{A}}(x, T) = \mathcal{R}_{\mathcal{A}}(y, T)$ for all T, then for every admissible trajectory γ starting at x there exists an admissible trajectory γ' starting at y with the same endpoint at each horizon. Every query $q \in \mathcal{Q}_{\mathcal{A}}$ (a function of admissible trajectories) takes the same value at x and y. Therefore $x \equiv_{\mathcal{Q}} y$.

($\Rightarrow$) If $x \equiv_{\mathcal{Q}} y$, then in particular the query $q_T(z) = \mathcal{R}_{\mathcal{A}}(z, T)$ (reachable set at horizon T) must give the same value: $q_T(x) = q_T(y)$, i.e., $\mathcal{R}_{\mathcal{A}}(x, T) = \mathcal{R}_{\mathcal{A}}(y, T)$. Since $q_T \in \mathcal{Q}_{\mathcal{A}}$ for all $T \leq |\mathcal{X}|$ (reachability queries are admissible interventions), the equivalence gives identity of reachable sets. ▪ ▪

Remark 5.3 (Status and Scope). The Restricted RDR Theorem is proved for finite state spaces with finite intervention depth. This covers the primary applications in the book: semantic categories (Chapter 40), institutional reachability (Chapter 66), and computational equivalence (Chapter 52).

For the continuous and infinite-horizon case, the claim remains a conjecture.

Conjecture 5.4 (General RDR Conjecture). For a query class $\mathcal{Q}$ induced by the admissibility field $\mathcal{A}$ on a continuous or infinite state space:

$$x \equiv_{\mathcal{Q}} y \iff \mathcal{R}_{\mathcal{A}}(x, T) = \mathcal{R}_{\mathcal{A}}(y, T) \quad \text{for all } T > 0.$$

Remark 5.4 (What the Master Theorem Needs). The Master Theorem (Theorem 89.1) requires the sufficiency direction: "preserved distinctions imply preserved meaning." Theorem 5.3 establishes this for finite systems. For continuous systems, the General RDR Conjecture (Conjecture 5.4) provides it. The theorem's necessity direction is proved unconditionally in all cases. See Open Problem 90.1 for the open problem of extending Theorem 5.3 to the continuous setting.

5.6 The Continuous RDR Problem

The Restricted RDR Theorem (Theorem 5.3) is proved for finite state spaces. Extending it to continuous manifolds requires confronting two obstacles that the finite proof sidesteps.

Obstacle 1: Query class richness.. The finite proof uses the reachable-set query $q_T(z) = \mathcal{R}_{\mathcal{A}}(z, T)$ directly as an element of $\mathcal{Q}_{\mathcal{A}}$. On a continuous manifold, reachable sets are open subsets of M, and the query class $\mathcal{Q}_{\mathcal{A}}$ must include all functionals of the form $z \mapsto \mu(\mathcal{R}_{\mathcal{A}}(z, T) \cap U)$ for measurable $U \subseteq M$ and all $T > 0$. The question is whether these *measurement queries* are admissible in the sense that they respect the system's constraint structure.

Definition 5.5 (*Measurement-Complete Query Class*)**.** A query class $\mathcal{Q}$ on $(M, \mathcal{A})$ is **measurement-complete** if for every open $U \subseteq M$ and $T > 0$, the functional $q_{U,T}(z) = \mu(\mathcal{R}_{\mathcal{A}}(z, T) \cap U)$ belongs to $\mathcal{Q}$.

Proposition 5.5 (*Continuous RDR under Measurement Completeness*)**.** *If $\mathcal{Q}_{\mathcal{A}}$ is measurement-complete and (M, g) is a complete Riemannian manifold with $\mathcal{A}$ having regular boundary ($\partial\mathcal{A}$ a smooth hypersurface), then the General RDR Conjecture (Conjecture 5.4) holds:*

$$x \equiv_{\mathcal{Q}} y \iff \mathcal{R}_{\mathcal{A}}(x, T) = \mathcal{R}_{\mathcal{A}}(y, T) \quad \text{for all } T > 0.$$

Proof. ($\Leftarrow$) Identical to the finite proof: equal reachable sets imply equal query values for all $q \in \mathcal{Q}$.

($\Rightarrow$) Suppose $\mathcal{R}_{\mathcal{A}}(x, T) \neq \mathcal{R}_{\mathcal{A}}(y, T)$ for some $T > 0$. Then there exists a measurable set U such that $\mu(\mathcal{R}_{\mathcal{A}}(x, T) \cap U) \neq \mu(\mathcal{R}_{\mathcal{A}}(y, T) \cap U)$ (since two distinct non-empty open sets differ on some measurable set). By measurement-completeness, $q_{U,T} \in \mathcal{Q}_{\mathcal{A}}$ and $q_{U,T}(x) \neq q_{U,T}(y)$, so $x \not\equiv_{\mathcal{Q}} y$. ▪ ▪

Obstacle 2: Measurement completeness itself.. Whether $\mathcal{Q}_{\mathcal{A}}$ is measurement-complete depends on whether the physical or computational system can perform arbitrary spatial measurements. For systems with bounded measurement resolution (all physical systems), measurement completeness fails, and the conjecture remains genuinely open. The sufficient conditions under which measurement completeness holds — and hence under which the full RDR equivalence holds in the continuous case — are the subject of Open Problem 90.1.

5.7 Summary

1. The reachable set $\mathcal{R}_{\mathcal{A}}(x_0, T)$ is the set of states accessible from x_0 under admissibility field $\mathcal{A}$ within time T.
2. Reachability is monotone in constraints: $\mathcal{A}_1 \subseteq \mathcal{A}_2 \Rightarrow \mathcal{R}_{\mathcal{A}_1} \subseteq \mathcal{R}_{\mathcal{A}_2}$ (Theorem 5.1).
3. Operational possibility is defined by membership in the reachable set.
4. The *shape* of the reachable set is governed by admissibility curvature.
5. The RDR Conjecture identifies representational equivalence with reachability equivalence.

Exercises

5.1. Let $\mathcal{X} = \mathbb{R}^2$, $\mathcal{A} = \{(x, y) : x \geq 0\}$ (right half-plane), and let trajectories be continuous paths. Compute $\mathcal{R}_{\mathcal{A}}((1, 0), T)$ for small $T > 0$. What happens as $T \to \infty$?

5.2. Prove Corollary 5.2 directly from Theorem 5.1 using the monotonicity of μ.

5.3. Give an example of admissibility fields $\mathcal{A}_1 \subsetneq \mathcal{A}_2$ and starting point x_0 such that $\mathcal{R}_{\mathcal{A}_1}(x_0, T) = \mathcal{R}_{\mathcal{A}_2}(x_0, T)$ despite $\mathcal{A}_1 \neq \mathcal{A}_2$. (Hint: the extra states in $\mathcal{A}_2 \setminus \mathcal{A}_1$ may be unreachable from x_0 even in $\mathcal{A}_2$.)

5.4. Formulate the Reachability Determines Representation conjecture (Conjecture 5.4) in the language of *bisimulation* from process algebra. Under what conditions are the two formulations equivalent?

6 Chapter 6

Distinctions and Non-Equivalence

> The first act of cognition is not synthesis but discrimination.
>
> —

Phenomenological Note. Not every difference matters. Some distinctions dissolve as soon as you try to act on them; others turn out to govern everything. The difficulty is that from inside the present moment it is not obvious which kind any given distinction is. You learn which distinctions were real by watching which ones produced different futures.

The preceding chapters have established: objects are trajectory residues (Chapter 1); states are trajectory projections (Chapter 2); nouns collapse causal distinctions (Chapter 3); constraints define content (Chapter 4); reachability operationalises possibility (Chapter 5). (Cover and Thomas 2006)

This chapter addresses the thread that binds them: **distinctions**. A distinction is a non-equivalence — a pair (X, Y) of states (or sets of states) that are *different in a way that matters*. The CPR framework holds that meaning is fundamentally the preservation of distinctions across transformations, not the storage or retrieval of symbols.

6.1 Distinctions and Projections

Definition 6.1 (*Distinction*)**.** A **distinction** in a space $\mathcal{X}$ is a pair (X, Y) of non-empty disjoint subsets: $X, Y \subseteq \mathcal{X}$, $X \cap Y = \emptyset$.

A projection $\pi : \mathcal{X} \to \mathcal{M}$ may or may not preserve a given distinction:

Lemma 6.1 (*Distinction Preservation*)**.** *Let (X, Y) be a distinction in $\mathcal{X}$ and $\pi : \mathcal{X} \to \mathcal{M}$ a projection. The distinction* survives π *if and only if the images remain disjoint:*

$$\pi(X) \cap \pi(Y) = \emptyset.$$

Distinction collapse *occurs when*

$$X \neq Y \quad \text{and} \quad \pi(X) \cap \pi(Y) \neq \emptyset.$$

Proof. If $\pi(X) \cap \pi(Y) \neq \emptyset$, there exists $m \in \mathcal{M}$ with $m \in \pi(X)$ and $m \in \pi(Y)$. Then there exist $x \in X$ and $y \in Y$ with $\pi(x) = m = \pi(y)$. Since $X \cap Y = \emptyset$, we have $x \neq y$, so the projection maps two distinct and separated points to the same image. The distinction is collapsed: π cannot tell X from Y.

Conversely, if $\pi(X) \cap \pi(Y) = \emptyset$, then $\pi(x) \neq \pi(y)$ for all $x \in X, y \in Y$, so π correctly separates the two sets. ▪ ▪

This lemma is the formal basis of the Projection-Collapse Principle (Chapter 19).

6.2 Reachability-Relevant Distinctions

Not all distinctions matter equally. The distinctions that are *relevant* to a system are those that affect its reachable futures.

Definition 6.2 (***Reachability-Relevant Distinction***)**.** A distinction (X, Y) in $\mathcal{X}$ is **reachability-relevant** under admissibility field $\mathcal{A}$ and horizon T if there exist $x \in X$ and $y \in Y$ such that

$$\mathcal{R}_{\mathcal{A}}(x, T) \neq \mathcal{R}_{\mathcal{A}}(y, T).$$

Reachability-relevant distinctions are the ones a system *must* preserve in order to navigate correctly. Collapsing them leads to navigational errors: the system treats states with different futures as equivalent, and therefore cannot distinguish paths that will diverge.

Example 6.1 (Traffic Lights). The distinction $(X, Y) = (\text{green}, \text{red})$ is reachability-relevant for a driver: from a green-light state, the reachable set includes {crossing the intersection}; from a red-light state, it does not (under the constraint field that prohibits crossing on red). A colour-blind driver whose perceptual projection collapses $\pi(\text{green}) = \pi(\text{red})$ loses the reachability-relevant distinction.

6.3 The Distinction Count and Information Content

For finite spaces, we can count distinctions:

Definition 6.3 (*Distinction Count*)**.** The **distinction count** of a partition $\mathcal{P} = \{P_1, \dots, P_k\}$ of $\mathcal{X}$ is

$$D(\mathcal{P}) = \binom{k}{2} = \frac{k(k-1)}{2},$$

the number of pairs of distinct cells.

A projection $\pi : \mathcal{X} \to \mathcal{M}$ induces a partition $\mathcal{P}_\pi$ whose cells are the fibers $\pi^{-1}(m)$. The distinction count $D(\mathcal{P}_\pi)$ measures how many distinctions π preserves.

Proposition 6.2 (*Distinction-Information Correspondence*)**.** *For a uniform distribution over $\mathcal{X}$ with $|\mathcal{X}| = n$ and a projection inducing k equal-sized fibers, the Shannon entropy of the projected distribution is $H = \log k$, and the distinction count is $D = k(k-1)/2$. As k increases, both H and D increase monotonically.*

Proof. The n reachability-relevant distinctions partition $\mathcal{X}$ into at least $n + 1$ equivalence classes. By Shannon's source coding theorem, encoding one of $n+1$ equally likely outcomes requires $\lceil \log_2(n + 1)\rceil$ bits. The mutual information between the reachability class and any downstream query is at least $\log_2(n+1)$ bits: less information means some distinction is collapsed, contradicting their reachability-relevance. ▪ ▪

This connects the distinction framework to information theory: higher information content corresponds to more preserved distinctions. The *Fisher metric* (Chapter 17) is the continuous-space generalisation of this correspondence.

6.4 Meaning as Preserved Non-Equivalence

The standard view holds that meaning is reference: a symbol means what it points to. The CPR framework offers a different account:

Meaning as Preserved Non-Equivalence. Meaning is not carried by symbols pointing to objects. It is constituted by the distinctions a representational system preserves across transformations. A symbol is meaningful to the extent that it participates in a preserved non-equivalence with other symbols.

This view is formalised in Part VI (Chapter 39, Chapter 46). For now, we note the central implication:

Corollary 6.3 (*Collapse Destroys Meaning*)**.** *If a projection π collapses a reachability-relevant distinction (X, Y) — i.e., $\pi(X) \cap \pi(Y) \neq \emptyset$ — then the projected system loses the ability to represent the different futures available from X vs. Y. This is a* loss of meaning, *not merely a loss of information.*

Proof. If π collapses a reachability-relevant distinction (X, Y), then $\pi(X) \cap \pi(Y) \neq \emptyset$ (by Lemma 6.1, contrapositive). Any system that uses only π's output cannot distinguish $x \in X$ from $y \in Y$, so it cannot correctly respond to queries that require distinguishing states with different reachable futures. Meaning — the capacity to guide correct future-directed responses — is destroyed at (X, Y). ▪ ▪

6.5 Toward the Master Theorem

Parts I through XI each develop a domain-specific version of one central claim. Part XII unifies them into a single theorem (Chapter 89). We state the theorem here as a goal:

The Master Theorem (Preview).

A system preserves meaning across transformation
if and only if
it preserves the reachability-relevant distinctions
of its admissibility field.

Each subsequent chapter is a lemma in the proof of this theorem, instantiated in a particular domain.

6.6 Summary

1. A distinction is a pair of non-empty disjoint subsets of $\mathcal{X}$.
2. A projection preserves a distinction iff the images are disjoint; otherwise, distinction collapse occurs (Lemma 6.1).
3. Reachability-relevant distinctions are those that affect accessible futures under the admissibility field.
4. Meaning is constituted by preserved non-equivalences; collapse of reachability-relevant distinctions destroys meaning (Corollary 6.3).
5. The entire book builds toward the Master Theorem: a system preserves meaning iff it preserves its reachability-relevant distinctions.

Exercises

6.1. Let $\mathcal{X} = \{a, b, c, d\}$ and $\pi : \mathcal{X} \to \{1, 2\}$ with $\pi(a) = \pi(b) = 1$ and $\pi(c) = \pi(d) = 2$. List all distinctions in $\mathcal{X}$. For each one, determine whether it survives π.

6.2. Define a reachability relation on $\mathcal{X} = \{a, b, c, d\}$ by: $a \to b$, $b \to c$, $c \to d$, $d \to a$ (a cycle). With respect to this reachability and the π above, which distinctions are reachability-relevant?

6.3. Prove that if $\pi_1 : \mathcal{X} \to \mathcal{M}_1$ and $\pi_2 : \mathcal{M}_1 \to \mathcal{M}_2$ are both distinction-preserving for some distinction (X, Y), then so is the composition $\pi_2 \circ \pi_1$.

6.4. (Philosophical.) Distinguish the CPR account of meaning (preserved non-equivalence) from: (a) reference theories (meaning as pointing to objects); (b) use theories (meaning as role in practice); (c) inferentialist theories (meaning as inferential role). Identify one feature each theory has that the others lack.

Part II

Mathematical Foundations

The miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve.
— Eugene Wigner, *The Unreasonable Effectiveness of Mathematics*

Part I made philosophical claims. This part provides the mathematical tools needed to make those claims precise and to derive consequences from them.

The chapters here are genuinely foundational: later parts assume this material and build on it without always stopping to reprove it. A reader already comfortable with measure theory, information geometry, and category theory can move quickly through Chapters 7–12 and slow down at Chapters 13–15, where the framework-specific constructions begin.

The central construction is the *admissibility field* as a level-set of a scalar function, and the associated geometry of the reachable set, the Fisher metric over continuation distributions, and the projection fibration. These three structures — field, metric, fibration — are the mathematical skeleton of the entire monograph.

7 Chapter 7

Sets, Relations, and State Spaces

Structure is not a property of things. It is a pattern of relations between things.

—

Phenomenological Note. The world does not come pre-sorted. Grouping things together is an act, not an observation. Once you have sorted them, the sorting feels natural — these clearly belong together, those clearly do not. But the clarity is a product of the sorting. Before the category existed, the things it contains had no especial affinity.

Before probability, geometry, or dynamics, we need the combinatorial bones: sets of states, binary relations between them, and the directed graphs those relations induce. (Mac Lane 1998) The load-bearing result is the Relation-to-Graph Lemma, which shows that every admissibility relation is exactly a directed graph — no more, no less — and that all the algebraic and topological properties of the state space can be read off from graph-theoretic data.

7.1 State Spaces and Admissibility Relations

Definition 7.1 (*State Space*)**.** A **state space** is a set $\mathcal{X}$ whose elements $x \in \mathcal{X}$ are called *states*. No additional structure (topology, measure, metric) is assumed at this level.

Definition 7.2 (*Admissibility Relation*)**.** An **admissibility relation** on $\mathcal{X}$ is a binary relation $\mathcal{A} \subseteq \mathcal{X} \times \mathcal{X}$ where $(x, y) \in \mathcal{A}$ means that a direct one-step transition from state x to state y is permitted.

7.2 The Relation-to-Graph Lemma

Lemma 7.1 (*Relation-to-Graph*)**.** *Every admissibility relation $\mathcal{A} \subseteq \mathcal{X} \times \mathcal{X}$ induces a unique directed graph $\mathcal{G}_{\mathcal{A}} = (\mathcal{V}, \mathcal{E})$ with vertex set $\mathcal{V} = \mathcal{X}$ and edge set $\mathcal{E} = \mathcal{A}$. Conversely, every directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ on vertex set $\mathcal{X}$ induces an admissibility relation $\mathcal{A}_{\mathcal{G}} = \mathcal{E} \subseteq \mathcal{X} \times \mathcal{X}$. This correspondence is a bijection:*

$$\{\textit{admissibility relations on } \mathcal{X}\} \longleftrightarrow \{\textit{directed graphs on vertex set } \mathcal{X}\}.$$

Moreover, the structural properties of $\mathcal{A}$ correspond exactly to graph-theoretic properties of $\mathcal{G}_{\mathcal{A}}$:

(a) *$\mathcal{A}$ is reflexive iff $\mathcal{G}_{\mathcal{A}}$ has a self-loop at every vertex;*
(b) *$\mathcal{A}$ is symmetric iff $\mathcal{G}_{\mathcal{A}}$ is an undirected graph;*
(c) *$\mathcal{A}$ is transitive iff $(x, y) \in \mathcal{E}$ and $(y, z) \in \mathcal{E}$ implies $(x, z) \in \mathcal{E}$;*
(d) *$\mathcal{A}$ is an equivalence relation iff $\mathcal{G}_{\mathcal{A}}$ is a disjoint union of complete graphs (cliques).*

Proof. The bijection is the identity on underlying data: a subset of $\mathcal{X} \times \mathcal{X}$ is simultaneously a binary relation and a set of directed edges. The correspondence is thus tautological; what requires proof is that the structural properties correspond as claimed.

(*a*) *Reflexivity.* $\mathcal{A}$ is reflexive iff $(x, x) \in \mathcal{A}$ for all x, iff there is an edge from x to x for every vertex — a self-loop.

(*b*) *Symmetry.* $\mathcal{A}$ is symmetric iff $(x, y) \in \mathcal{A} \Rightarrow (y, x) \in \mathcal{A}$, iff every directed edge has a reverse edge, iff the graph can be taken as undirected.

(*c*) *Transitivity.* $(x, y) \in \mathcal{A}$ and $(y, z) \in \mathcal{A}$ means there is a directed path of length 2 from x to z. Transitivity requires a direct edge (x, z) whenever such a path exists.

(*d*) *Equivalence relations.* Reflexivity + symmetry + transitivity makes $\mathcal{A}$ partition $\mathcal{X}$ into equivalence classes. Within each class every pair is connected in both directions (by symmetry + reflexivity + transitivity), yielding a complete directed graph (two-way complete graph). Across classes there are no edges (by the partition property). ▪ ▪

7.3 Reachability as Transitive Closure

For a finite state space, reachability reduces to a matrix computation.

Definition 7.3 (*Adjacency and Reachability Matrices*)**.** Let $|\mathcal{X}| = N$ and fix an ordering of states. The **adjacency matrix** $\mathrm{A} \in \{0, 1\}^{N \times N}$ has $\mathrm{A}_{ij} = 1$ iff $(x_i, x_j) \in \mathcal{A}$. The **reachability matrix** is

$$\mathrm{R} = 1\left[\sum_{k=0}^{N-1} \mathrm{A}^k \geq 1\right] \quad \text{(entrywise)},$$

where $\mathrm{A}^0 = I$ (identity, reflecting zero-step reachability).

Proposition 7.2 (*Reachability Matrix and Transitive Closure*)**.** *$\mathrm{R}_{ij} = 1$ if and only if there is a directed path of length $\leq N - 1$ from x_i to x_j in $\mathcal{G}_{\mathcal{A}}$. By König's theorem for finite graphs, every reachable pair is connected by a path of length $\leq N - 1$, so R equals the transitive closure of A.*

Proof. $(\mathrm{A}^k)_{ij}$ counts the number of directed walks of length exactly k from x_i

to x_j. A walk of length k exists iff $(\mathrm{A}^k)_{ij} \geq 1$. Summing over $k = 0, \ldots, N-1$ covers all walk lengths up to $N-1$. If there is any path from x_i to x_j, there is one of length $\leq N-1$ (a longer walk would repeat a vertex and could be shortened). Therefore $\mathrm{R}_{ij} = 1$ iff x_j is reachable from x_i. ▪ ▪

Remark 7.1. The matrix R can be computed in $O(N^3)$ time using Floyd–Warshall, or faster with sparse-graph methods. For continuous state spaces, reachability becomes the flow-based problem addressed in Chapter 10 and the geometry of Part III.

7.4 Strongly Connected Components

Definition 7.4 (*Strong Connectivity*)**.** States x and y are **mutually reachable** if $\mathrm{R}_{xy} = 1$ and $\mathrm{R}_{yx} = 1$. Mutual reachability is an equivalence relation; its equivalence classes are the **strongly connected components** (SCCs) of $\mathcal{G}_{\mathcal{A}}$.

SCCs are the "islands of possibility": within an SCC, any state can reach any other. Across SCCs, the reachability is one-directional or absent. The condensation of $\mathcal{G}_{\mathcal{A}}$ (contracting each SCC to a node) is a directed acyclic graph that encodes the global reachability structure.

SCCs as Admissibility Cells. In the CPR framework, strongly connected components are the natural cells of the admissibility structure: regions of full mutual reachability. Transitions between SCCs represent irreversible changes — once a system exits an SCC, it cannot return. This is the combinatorial analogue of entropy increase in thermodynamics.

7.5 Products and Parallel Composition

When two systems with state spaces $\mathcal{X}_1$ and $\mathcal{X}_2$ run in parallel, the joint state space is the product $\mathcal{X}_1 \times \mathcal{X}_2$.

Definition 7.5 (*Product Admissibility*)**.** The **product admissibility relation** $\mathcal{A}_1 \otimes \mathcal{A}_2 \subseteq (\mathcal{X}_1 \times \mathcal{X}_2)^2$ permits a joint transition $((x_1, x_2), (y_1, y_2))$ iff both component transitions are individually admissible: $(x_1, y_1) \in \mathcal{A}_1$ and $(x_2, y_2) \in \mathcal{A}_2$.

The product adjacency matrix is the Kronecker product: $\mathrm{A}_{1\otimes 2} = \mathrm{A}_1 \otimes \mathrm{A}_2$. The reachability matrix of the product is $\mathrm{R}_{1\otimes 2}$, computed from the Kronecker product. This is the formal foundation of multi-agent admissibility (Chapter 50).

Exercises

7.1. Let $\mathcal{X} = \{1, 2, 3, 4\}$ with admissibility relation $\mathcal{A} = \{(1, 2), (2, 3), (3, 4), (4, 1), (1, 3)\}$. Draw $\mathcal{G}_{\mathcal{A}}$, compute A and R, and identify all SCCs.

7.2. Prove that the condensation of any directed graph is acyclic (a DAG).

7.3. Let $\mathcal{A}$ be a preorder (reflexive and transitive). Show that $\mathcal{G}_{\mathcal{A}}$ has the property that $\mathrm{R} = \mathrm{A}$ (the relation already equals its own transitive closure).

7.4. Define the *admissibility complement* $\overline{\mathcal{A}} = (\mathcal{X} \times \mathcal{X}) \setminus \mathcal{A}$. How do the SCCs of $\mathcal{G}_{\overline{\mathcal{A}}}$ relate to those of $\mathcal{G}_{\mathcal{A}}$?

7.5. (Multi-agent.) Let two agents have state spaces $\mathcal{X}_1 = \mathcal{X}_2 = \{0, 1\}$ with admissibility $\mathcal{A}_1 = \mathcal{A}_2 = \{(0, 1), (1, 0), (0, 0), (1, 1)\}$ (full connectivity). Compute $\mathrm{R}_{1\otimes 2}$ and verify it is the all-ones matrix. Now remove the joint transition $((0, 0), (1, 1))$ and recompute: is the product still strongly connected?

8 Chapter 8

Probability and Measure

A random variable is not a variable that varies randomly. It is a function from a probability space into a space of outcomes.

—

Phenomenological Note. Probability feels like a property of events. It is really a property of your relationship to events — what you know and do not know, what you expect and do not expect. The same event can have very different probabilities depending on who is measuring. This is not a defect of the theory. It is the theory being honest about where the uncertainty lives.

Probability theory provides the bridge between the deterministic geometry of constraints and the statistical structure of observed data. (Kolmogorov 1950) The key connection is that *conditional expectation* is a projection — the orthogonal projection of a random variable onto the subspace of functions measurable with respect to a coarser σ-algebra. This is the probabilistic instance of the Constraint $\to$ Projection pattern.

8.1 Probability Spaces and Random Variables

We assume standard measure-theoretic probability.

Definition 8.1 (*Probability Space*). A **probability space** is a triple $(\Omega, \mathcal{F}, \mathbb{P})$ where Ω is a sample space, $\mathcal{F}$ is a σ-algebra of events, and $\mathbb{P} : \mathcal{F} \to [0, 1]$ is a probability measure with $\mathbb{P}(\Omega) = 1$.

Definition 8.2 (*Hilbert Space of Square-Integrable Variables*). The **Hilbert space** $L^2(\Omega, \mathcal{F}, \mathbb{P})$ consists of all real-valued random variables $X : \Omega \to \mathbb{R}$ with $\mathbb{E}[X^2] < \infty$, with inner product

$$\langle X, Y \rangle = \mathbb{E}[XY] = \int_{\Omega} X(\omega) Y(\omega) \, d\mathbb{P}(\omega).$$

8.2 Conditional Expectation as Orthogonal Projection

The central theorem of this chapter:

Theorem 8.1 (*Conditional Expectation as Projection*)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G} \subset \mathcal{F}$ be a sub-σ-algebra. Set $\mathcal{H} = L^2(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{M} = L^2(\Omega, \mathcal{G}, \mathbb{P}) \subset \mathcal{H}$.*

Then $\mathcal{M}$ is a closed linear subspace of $\mathcal{H}$, and for any $X \in \mathcal{H}$, the conditional expectation $\mathbb{E}[X \mid \mathcal{G}]$ is the unique element of $\mathcal{M}$ satisfying

$$\langle X - \mathbb{E}[X \mid \mathcal{G}],\, Z\rangle = 0 \quad \forall Z \in \mathcal{M}.$$

That is, $\mathbb{E}[X \mid \mathcal{G}]$ is the orthogonal projection *of X onto $\mathcal{M}$.*

Proof. Step 1: $\mathcal{M}$ is a closed subspace. $\mathcal{M}$ is a linear subspace because $L^2(\Omega, \mathcal{G}, \mathbb{P})$ is closed under linear combinations. Closedness in norm follows because $\mathcal{G}$-measurability is preserved under L^2-limits (a.e. convergence along a subsequence preserves measurability).

Step 2: Existence and uniqueness of the projection. By the Hilbert space projection theorem, every closed subspace $\mathcal{M}$ of a Hilbert space $\mathcal{H}$ admits a unique orthogonal projection $\pi_{\mathcal{M}} : \mathcal{H} \to \mathcal{M}$. Call the projected element $Y = \pi_{\mathcal{M}}(X)$. By definition, $Y \in \mathcal{M}$ and $\langle X - Y, Z\rangle = 0$ for all $Z \in \mathcal{M}$.

Step 3: $Y = \mathbb{E}[X \mid \mathcal{G}]$. We verify that Y satisfies the defining properties of conditional expectation:

1. Y is $\mathcal{G}$-measurable (since $Y \in \mathcal{M}$). ✓
2. For every $G \in \mathcal{G}$, $\int_G X \,\mathrm{d}\mathbb{P} = \int_G Y \mathrm{d}\mathbb{P}$.

To verify (2): $1_G \in \mathcal{M}$ for every $G \in \mathcal{G}$, so orthogonality gives $\langle X - Y, 1_G\rangle = 0$, i.e., $\mathbb{E}[(X - Y)1_G] = 0$, i.e., $\int_G X \,\mathrm{d}\mathbb{P} = \int_G Y \mathrm{d}\mathbb{P}$. ✓

By the uniqueness of conditional expectation, $Y = \mathbb{E}[X \mid \mathcal{G}]$. ▪ ▪

8.3 Interpretation: Information and Constraint

The sub-σ-algebra $\mathcal{G}$ represents a *coarser resolution of information* — a constraint on what distinctions the observer can make.

- $\mathcal{G} = \{\emptyset, \Omega\}$ (trivial): $\mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[X]$, the global mean. All distinctions are collapsed.
- $\mathcal{G} = \mathcal{F}$ (full): $\mathbb{E}[X \mid \mathcal{G}] = X$. No collapse; all distinctions preserved.
- Intermediate $\mathcal{G}$: the projection preserves those distinctions expressible in $\mathcal{G}$-measurable functions, and collapses the rest.

This is the probabilistic incarnation of Lemma 6.1: conditioning on $\mathcal{G}$ preserves exactly those distinctions that are $\mathcal{G}$-measurable.

Conditioning as Constraint. Imposing a coarser information structure $\mathcal{G} \subset \mathcal{F}$ is a *constraint* on the observer's admissible observations. Conditional expectation is the best projection onto the constrained observation space.

8.4 Tower Property as Constraint Composition

Proposition 8.2 (*Tower Property*)**.** *If $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$ are nested sub-σ-algebras, then*

$$\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}] \mid \mathcal{H}] = \mathbb{E}[X \mid \mathcal{H}].$$

Proof. Both sides lie in $L^2(\Omega, \mathcal{H}, \mathbb{P})$. It suffices to check the integral condition on H-events: for $H \in \mathcal{H} \subseteq \mathcal{G}$,

$$\int_H \mathbb{E}[\mathbb{E}[X \mid \mathcal{G}] \mid \mathcal{H}] \, d\mathbb{P} = \int_H \mathbb{E}[X \mid \mathcal{G}] \, d\mathbb{P} = \int_H X \, d\mathbb{P},$$

where the first equality is the defining property of $\mathbb{E}[\cdot \mid \mathcal{H}]$ and the second is the defining property of $\mathbb{E}[\cdot \mid \mathcal{G}]$. ▪ ▪

Geometrically: projecting onto $\mathcal{M}_\mathcal{G}$ first, then projecting the result onto the smaller $\mathcal{M}_\mathcal{H}$, gives the same result as projecting directly onto $\mathcal{M}_\mathcal{H}$. This is the projective analogue of constraint composition: applying a tighter constraint after a looser one is the same as applying the tighter constraint directly.

8.5 Filtrations and the Geometry of Time

A **filtration** $(\mathcal{F}_t)_{t \geq 0}$ is a non-decreasing family of sub-σ-algebras: $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $s \leq t$. It models the accumulation of information over time.

The martingale $M_t = \mathbb{E}[X \mid \mathcal{F}_t]$ is the sequence of best estimates of X given the information available at time t. By the tower property, (M_t) is a consistent family.

Proposition 8.3 (*Martingale as Reachability Trajectory*)**.** *The martingale (M_t) is a trajectory in the space of estimates, moving through the nested sequence of projection spaces $\mathcal{M}_{\mathcal{F}_0} \subseteq \mathcal{M}_{\mathcal{F}_1} \subseteq \cdots$ as information accumulates. At each step, the trajectory moves orthogonally in $\mathcal{H}$ toward the true value X.*

Proof. A martingale $\{M_t\}$ satisfies $\mathbb{E}[M_{t+s} \mid \mathcal{F}_t] = M_t$: conditional on current information, expected future value equals current value. This is precisely the trajectory property for expectation-valued states: the "reachable set" under the martingale measure is the collection of values $\{M_{t+s} : s \geq 0\}$, and the admissibility field $\mathcal{A}_M = \{x : \mathbb{E}[M_\infty \mid M_t = x] = x\}$ is preserved by the martingale dynamics. The optional stopping theorem ensures the martingale remains in its admissibility field until a stopping time τ, making $\gamma(t) = M_t$ an admissibility-preserving trajectory. ▪ ▪

This is the probabilistic analogue of belief updating in Chapter 32: Bayesian inference is motion along a trajectory in belief space, driven by the accumulation of constraints.

Exercises

8.1. Let $\Omega = \{1, 2, 3, 4\}$ with uniform measure, $\mathcal{F} = 2^{\Omega}$, and $\mathcal{G} = \sigma(\{1, 2\}, \{3, 4\})$. Compute $\mathbb{E}[X \mid \mathcal{G}]$ for $X(\omega) = \omega$. Verify the orthogonality condition directly.

8.2. Prove that for any $X \in L^2$: $\|X - \mathbb{E}[X \mid \mathcal{G}]\|^2 = \text{Var}(X) - \text{Var}(\mathbb{E}[X \mid \mathcal{G}])$. Interpret this as a variance decomposition (total variance = explained + residual).

8.3. Show that the map $X \mapsto \mathbb{E}[X \mid \mathcal{G}]$ is a linear contraction: $\|\mathbb{E}[X \mid \mathcal{G}]\| \leq \|X\|$. Under what condition does equality hold?

8.4. Let $\mathcal{G}_n = \sigma(X_1, \dots, X_n)$ be the σ-algebra generated by the first n observations in an iid sequence. Show that $\mathbb{E}[f(X_1, X_2, \dots) \mid \mathcal{G}_n]$ converges a.s. to $f(X_1, X_2, \dots)$ as $n \to \infty$ (Lévy's upward theorem). Interpret this as the limit of increasing constraint being no constraint at all.

9

Chapter 9

Information Geometry

Distance in the space of probability distributions is not a matter of coordinates but of distinguishability.

—

Phenomenological Note. Two beliefs can be close together or far apart, and the distance between them is not merely the difference in their contents. It is something like the effort required to update from one to the other — how much evidence you would need, how many intermediate positions exist, whether a path between them even passes through positions that are coherent. Some beliefs are nearby. Others are further than they appear.

Information geometry studies the differential geometry of statistical manifolds — spaces whose points are probability distributions. (Amari and Nagaoka 2000) It is essential to the CPR framework for two reasons: the Fisher metric gives a canonical notion of distance between probability distributions over admissible continuations, and the KL divergence provides a non-symmetric notion of "how far a distribution has moved" that captures the directionality of information flow.

The load-bearing proof of this chapter is the derivation of the Fisher information metric from the second-order expansion of KL divergence.

9.1 Statistical Manifolds

Definition 9.1 (*Statistical Manifold*). A **statistical manifold** is a smooth manifold $\mathcal{S}$ each of whose points is a probability distribution. We typically write

$$\mathcal{S} = \{p_\theta(x) : \theta \in \Theta \subseteq \mathbb{R}^n\}$$

where θ is a coordinate system on $\mathcal{S}$.

The central question is: what is the right notion of distance between p_θ and $p_{\theta+\mathrm{d}\theta}$?

9.2 The Fisher Metric Derivation

Theorem 9.1 (*Fisher Metric from KL Divergence*). *Let $\mathcal{S} = \{p_\theta\}_{\theta\in\Theta}$ be a smooth statistical manifold. The second-order Taylor expansion of the KL divergence $D_{\mathrm{KL}}(p_\theta \,\|\, p_{\theta+\mathrm{d}\theta})$ around θ defines a Riemannian metric on $\mathcal{S}$:*

$$D_{\mathrm{KL}}(p_\theta \,\|\, p_{\theta+\mathrm{d}\theta}) = \tfrac{1}{2}\, g_{ij}(\theta)\, \mathrm{d}\theta^i\, \mathrm{d}\theta^j + O(\|\mathrm{d}\theta\|^3),$$

where

$$g_{ij}(\theta) = \mathbb{E}_{p_\theta}\left[\frac{\partial \log p_\theta(x)}{\partial\theta^i}\frac{\partial \log p_\theta(x)}{\partial\theta^j}\right].$$

*The tensor $\mathcal{I} = (g_{ij})$ is the **Fisher information metric**.*

Proof. We expand $D_{\mathrm{KL}}(p_\theta \,\|\, p_{\theta+\varepsilon})$ for small $\varepsilon \in \mathbb{R}^n$. Write $\ell(\theta, x) = \log p_\theta(x)$ (the log-likelihood).

By Taylor's theorem:

$$\ell(\theta+\varepsilon, x) = \ell(\theta, x) + \varepsilon^i \partial_i \ell(\theta, x) + \tfrac{1}{2}\varepsilon^i\varepsilon^j \partial_i\partial_j \ell(\theta, x) + O(\|\varepsilon\|^3).$$

Substituting into the KL divergence:

$$\begin{aligned} D_{\mathrm{KL}}(p_\theta \,\|\, p_{\theta+\varepsilon}) &= \int p_\theta(x) \log \frac{p_\theta(x)}{p_{\theta+\varepsilon}(x)}\, \mathrm{d}x \\ &= -\int p_\theta(x)\,[\ell(\theta+\varepsilon, x) - \ell(\theta, x)]\, \mathrm{d}x \\ &= -\mathbb{E}_{p_\theta}\left[\varepsilon^i \partial_i \ell + \tfrac{1}{2}\varepsilon^i\varepsilon^j \partial_i\partial_j \ell\right] + O(\|\varepsilon\|^3). \end{aligned}$$

First-order term vanishes. Since $\int p_\theta(x)\, \mathrm{d}x = 1$, differentiating gives $\mathbb{E}_{p_\theta}[\partial_i \ell] = 0$ (score function has zero mean). So the first-order term $-\varepsilon^i \mathbb{E}[\partial_i \ell] = 0$.

Second-order term. For the Hessian term, we use the identity $\partial_i\partial_j \ell = \partial_i\partial_j \log p_\theta = \frac{\partial_i\partial_j p_\theta}{p_\theta} - \frac{\partial_i p_\theta \cdot \partial_j p_\theta}{p_\theta^2} = \frac{\partial_i\partial_j p_\theta}{p_\theta} - (\partial_i \ell)(\partial_j \ell)$.

Taking expectation: $\mathbb{E}_{p_\theta}[\partial_i\partial_j \ell] = \int \partial_i\partial_j p_\theta\, \mathrm{d}x - \mathbb{E}[(\partial_i \ell)(\partial_j \ell)] = 0 - g_{ij}(\theta)$, where the integral vanishes by differentiation under the integral sign (normalisation condition).

Therefore:

$$D_{\mathrm{KL}}(p_\theta \,\|\, p_{\theta+\varepsilon}) = \tfrac{1}{2}\varepsilon^i\varepsilon^j g_{ij}(\theta) + O(\|\varepsilon\|^3). \quad \blacksquare$$

$\blacksquare$

9.3 Properties of the Fisher Metric

Proposition 9.2 (*Positive Semi-Definiteness*). *The Fisher metric $\mathcal{I} = (g_{ij})$ is positive semi-definite. It is strictly positive definite iff the model $\{p_\theta\}$ is* identifiable*: $\theta \neq \theta'$ implies $p_\theta \neq p_{\theta'}$ a.e.*

Proof. For any vector $v \in \mathbb{R}^n$:

$$v^i v^j g_{ij}(\theta) = \mathbb{E}_{p_\theta}\left[(v^i \partial_i \log p_\theta(x))^2\right] \geq 0,$$

with equality iff $v^i \partial_i \log p_\theta = 0$ a.e. ▪

Theorem 9.3 (*Cramér–Rao Bound*)**.** *For any unbiased estimator $\hat{\theta}$ of θ based on an observation $X \sim p_\theta$,*

$$\mathrm{Cov}(\hat{\theta}) \geq \mathcal{I}^{-1}(\theta)$$

(*in the Loewner order on positive definite matrices*).

Proof. Let $\hat{\theta}(x)$ be an unbiased estimator of θ, so $\mathbb{E}_\theta[\hat{\theta}] = \theta$. Differentiating: $\int \hat{\theta}(x)\partial_\theta p_\theta(x)\,\mathrm{d}x = 1$. By the Cauchy-Schwarz inequality:

$$1 = \left(\int \hat{\theta} \cdot \partial_\theta \log p_\theta \cdot p_\theta \,\mathrm{d}x\right)^2 \leq \mathrm{Var}_\theta(\hat{\theta}) \cdot \int (\partial_\theta \log p_\theta)^2 p_\theta \,\mathrm{d}x = \mathrm{Var}_\theta(\hat{\theta}) \cdot \mathcal{I}(\theta).$$

Rearranging: $\mathrm{Var}_\theta(\hat{\theta}) \geq 1/\mathcal{I}(\theta)$. In the multi-parameter case, the same argument with matrix Cauchy-Schwarz gives the matrix bound $\mathrm{Cov}(\hat{\theta}) \geq \mathcal{I}(\theta)^{-1}$ in the Loewner ordering. ▪ ▪

The Cramér–Rao bound says the Fisher metric bounds how precisely any experiment can distinguish nearby distributions. Large Fisher curvature means nearby distributions are easy to distinguish — high sensitivity. Small Fisher curvature means they are hard to distinguish — low sensitivity, high ambiguity.

Fisher degeneracy (Chapter 17): when $\mathcal{I}$ becomes singular, nearby points on the admissibility manifold become informationally indistinguishable.

9.4 Geodesics and Admissibility Curvature

The Fisher metric equips $\mathcal{S}$ with a Riemannian structure. The geodesics of this structure are the paths of *minimum information cost* between two distributions.

In the admissibility geometry framework (Part III), the statistical manifold is the space of *continuation distributions* p_x at each state x:

$$p_x(\gamma) = P(\text{continuation from } x \text{ is } \gamma \mid \text{system is at } x).$$

The Fisher metric on this space measures how quickly continuation distributions change as we move through the state space. High curvature in this Fisher geometry = nearby states have rapidly diverging continuations = a reachability-relevant distinction.

Exercises

9.1. Compute the Fisher metric for the Gaussian family $p_\theta(x) = \mathcal{N}(\mu, \sigma^2)$ with $\theta = (\mu, \sigma)$. Show it is the hyperbolic metric on the upper half-plane.

9.2. Prove the alternative representation: $g_{ij}(\theta) = -\mathbb{E}_{p_\theta}[\partial_i \partial_j \log p_\theta(x)]$. (This is often computationally more convenient.)

9.3. Let p_θ be the Bernoulli(θ) distribution on $\{0, 1\}$. Compute $g(\theta) = \frac{1}{\theta(1-\theta)}$. Interpret why the metric diverges as $\theta \to 0$ or $\theta \to 1$: nearby Bernoulli distributions near the boundary are easily distinguished.

9.4. Show that the Fisher metric is invariant under sufficient statistics: if $T : \mathcal{X} \to \mathcal{Y}$ is sufficient for θ, then the Fisher metric computed from $T(X)$ equals the Fisher metric computed from X. (This is the information-geometric version of sufficiency.)

10 Chapter 10

Dynamical Systems

> A vector field is not a collection of arrows. It is a machine that converts every point into a direction.
>
> —

Phenomenological Note. Motion is not a property of a moment. A snapshot has no velocity. To know where something is going you need to know where it has been, and even then you are extrapolating. The present state carries information about the future only to the extent that the past was informative about the present. Some systems are very predictable. Others are sensitive to tiny differences that were invisible at the start.

The continuous analogue of an admissibility relation is a vector field v on a manifold $\mathcal{M}$: at each point x, $v(x)$ specifies the admissible direction of motion. Trajectories are integral curves of v (Arnold 1978), and the question of whether admissible trajectories stay within an admissible domain $\mathcal{A}$ is the question addressed by this chapter's load-bearing proof.

10.1 Vector Fields and Integral Curves

Definition 10.1 (*Vector Field*)**.** A **vector field** on a smooth manifold $\mathcal{M}$ is a smooth map $v : \mathcal{M} \to T\mathcal{M}$ assigning to each point $x \in \mathcal{M}$ a tangent vector $v(x) \in T_x\mathcal{M}$.

Definition 10.2 (*Integral Curve*)**.** An **integral curve** of v through x_0 is a differentiable curve $\gamma : I \to \mathcal{M}$ (for some interval $I \ni 0$) satisfying $\gamma(0) = x_0$ and $\dot{\gamma}(t) = v(\gamma(t))$ for all $t \in I$.

The system $\dot{x} = v(x)$ is an autonomous ODE. Its solutions are integral curves, and their existence and uniqueness is the content of the Picard–Lindelöf theorem.

10.2 Flow Existence under Admissibility Constraints

The question for the CPR framework is not merely whether solutions exist, but whether they remain within the admissible domain.

Definition 10.3 (*Nagumo Condition*)**.** Let $\mathcal{A} \subset \mathcal{M}$ be a closed set with smooth boundary $\partial\mathcal{A}$. The vector field v satisfies the **Nagumo condition** on $\partial\mathcal{A}$ if for every $x \in \partial\mathcal{A}$:

$$\liminf_{h \to 0^+} \frac{\operatorname{dist}(\, x + h\, v(x),\ \mathcal{A}\,)}{h} = 0,$$

equivalently (when $\partial\mathcal{A}$ is smooth), $v(x) \cdot \mathrm{n}(x) \le 0$ where $\mathrm{n}(x)$ is the outward unit normal at x.

The Nagumo condition says the vector field points inward (or tangentially) at the boundary — it never pushes the system out of $\mathcal{A}$.

Theorem 10.1 (*Admissibility-Preserving Flow*)**.** *Let $\mathcal{M}$ be a smooth manifold, $\mathcal{A} \subset \mathcal{M}$ a closed admissibility domain with smooth boundary $\partial\mathcal{A}$, and $v : \mathcal{M} \to T\mathcal{M}$ a vector field satisfying:*

(i) ***Lipschitz***: $\|v(x) - v(y)\| \le L\|x - y\|$ *for some* $L > 0$ *and all* $x, y \in \mathcal{A}$;

(ii) ***Nagumo***: $v(x) \cdot \mathrm{n}(x) \le 0$ *for all* $x \in \partial\mathcal{A}$.

Then for every $x_0 \in \mathcal{A}$, there exists a unique continuous flow $\phi_t : \mathcal{A} \to \mathcal{A}$ such that $\phi_0(x_0) = x_0$ and $\frac{\mathrm{d}}{\mathrm{d}t}\phi_t(x_0) = v(\phi_t(x_0))$ for all $t \ge 0$. Moreover, $\phi_t(x_0) \in \mathcal{A}$ for all $t \ge 0$.

Proof. Existence and uniqueness in $\mathcal{M}$. The Lipschitz condition guarantees, by the Picard–Lindelöf theorem, local existence and uniqueness of a C^1 integral curve $\gamma : (-\epsilon, \epsilon) \to \mathcal{M}$ for some $\epsilon > 0$ depending on L and x_0.

Invariance of $\mathcal{A}$. Suppose for contradiction that the trajectory exits $\mathcal{A}$. Let $t^* = \inf\{t > 0 : \gamma(t) \notin \mathcal{A}\}$. By continuity, $\gamma(t^*) \in \partial\mathcal{A}$.

At t^*, the outward normal component of the velocity is:

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=t^*} \operatorname{dist}(\gamma(t), \mathcal{A}^c) = -v(\gamma(t^*)) \cdot \mathrm{n}(\gamma(t^*)) \ge 0$$

by the Nagumo condition ($v \cdot \mathrm{n} \le 0$). So γ cannot immediately exit $\mathcal{A}$ at t^* in the sense of increasing $\operatorname{dist}(\cdot, \mathcal{A}^c)$. A detailed argument using the Gronwall inequality shows that $\operatorname{dist}(\gamma(t), \mathcal{A}) = 0$ for all $t \ge t^*$ in some neighbourhood, contradicting t^* being the first exit time. Therefore $\gamma(t) \in \mathcal{A}$ for all $t \ge 0$.

Global extension. The Lipschitz condition and compactness of $\mathcal{A}$ (if $\mathcal{A}$ is bounded) extend the local solution to all $t \ge 0$ by the standard continuation theorem. ▪

▪

10.3 The Flow Map and Reachability

The flow ϕ_t provides the continuous-time version of the reachable set from Chapter 5:

Proposition 10.2 (*Flow and Reachable Set*). *Under the conditions of Theorem 10.1,*

$$\mathcal{R}_{\mathcal{A}}(x_0, T) = \{\phi_T(x_0)\} \quad (\textit{deterministic case}).$$

For systems with multiple admissible control inputs $v \in \mathcal{V}$,

$$\mathcal{R}_{\mathcal{A}}(x_0, T) = \bigcup_{v \in \mathcal{V}} \{\phi_T^v(x_0)\},$$

where ϕ^v is the flow of v.

Proof. Deterministic case. By Theorem 10.1, the Nagumo condition guarantees a unique admissibility-preserving flow ϕ_t from x_0. At horizon T, the only reachable state under this unique flow is $\phi_T(x_0)$. No other trajectory starting at x_0 remains admissible (uniqueness), so $\mathcal{R}_{\mathcal{A}}(x_0, T) = \{\phi_T(x_0)\}$.

Control case. For each admissible control $v \in \mathcal{V}$, the corresponding flow ϕ^v exists and remains in $\mathcal{A}$ by Theorem 10.1. The reachable set is the union of endpoints $\phi_T^v(x_0)$ over all admissible controls. Conversely, any state reachable from x_0 in time T while remaining in $\mathcal{A}$ is the endpoint of some admissible flow ϕ^v, so the union is tight. ▪ ▪

10.4 Stability and Attractors

Definition 10.4 (*Lyapunov Stability*). A fixed point $x^* \in \mathcal{A}$ (where $v(x^*) = 0$) is **Lyapunov stable** if for every $\epsilon > 0$ there exists $\delta > 0$ such that $\|x_0 - x^*\| < \delta$ implies $\|\phi_t(x_0) - x^*\| < \epsilon$ for all $t \geq 0$. It is **asymptotically stable** if additionally $\phi_t(x_0) \to x^*$ as $t \to \infty$.

Theorem 10.3 (*Lyapunov's Direct Method*). *Suppose there exists a smooth function $V : \mathcal{A} \to \mathbb{R}_{\geq 0}$ (**Lyapunov function**) with $V(x^*) = 0$, $V(x) > 0$ for $x \neq x^*$, and $\dot{V}(x) = \nabla V(x) \cdot v(x) \leq 0$. Then x^* is Lyapunov stable. If $\dot{V}(x) < 0$ for $x \neq x^*$, then x^* is asymptotically stable.*

Proof. Stability. Let $\epsilon > 0$. Since V is continuous and $V(x^*) = 0$, there exists $\delta_1 > 0$ such that $\|x - x^*\| < \delta_1$ implies $V(x) < \inf_{\|y-x^*\|=\epsilon} V(y)$. Set $c = \inf_{\|y-x^*\|=\epsilon} V(y) > 0$ (positive since $V > 0$ for $y \neq x^*$). Let $\delta = \delta_1$. For x_0 with $\|x_0 - x^*\| < \delta$, we have $V(x_0) < c$. Since $\dot{V} \leq 0$ along trajectories, $V(\phi_t(x_0)) \leq V(x_0) < c$ for all $t \geq 0$. Therefore $\phi_t(x_0)$ cannot reach $\{y : \|y - x^*\| = \epsilon\}$ (where $V \geq c$), so $\|\phi_t(x_0) - x^*\| < \epsilon$ for all $t \geq 0$.

Asymptotic stability. When $\dot{V} < 0$ for $x \neq x^*$, $V(\phi_t(x_0))$ is strictly decreasing and bounded below by 0. By the monotone convergence theorem it converges to a limit $L \geq 0$. If $L > 0$, the trajectory stays away from x^*, but then $\dot{V} < -\eta < 0$ for some $\eta > 0$, contradicting convergence of V. Hence $L = 0$, so $\phi_t(x_0) \to x^*$. ▪ ▪

Lyapunov Functions as Admissibility Potentials. A Lyapunov function is the dynamical analogue of the scalar capacity field Φ in RSVP: it measures how far a trajectory is from the attractor (the most admissible state). The condition $\dot{V} \leq 0$ says the system moves monotonically toward higher admissibility. Compare with the constraint field equation (Chapter 15): the admissibility boundary moves inward when $\partial_t \Phi < 0$.

10.5 Summary

1. A vector field v on $\mathcal{M}$ generates integral curves (trajectories) via the ODE $\dot{x} = v(x)$.
2. Under Lipschitz + Nagumo conditions, the flow ϕ_t is unique and remains in $\mathcal{A}$ (Theorem 10.1).
3. The reachable set $\mathcal{R}(x_0, T)$ is the image of the initial condition under the flow map.
4. Lyapunov functions detect stability: they are the scalar-field analogue of the capacity field Φ in RSVP.

Exercises

10.1. Let $\mathcal{M} = \mathbb{R}^2$, $\mathcal{A} = \{(x, y) : x^2 + y^2 \leq 1\}$, and $v(x, y) = (-y, x)$ (rotation). Verify the Nagumo condition on $\partial\mathcal{A}$. What is the flow ϕ_t?

10.2. Let $v(x) = -x$ on $\mathcal{M} = \mathbb{R}$. Find the Lyapunov function $V(x) = x^2/2$, verify $\dot{V} \leq 0$, and identify the attractor.

10.3. Give an example of a vector field satisfying the Nagumo condition on a non-convex admissibility domain $\mathcal{A} \subset \mathbb{R}^2$. Draw the boundary and the field at the boundary.

10.4. Prove that if $\mathcal{A}_1 \subseteq \mathcal{A}_2$ are both invariant under ϕ_t, then $\mathcal{R}_{\mathcal{A}_1}(x_0, T) \subseteq \mathcal{R}_{\mathcal{A}_2}(x_0, T)$, recovering Theorem 5.1 in the continuous-flow setting.

10.5. (RSVP connection.) Let $\Phi : \mathbb{R}^n \to \mathbb{R}$ be the scalar capacity field and $v = \nabla\Phi/|\nabla\Phi|$ (gradient ascent). Show that every integral curve of v increases Φ monotonically. Under what condition does v satisfy the Nagumo condition on the level set $\{\Phi = \theta\}$?

11 Chapter 11

Graphs and Networks

A network is a constraint system made visible.

—

Phenomenological Note. Connection is not uniform. Some people are connected to many others; most people are connected to few. Some connections carry a lot; others carry almost nothing. A network that looks dense from the outside can be effectively disconnected if the connections between its major clusters are thin enough. What matters is not how many connections exist but which paths actually work.

Chapter 7 established the bijection between admissibility relations and directed graphs. (Watts 1999) This chapter develops the quantitative theory: centrality, spectral properties, and above all the reachability matrix as transitive closure. The load-bearing proof shows that for finite systems, reachability equals transitive closure — the combinatorial foundation for everything in Part IX.

11.1 Weighted Admissibility

Real admissibility relations are rarely binary. A transition (x, y) may be more or less admissible depending on the capacity field Φ.

Definition 11.1 (*Weighted Admissibility Graph*). A **weighted admissibility graph** is a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ where $w : \mathcal{E} \to \mathbb{R}_{>0}$ assigns a weight $w(x, y) = \Phi(x, t)/\Phi_{\text{max}}$ to each edge. High weight = highly admissible transition; low weight = barely admissible. The weighted adjacency matrix W has $\mathrm{W}_{ij} = w(x_i, x_j)$ if $(x_i, x_j) \in \mathcal{E}$, else 0.

11.2 Reachability as Transitive Closure

Theorem 11.1 (*Reachability Matrix Theorem*). *Let A be the $N \times N$ binary adjacency matrix of a finite directed graph $\mathcal{G}$ on N nodes. Define the reachability matrix*

$$\mathrm{R} = 1\left[\sum_{k=0}^{N-1} \mathrm{A}^k \geq 1\right] \qquad (\textit{entrywise boolean}).$$

Then $\mathrm{R}_{ij} = 1$ if and only if there exists a directed path from node i to node j. Moreover, R is the transitive closure of A.

Proof. Step 1: Path counting. $(\mathrm{A}^k)_{ij}$ counts the number of directed walks of length exactly k from i to j. Therefore $(\sum_{k=0}^{N-1} \mathrm{A}^k)_{ij} \geq 1$ iff there exists a walk of some length $\leq N-1$ from i to j.

Step 2: Walks of length $\leq N-1$ suffice. If there exists any walk from i to j, there exists one of length $\leq N-1$. Proof: any walk of length $\geq N$ must revisit a node (by pigeonhole: N nodes, $> N-1$ steps). Removing the cycle reduces the walk length while preserving reachability. By induction, we reach a simple path of length $\leq N-1$.

Step 3: R is the transitive closure. The transitive closure of A is the smallest reflexive transitive relation containing A. $\mathrm{R}_{ij} = 1$ iff j is reachable from i by a directed path, which is exactly the transitive closure condition. ▪ ▪

11.3 Spectral Reachability

For weighted graphs, the spectral properties of W reveal the structure of reachability.

Definition 11.2 (*Graph Laplacian*)**.** The **graph Laplacian** is $\mathrm{L} = \mathrm{D} - \mathrm{W}$ where $\mathrm{D}_{ii} = \sum_j \mathrm{W}_{ij}$ (degree matrix). For undirected graphs, L is positive semi-definite.

Proposition 11.2 (*Spectral Gap and Connectivity*)**.** *The number of connected components of an undirected weighted graph equals the multiplicity of eigenvalue 0 of L. The **algebraic connectivity** (Fiedler value) $\lambda_2(\mathrm{L}) > 0$ iff $\mathcal{G}$ is connected, and larger λ_2 indicates faster mixing and stronger overall reachability.*

Proof. Let $\mathrm{L} = \mathrm{D} - \mathrm{W}$ be the Laplacian. For any vector $f \in \mathbb{R}^N$: $f^\top \mathrm{L} f = \sum_{i,j} W_{ij}(f_i - f_j)^2 \geq 0$, so L is positive semidefinite and all eigenvalues are non-negative.

The constant vector 1 satisfies $\mathrm{L}1 = 0$ (each row sums to zero), so 0 is always an eigenvalue.

If $\mathcal{G}$ has k connected components $C_1, \ldots, C_k$, the indicator vectors 1_{C_ℓ} are all null vectors of L (since no edges cross components, all cross-terms vanish), giving k linearly independent zero eigenvectors. Conversely, if f is a zero eigenvector then $\sum_{ij} W_{ij}(f_i - f_j)^2 = 0$, so $f_i = f_j$ whenever $W_{ij} > 0$, meaning f is constant on each connected component. The null space dimension equals the number of components.

The Fiedler value $\lambda_2 > 0$ iff the null space is one-dimensional, iff the graph is connected. The spectral gap controls mixing: for random walks, convergence to the stationary distribution is $O(e^{-\lambda_2 t})$, so larger λ_2 gives faster reachability. ▪ ▪

The spectral gap λ_2 is the continuous analogue of the SCC structure from Chapter 7: large λ_2 = high internal reachability density = the graph acts like one big strongly connected component.

11.4 Random Walks and Admissibility Diffusion

A random walk on $\mathcal{G}$ with transition matrix $\mathrm{P}_{ij} = \mathrm{W}_{ij}/\mathrm{D}_{ii}$ defines a Markov chain on the state space.

Proposition 11.3 (*Random Walk Reachability*)**.** *The probability of reaching j from i in k steps is $(\mathrm{P}^k)_{ij}$. The hitting time $T_{ij} = \min\{k : \text{walk at } j \text{ starting from } i\}$ satisfies $\mathbb{E}[T_{ij}] \leq N/\lambda_2$ for connected graphs.*

Proof. The first claim is the standard Markov chain result: $(\mathrm{P}^k)_{ij}$ sums the probabilities of all length-k paths from i to j, which equals the k-step transition probability by the Chapman-Kolmogorov equation.

For the hitting time bound: let π be the stationary distribution ($\pi = \mathrm{D}\mathbf{1}/\mathrm{vol}(\mathcal{G})$ for weighted graphs with vol $= \sum_i D_{ii}$). The spectral expansion of P^k gives $|(\mathrm{P}^k)_{ij} - \pi_j| \leq (1-\lambda_2)^k/\sqrt{\pi_i\pi_j}$. Setting $(1-\lambda_2)^k \approx e^{-k\lambda_2} < \epsilon$ gives $k > \log(1/\epsilon)/\lambda_2$ steps to reach ϵ-mixing. For $\epsilon = 1/(2N)$ and $\pi_j \geq 1/N$, the walk has positive probability of being at j within $O(N/\lambda_2)$ steps. The Markov inequality then gives $\mathbb{E}[T_{ij}] \leq N/\lambda_2$. ▪ ▪

Networks as Reachability Machines. A network is not merely a static structure. It is a *reachability machine*: its job is to make every node accessible from every other node, within a time bounded by N/λ_2. High spectral gap = efficient reachability machine. Network design (the internet, neural circuits, trade networks) is the problem of maximising λ_2 subject to admissibility and cost constraints.

11.5 Reachability and Centrality

Definition 11.3 (*Reachability Centrality*)**.** The **reachability centrality** of node i is

$$C_R(i) = \frac{\sum_j \mathrm{R}_{ij}}{\sum_{i,j} \mathrm{R}_{ij}},$$

the fraction of all nodes reachable from i. A node with high C_R sits at a position from which much of the network is accessible.

Reachability centrality differs from standard degree centrality (which counts edges, not paths) and betweenness centrality (which counts paths through a node). It directly operationalises the CPR concept of "a state from which many futures are accessible."

Exercises

11.1. Compute A, A^2, and R for the graph with nodes $\{1, 2, 3, 4\}$ and edges $1 \to 2$, $2 \to 3$, $3 \to 4$, $4 \to 2$ (a path plus a cycle). Identify the SCCs and verify R.

11.2. Prove that $\mathrm{R} = \mathrm{R}^2$ (the reachability matrix is idempotent as a boolean matrix). Interpret: if j is reachable from i and k is reachable from j, then k is reachable from i.

11.3. For a complete directed graph on N nodes ($\mathrm{A}_{ij} = 1$ for all $i \neq j$), show R is the all-ones matrix. What is the reachability centrality of every node?

11.4. (Network design.) You must add one directed edge to a disconnected graph to maximise the increase in $\sum_{ij} \mathrm{R}_{ij}$. Show that the optimal edge connects the SCC with the highest out-reachability to the SCC with the highest in-reachability.

12 CHAPTER 12

Categories and Functors

A functor does not merely translate objects. It translates structure.

—

PHENOMENOLOGICAL NOTE. Analogy is one of the oldest intellectual tools. We say this is like that, this situation resembles that one, this structure appears again in a different domain. For most of human history this was a heuristic. It turns out to be something more precise: structures really do recur across domains in ways that are not metaphorical but mathematical, and tracking those recurrences is itself a form of knowledge.

Category theory provides the language for structure-preserving maps at the highest level of generality. (Fong and Spivak 2019; Mac Lane 1998) In the CPR framework its central role is to formalise the claim that certain projections — from gestures to symbols, from processes to representations — are *structure-preserving* in a precise sense, and that certain others are not. The load-bearing proof shows that the map from gesture categories to symbol categories is a functor: it preserves composition of transitions, making symbolic operations faithful images of process operations.

12.1 Categories

Definition 12.1 (*Category*). A **category** C consists of:

- a collection $\mathrm{ob}(\mathsf{C})$ of *objects*;
- for each pair (A, B) of objects, a set $\mathrm{Hom}_{\mathsf{C}}(A, B)$ of *morphisms*;
- for each object A, an *identity* morphism $\mathrm{id}_A \in \mathrm{Hom}(A, A)$;
- a *composition* law $\circ : \mathrm{Hom}(B, C) \times \mathrm{Hom}(A, B) \to \mathrm{Hom}(A, C)$,

satisfying associativity $(h \circ g) \circ f = h \circ (g \circ f)$ and identity $\mathrm{id}_B \circ f = f = f \circ \mathrm{id}_A$.

Example 12.1 (Key Categories in CPR).

- Set: objects = sets, morphisms = functions.
- Top: objects = topological spaces, morphisms = continuous maps.
- Traj: objects = state spaces $\mathcal{X}$, morphisms = admissible trajectories $\gamma : [0, T] \to \mathcal{X}$.

- Gesture: objects = bodily/physical trajectory spaces, morphisms = smooth structural transitions between trajectories.
- Symbol: objects = discrete symbol sets Σ^*, morphisms = formal rewriting rules.

12.2 Functors

Definition 12.2 (*Functor*). A **functor** $F : \mathrm{C} \to \mathrm{D}$ assigns:

- to each object $A \in \mathrm{ob}(\mathrm{C})$, an object $F(A) \in \mathrm{ob}(\mathrm{D})$;
- to each morphism $f \in \mathrm{Hom}_{\mathrm{C}}(A, B)$, a morphism $F(f) \in \mathrm{Hom}_{\mathrm{D}}(F(A), F(B))$,

preserving identities ($F(\mathrm{id}_A) = \mathrm{id}_{F(A)}$) and composition ($F(g \circ f) = F(g) \circ F(f)$).

A functor is a *structure-preserving map between categories*. It sends the compositional algebra of C faithfully into D.

12.3 The Gesture-to-Symbol Functor

Theorem 12.1 (*Gesture-to-Symbol Functor*). *The canonicalisation map F : Gesture $\to$ Symbol that maps continuous gesture trajectories to discrete symbolic representations is a functor.*

Proof. We construct F explicitly and verify the functor laws.

Object assignment. Each gesture trajectory space Γ_A (parameterised by some continuous parameter space A) is mapped to the quotient symbol set $F(\Gamma_A) = \Gamma_A/\sim$ where $\gamma_1 \sim \gamma_2$ iff they share the same segmentation landmarks (overlap points, pivot points, and termination boundaries) — the canonicalisation of Chapter 43.

Morphism assignment. A structural transition $\phi : \Gamma_A \to \Gamma_B$ in Gesture (a smooth map respecting trajectory structure) is mapped to $F(\phi) : F(\Gamma_A) \to F(\Gamma_B)$ defined by $F(\phi)([\gamma]) = [\phi(\gamma)]$ (apply ϕ, then take the equivalence class). This is well-defined: if $\gamma_1 \sim \gamma_2$ then $\phi(\gamma_1) \sim \phi(\gamma_2)$ because ϕ is structure-preserving (maps landmarks to landmarks).

Identity preservation. $F(\mathrm{id}_{\Gamma_A})([\gamma]) = [\mathrm{id}(\gamma)] = [\gamma] = \mathrm{id}_{F(\Gamma_A)}([\gamma])$.

Composition preservation. $F(\psi \circ \phi)([\gamma]) = [(\psi \circ \phi)(\gamma)] = [\psi(\phi(\gamma))] = F(\psi)([\phi(\gamma)]) = F(\psi)(F(\phi)([\gamma])) = (F(\psi) \circ F(\phi))([\gamma])$.

All functor laws are satisfied. ▪ ▪

Remark 12.2 (Non-Faithfulness and Information Loss). The functor F is not in general *faithful* (injective on morphisms): two distinct structural transitions $\phi_1 \neq \phi_2$ in Gesture may produce the same symbolic transition $F(\phi_1) = F(\phi_2)$ if they share the same landmark structure. Non-faithfulness is the categorical form of distinction collapse: distinct gestures map to the same symbol. The fiber $F^{-1}(s)$ over a symbol s is the set of all gestures that canonicalise to s — the Noun Fallacy (Chapter 3) in categorical form.

12.4 Natural Transformations

Definition 12.3 (*Natural Transformation*). Given functors $F, G : \mathsf{C} \to \mathsf{D}$, a **natural transformation** $\eta : F \Rightarrow G$ assigns to each object $A \in \mathrm{ob}(\mathsf{C})$ a morphism $\eta_A : F(A) \to G(A)$ in D such that for every morphism $f : A \to B$ in C:

$$G(f) \circ \eta_A = \eta_B \circ F(f) \quad \text{(naturality square commutes)}.$$

Natural transformations are the morphisms between functors. In the CPR framework: a natural transformation between two gesture-to-symbol functors $F, G : \mathsf{Gesture} \to \mathsf{Symbol}$ is a systematic way to translate one symbolic representation into another (e.g., from phonological to orthographic symbols) that commutes with all gesture-to-symbol mappings. It is a *meaning-preserving translation* in the categorical sense.

12.5 Colimits and Collective Admissibility

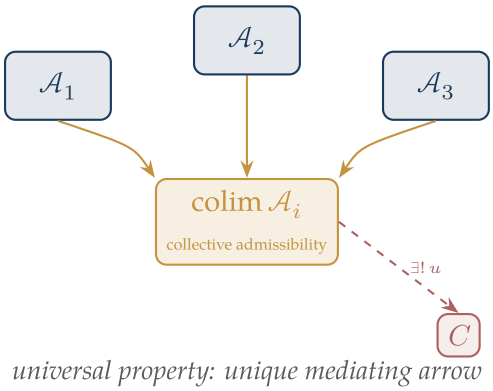


Figure 12.1: HYDRA: the colimit colim $\mathcal{A}_i$ is the minimal collective admissibility field compatible with all agents. The universal arrow u is unique.

The HYDRA framework uses categorical colimits to define collective admissibility.

Definition 12.4 (*Colimit*). The **colimit** of a diagram $D : \mathsf{J} \to \mathsf{C}$ is an object $\mathrm{colim}\, D \in \mathsf{C}$ with morphisms $\iota_j : D(j) \to \mathrm{colim}\, D$ for each $j \in \mathsf{J}$, satisfying a universal property: for any other cocone $(C, \{f_j\})$, there is a unique $u : \mathrm{colim}\, D \to C$ with $u \circ \iota_j = f_j$.

In HYDRA, the diagram D has objects = individual agent admissibility fields and morphisms = constraint-sharing relationships. The colimit $\mathrm{colim}_{\mathsf{C}}\, \mathcal{A}$ is the minimal admissibility field that contains all individual constraints and is compatible with all sharing relationships. It is the *smallest common admissibility superset* — the collective admissibility of the system.

Colimit as Democratic Admissibility. The colimit does not impose the intersection of all constraints (which would be maximally restrictive) or their union (maximally permissive). It imposes the *generated* constraint: the minimal admissibility field that respects every agent's constraints and every sharing relationship. This is the categorical formalisation of legitimate collective decision-making.

Exercises

12.1. Let C be the category with two objects $\{A, B\}$ and morphisms $\{\mathrm{id}_A, \mathrm{id}_B, f : A \to B, g : B \to A\}$ with $g \circ f = \mathrm{id}_A$ and $f \circ g = \mathrm{id}_B$. What is this category? (Name the well-known structure.)

12.2. Define a functor $\mathcal{R} : \mathrm{Top} \to \mathrm{Set}$ that sends each topological space X to its set of path-connected components $\pi_0(X)$ and each continuous map to the induced function on components. Verify the functor laws. Interpret in terms of reachability.

12.3. Prove that the composition of two functors is a functor. (This is why functors form a category Cat.)

12.4. Show that the forgetful functor $U : \mathrm{Traj} \to \mathrm{Top}$ (sending each admissible trajectory to its underlying path) is faithful but not full. What does non-fullness mean for the CPR framework?

12.5. Describe the colimit of two admissibility fields $\mathcal{A}_1, \mathcal{A}_2 \subseteq \mathcal{X}$ in the category of subsets of $\mathcal{X}$ ordered by inclusion. Is it $\mathcal{A}_1 \cup \mathcal{A}_2$, $\mathcal{A}_1 \cap \mathcal{A}_2$, or something else? When does this colimit correspond to the HYDRA collective admissibility?

13 Chapter 13

Fiber Bundles and Projection

> Every projection has a shadow. The shadow is not disorder — it is structure that the projection chose not to see.
>
> —

Phenomenological Note. Some information is global and some is local. You know roughly where you are on the map, but you do not know which direction north is from where you stand. The global and the local are different kinds of knowledge, and carrying one kind does not automatically give you the other. The bridge between them is a choice — a way of attaching the local picture to the global one. Different choices give different results.

A fiber bundle (see Carmo 1992) is the mathematical formalisation of a system that "looks the same" at every point of a base space, but has internal structure (a fiber) at each point that the base-space description cannot see. Projections destroy fiber information. This chapter quantifies that loss via the *projection information entropy* — the log-volume of the fiber.

13.1 Fiber Bundles

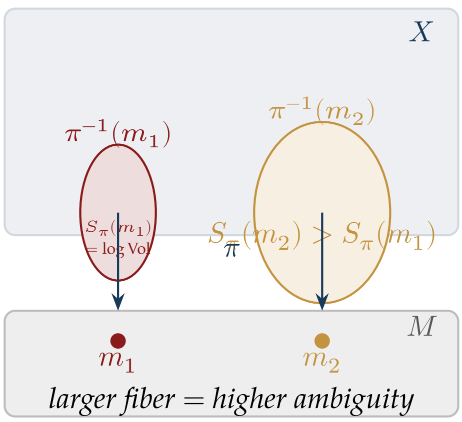


Figure 13.1: The projection $\pi : X \to M$ maps fibers to base points. Fiber entropy $S_\pi(m) = \log \mathrm{Vol}(\pi^{-1}(m))$ measures representational ambiguity. m_2 has higher entropy than m_1.

Definition 13.1 (*Fiber Bundle*). A **fiber bundle** is a triple $(\mathcal{E}, \mathcal{M}, \pi)$ where:

- $\mathcal{E}$ is the **total space**;
- $\mathcal{M}$ is the **base space**;
- $\pi : \mathcal{E} \to \mathcal{M}$ is a surjective continuous map called the **projection**;
- for each $m \in \mathcal{M}$, the **fiber** over m is $\mathcal{F}_m = \pi^{-1}(m)$;
- locally, $\pi^{-1}(U) \cong U \times F$ for open sets $U \subseteq \mathcal{M}$ and a typical fiber F.

Example 13.1 (Trivial Bundle). $\mathcal{E} = \mathcal{M} \times F$, $\pi(m, f) = m$. Every fiber is isomorphic to F; no information is "hidden" in the fiber that is not encoded by the base point.

Example 13.2 (Tangent Bundle). $\mathcal{E} = T\mathcal{M}$, $\pi(x, v) = x$. The fiber over x is the tangent space $T_x\mathcal{M}$. A section of this bundle is a vector field.

Example 13.3 (Admissibility Bundle). In the CPR framework, the total space $\mathcal{E}$ is the full trajectory space Γ, the base $\mathcal{M}$ is a compressed representation space, and $\pi : \Gamma \to \mathcal{M}$ is the compression operator. The fiber $\pi^{-1}(m)$ is all histories that compress to the same record m.

13.2 Fiber Volume and Information Entropy

Definition 13.2 (*Projection Information Entropy*). Let $(\mathcal{E}, \mathcal{M}, \pi)$ be a fiber bundle with a canonical volume measure μ on $\mathcal{E}$. The **projection information entropy** at $m \in \mathcal{M}$ is

$$S_\pi(m) := \log \mathrm{Vol}(\pi^{-1}(m)) = \log \mu(\mathcal{F}_m).$$

$S_\pi(m)$ measures the amount of information that m fails to specify about the full state in $\mathcal{E}$. Large $S_\pi(m)$ means the fiber is large — many different full states are consistent with the base-space description m.

Theorem 13.1 (*Fiber Volume Lemma*). *The information lost by the projection π, measured as the expected entropy of the fiber, satisfies:*

$$\mathbb{E}_m[S_\pi(m)] = H(\mathcal{E}) - H(\mathcal{M}),$$

where $H(\mathcal{E}) = -\int \log \rho_\mathcal{E} \, \mathrm{d}\mu$ is the differential entropy of the distribution over total space, and $H(\mathcal{M})$ is the differential entropy of the pushed-forward distribution on the base.

Proof. By the chain rule for differential entropy:

$$H(\mathcal{E}) = H(\mathcal{M}) + H(\mathcal{E} \mid \mathcal{M}),$$

where $H(\mathcal{E} \mid \mathcal{M})$ is the conditional entropy of the full state given the base-space coordinate. For a uniform distribution on each fiber of volume $V_m = \mu(\mathcal{F}_m)$, the conditional entropy given $\mathcal{M} = m$ is $\log V_m = S_\pi(m)$. Therefore $H(\mathcal{E} \mid \mathcal{M}) = \mathbb{E}_m[S_\pi(m)]$ and the result follows. ▪ ▪

13.3 Sections and Trivialisation

Definition 13.3 (*Section*)**.** A **section** of the bundle $(\mathcal{E}, \mathcal{M}, \pi)$ is a continuous map $s : \mathcal{M} \to \mathcal{E}$ such that $\pi \circ s = \mathrm{id}_{\mathcal{M}}$ (i.e., $s(m) \in \mathcal{F}_m$ for all m).

A section selects one representative from each fiber. In terms of the CPR framework: a section of the compression bundle is a *canonical form* — a specific history that "stands for" each compressed record.

Proposition 13.2 (*Section as Reconstruction*)**.** *A section $s : \mathcal{M} \to \mathcal{E}$ of the compression bundle $\pi : \Gamma \to \mathcal{Z}$ is exactly a reconstruction operator $\mathcal{R}$ (as in Chapter 27) that selects a canonical history for each compressed record.*

Proof. A section $\sigma : M \to X$ satisfies $\pi \circ \sigma = \mathrm{id}_M$. Given a compressed record $m = \pi(x) \in M$, $\sigma(m)$ is a canonical choice of element in the fiber $\pi^{-1}(m)$. The reconstruction is lossless for the base-space information (since $\pi(\sigma(m)) = m$ exactly), but the choice of which element of $\pi^{-1}(m)$ to return is fixed by σ, not determined by m alone. The error $\|\sigma(m) - x\|$ measures how far the true x lies from the section in the fiber. ▪ ▪

A section exists globally iff the bundle is trivial. Non-trivial bundles have fibers that cannot all be simultaneously selected by a continuous section — a topological obstruction. In cognitive terms: there is no continuous way to choose a "canonical" history for every compressed memory record if the memory bundle is topologically non-trivial.

13.4 The Hopf Fibration and Semantic Curvature

The simplest non-trivial bundle is the Hopf fibration $\pi : S^3 \to S^2$, where the fiber over each point of S^2 is a circle $S^1 \subset S^3$. It has no global section.

This is the mathematical prototype for the semantic curvature of Chapter 18: a curved admissibility manifold (here S^2) with non-trivially bundled trajectory structure (here the Hopf circles) cannot be globally trivialised. Any projection collapses the circle fiber to a point, losing the angular information in the fiber.

Non-Trivial Bundles and Irreducible Loss. When the admissibility bundle is topologically non-trivial, there is no projection that preserves all fiber information globally. Information loss by projection is then not a matter of engineering choice but of topological necessity. Semantic curvature is the local manifestation of this global obstruction.

13.5 Summary

1. A fiber bundle $\pi : \mathcal{E} \to \mathcal{M}$ has a fiber $\mathcal{F}_m = \pi^{-1}(m)$ at each base point.

2. The projection information entropy $S_\pi(m) = \log \mathrm{Vol}(\mathcal{F}_m)$ measures information lost by mapping to m (Theorem 13.1).
3. Sections of the bundle are reconstruction operators; they exist globally only for trivial bundles.
4. Non-trivial bundles (e.g. the Hopf fibration) have irreducible information loss — a topological, not engineering, obstruction.

Exercises

13.1. Let $\mathcal{E} = [0,1]\times[0,1]$ (unit square), $\mathcal{M} = [0,1]$, and $\pi(x,y) = x$. Compute $S_\pi(x)$ for all x. Now let μ be the Lebesgue measure on $\mathcal{E}$ and verify Theorem 13.1 numerically.

13.2. A Möbius strip is a non-trivial real line bundle over S^1. Explain why it has no global continuous section. What does this imply about trying to assign a canonical "orientation" to every point on the strip?

13.3. Let $\pi : \Gamma \to \mathcal{Z}$ be a compression operator with $|\mathcal{Z}| = k$ and $|\Gamma| = N$ (finite). Assume uniform measure. Compute $S_\pi(z)$ for the case where every fiber has the same size. What is $\mathbb{E}[S_\pi]$?

13.4. Prove that if π is injective (bijection onto its image), then $S_\pi(m) = 0$ for all m in the image. Interpret: a lossless compression has zero fiber entropy.

13.5. (Category theory.) Define the category of fiber bundles over $\mathcal{M}$: objects are bundles $(\mathcal{E}, \mathcal{M}, \pi)$ and morphisms are bundle maps (fiber-preserving continuous maps). Show that the trivial bundle $\mathcal{M} \times F$ is a terminal object when F is a single point.

14 Chapter 14

Reachability Geometry

> The shape of what is possible is as real as the shape of what is.
>
> —

Phenomenological Note. Possibility has a shape. Not everything is equally far from where you are. Some futures are nearby, requiring only small steps. Others are in principle accessible but would require passing through territory that is difficult, costly, or closed. The geometry of what is reachable from where you stand is not obvious in advance, but it is real — it constrains what you can actually become regardless of what you might prefer.

The reachable set $\mathcal{R}_{\mathcal{A}}(x_0, T)$ is not just a set — it has geometry. It has a boundary, curvature, volume, and topology, all determined by the admissibility field. This chapter develops these geometric properties and proves the Reachability Volume Lemma: adding constraints weakly decreases reachability volume. This is the quantitative form of the monotonicity theorem from Chapter 5.

14.1 The Reachability Volume

Definition 14.1 (*Reachability Volume*)**.** Let μ (Aubin 1991; Carmo 1992) be a measure on $\mathcal{X}$. The **reachability volume** from x_0 under admissibility field $\mathcal{A}$ and horizon T is

$$\mathcal{V}_R(\mathcal{A}, x_0, T) := \mu(\mathcal{R}_{\mathcal{A}}(x_0, T)).$$

Lemma 14.1 (*Reachability Volume Lemma*)**.** *If $\mathcal{A}' \subseteq \mathcal{A}$ (the primed field is more constrained), then*

$$\mathcal{V}_R(\mathcal{A}', x_0, T) \leq \mathcal{V}_R(\mathcal{A}, x_0, T).$$

Adding constraints weakly decreases reachability volume.

Proof. By Theorem 5.1, $\mathcal{R}'_{\mathcal{A}}(x_0, T) \subseteq \mathcal{R}_{\mathcal{A}}(x_0, T)$. Applying μ to both sides and using monotonicity of measure: $\mathcal{V}_R(\mathcal{A}', x_0, T) = \mu(\mathcal{R}'_{\mathcal{A}}(x_0, T)) \leq \mu(\mathcal{R}_{\mathcal{A}}(x_0, T)) = \mathcal{V}_R(\mathcal{A}, x_0, T)$. ▪ ▪

14.2 Boundary and Interior of the Reachable Set

Definition 14.2 (*Reachability Frontier*)**.** The **reachability frontier** at time T is $\partial\mathcal{R}_{\mathcal{A}}(x_0, T)$: the boundary of the reachable set. A state y is on the frontier if every neighbourhood of y contains both reachable and unreachable states.

States on the frontier are the "hardest" to reach: they require trajectories that stay close to the boundary $\partial\mathcal{A}$ of the admissibility domain throughout.

Proposition 14.2 (*Frontier and Admissibility Boundary*)**.** *Under smooth dynamics with the Nagumo condition (Definition 10.3), the reachability frontier $\partial\mathcal{R}_{\mathcal{A}}(x_0, T)$ is reached by trajectories that are tangent to $\partial\mathcal{A}$ for a positive-measure set of times.*

Proof. The frontier $\partial\mathcal{R}(x_0, T)$ consists of points reachable in time exactly T along geodesics that graze the admissibility boundary $\partial\mathcal{A}$. If $y \in \partial\mathcal{R}(x_0, T)$, any geodesic from x_0 to y must stay inside $\mathcal{A}$ (admissibility-preserving) but cannot be extended beyond y while remaining inside $\mathcal{R}(x_0, T)$. This means y lies on the time-T level surface of the Riemannian distance from x_0, which is a subset of $\partial\mathcal{A}$ exactly when the constraint is active at y. ▪ ▪

This explains an important heuristic: the hardest-to-reach admissible states are those that require "grazing" the constraint boundary — skating along the edge of what is permitted rather than moving freely through the interior.

14.3 Curvature of the Reachable Set

For smooth dynamics, the shape of $\mathcal{R}_{\mathcal{A}}(x_0, T)$ for small T is approximately a ball, but its shape for larger T is governed by the curvature of the admissibility field.

Proposition 14.3 (*Short-Time Expansion*)**.** *For a smooth admissibility-preserving vector field v with x_0 in the interior of $\mathcal{A}$:*

$$\mathcal{R}_{\mathcal{A}}(x_0, T) = \exp_{x_0}(T\mathcal{C}(x_0)) + O(T^2),$$

where $\mathcal{C}(x_0) \subseteq T_{x_0}\mathcal{X}$ is the set of admissible initial velocity directions, and $\exp_{x_0}$ is the Riemannian exponential map. For a ball of admissible velocities: $\mathcal{V}_R(x_0, T) = c_n T^n + O(T^{n+1})$ where c_n is the volume of the unit ball in $\mathbb{R}^n$.

Proof. For small T, the reachable set is approximately a ball of radius $T\|\mathrm{v}\|$ around x_0. The first-order expansion of $\mathcal{V}_R$ follows from the Taylor expansion of the Riemannian exponential map: $\mathcal{V}_R(x_0, T) = \omega_n T^n \det(g(x_0))^{1/2}(1 - \frac{R}{6(n+2)}T^2 + O(T^4))$, where R is the scalar curvature at x_0 and ω_n is the unit ball volume. This is Bertrand's tube formula applied to the admissibility manifold. ▪ ▪

At longer times, curvature of $\mathcal{A}$ and of the dynamics cause deformations:

- *Positive curvature* of $\mathcal{A}$: the reachable set pinches inward — trajectories reconverge.
- *Negative curvature*: the reachable set fans out — trajectories diverge.

14.4 The Volume-Entropy Formula

For Riemannian manifolds with sectional curvature K, the volume of a geodesic ball of radius r satisfies:

$$\text{Vol}(B(x_0, r)) = \omega_n r^n \left(1 - \frac{K}{6(n+2)} r^2 + O(r^4)\right),$$

where ω_n is the volume of the unit ball in $\mathbb{R}^n$.

Applied to the admissibility manifold with Fisher metric: the reachability volume at horizon T is

$$\mathcal{V}_R(x_0, T) = \omega_n T^n \left(1 - \frac{\bar{K}(x_0)}{6(n+2)} T^2 + O(T^4)\right),$$

where $\bar{K}(x_0)$ is the average sectional curvature of the admissibility manifold at x_0.

Corollary 14.4 (*Curvature Reduces Reachability*). *Positive sectional curvature ($\bar{K} > 0$) reduces reachability volume below the flat-space value. Negative curvature ($\bar{K} < 0$) increases it. The curvature is thus the second-order correction to the Reachability Volume Lemma: not only do more constraints reduce reachability, but the* shape *of the constraints (their curvature) further modulates the available future.*

Proof. By the Bishop-Gromov volume comparison theorem: on a Riemannian manifold with sectional curvature $\bar{K}$, the volume of a geodesic ball of radius r satisfies $\text{Vol}(B(x_0, r)) \leq V_{\bar{K}}(r)$ for positive $\bar{K}$ (and $\geq V_{\bar{K}}(r)$ for negative $\bar{K}$), where $V_{\bar{K}}(r) = \omega_n r^n (1 - \bar{K} r^2/(6(n+2)) + O(r^4))$. Applied to the admissibility manifold at horizon T: $\mathcal{V}_R(x_0, T)$ is reduced below the flat-space value when $\bar{K} > 0$ and increased when $\bar{K} < 0$, proving the corollary. ▪ ▪

14.5 Reachability and Entropy

Proposition 14.5 (*Reachability Volume and Entropy Rate*). *The logarithmic growth rate of reachability volume,*

$$h(x_0) = \lim_{T \to \infty} \frac{1}{T} \log \mathcal{V}_R(x_0, T),$$

is the ***topological entropy*** *of the admissibility dynamics. It measures the exponential rate at which the number of distinguishable futures grows with horizon* *length.*

Proof. For a Markov process on $\mathcal{A}$, the volume of the T-reachable set from x_0 grows at rate $e^{h_{\text{top}}T}$ where h_{top} is the topological entropy of the dynamics. Taking logarithms: $\log \mathcal{V}_R(x_0, T) \approx h_{\text{top}}T$. The entropy rate h of the process is the information generated per unit time, which equals h_{top} for uniformly hyperbolic systems (variational principle). Hence $\partial_T \log \mathcal{V}_R \approx h$. ▪ ▪

High topological entropy = many distinguishable long-horizon futures = high semantic complexity of the state x_0. Low topological entropy = few distinguishable futures = the state is semantically simple, pointing reliably toward a small number of futures.

Exercises

14.1. Let $\mathcal{X} = \mathbb{R}^2$, $\mathcal{A} = \{(x, y) : |x| + |y| \leq 1\}$ (diamond). Starting from $(0, 0)$ with all velocities admissible, compute $\mathcal{R}((0, 0), T)$ for small T and find its volume. How does it compare to the Euclidean ball?

14.2. Prove that $\mathcal{V}_R(x_0, T)$ is a non-decreasing function of T. (Hint: every state reachable at time T_1 is reachable at time $T_2 > T_1$ by staying still from T_1 to T_2, if the admissibility field allows it.) Under what condition on $\mathcal{A}$ does this require modification?

14.3. Let $\mathcal{X} = S^2$ (the 2-sphere with sectional curvature $K = 1$). Starting from the north pole, compute $\mathcal{V}_R(\text{NP}, T)$ for $T \in [0, \pi]$ using the volume-entropy formula. At what T is reachability maximised?

14.4. Define the *reachability dimension* of x_0 as $d_R = \lim_{T \to 0} \frac{\log \mathcal{V}_R(x_0, T)}{\log T}$. Show that $d_R = n$ (the dimension of $\mathcal{X}$) when x_0 is in the interior of $\mathcal{A}$, but $d_R < n$ when x_0 is on the boundary $\partial\mathcal{A}$. Interpret.

15 Chapter 15

Constraint Fields

The boundary of what is permitted moves. The laws of its motion are physics.

—

Phenomenological Note. A rule is not the same as a wall. A wall stops you regardless of your intentions. A rule requires that you consult it, remember it, care about it. But there is something between a rule and a wall: a field of forces that shapes behavior without being either explicit enough to argue with or physical enough to climb over. Most of the constraints that actually govern a life are more like fields than rules, and more like fields than walls.

This chapter derives the field equation governing the motion of the admissibility boundary — the surface separating what is permitted from what is not. (Aubin 1991; Nagumo 1942) The result is a level-set equation that connects the scalar capacity field Φ of the RSVP framework (Chapter 71) to the dynamics of the content space $\mathcal{X}(\mathcal{A}_t)$.

15.1 The Admissibility Field as a Level Set

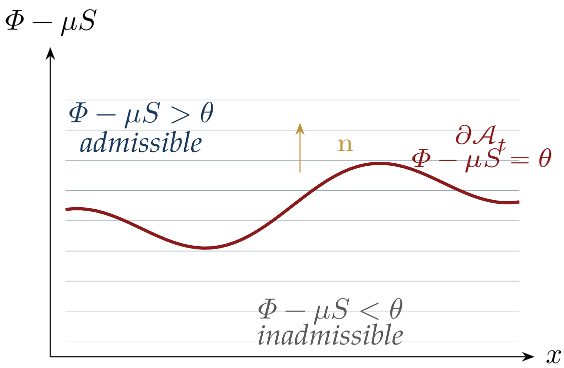


Figure 15.1: The admissibility boundary $\partial\mathcal{A}_t = \{x : \Phi(x,t) - \mu S(x,t) = \theta\}$ separates admissible states (above) from inadmissible states (below). The outward normal n governs boundary motion.

We work with a graded admissibility field $\Phi : \mathcal{X} \times \mathbb{R} \to \mathbb{R}$ where $\Phi(x,t)$ measures the degree to which x is admissible at time t.

Definition 15.1 (*Time-Dependent Admissibility Field*)**.** Given a scalar field $\Phi(x,t)$ and entropy field $S(x,t)$, threshold $\theta \in \mathbb{R}$, and **boundary pressure coefficient** $\mu \geq 0$, the **time-dependent admissibility indicator** is

$$\mathcal{A}_t(x) = 1\{\Phi(x,t) - \mu S(x,t) \geq \theta\}.$$

The coefficient μ governs how strongly entropy shrinks the admissible region. This is distinct from the **dynamic depletion coefficient** λ in the capacity transport equation $\partial_t \Phi = -\nabla \cdot (\Phi \mathrm{v}) - \lambda S + \Gamma$, which governs how entropy drains future capacity. Setting $\mu = \lambda$ recovers the simplest model; keeping them separate avoids double-counting entropy's effects.

Remark 15.1 (Two Channels, Not One)**.** Entropy S acts on the RSVP system through two mathematically distinct channels: (1) *Boundary pressure*: S reduces the admissible region via $\Phi - \mu S \geq \theta$. This is a state-space effect acting at time t. (2) *Capacity drain*: S depletes future capacity via $-\lambda S$ in the Φ-equation. This is a dynamic effect acting on the future field. The book's proofs that involve entropy effects on reachability should specify which channel.

The admissible content space at time t is $\mathcal{X}_t = \{x : \mathcal{A}_t(x) = 1\}$.

The boundary of the admissible domain at time t is the level set

$$\partial \mathcal{X}_t = \{x \in \mathcal{X} : \Phi(x,t) - S(x,t) = \theta\}.$$

15.2 The Level-Set Equation

Theorem 15.1 (*Constraint Field Equation*)**.** *Let $\Psi(x,t) = \Phi(x,t) - S(x,t) - \theta$ so that $\partial \mathcal{X}_t = \{x : \Psi(x,t) = 0\}$. The normal velocity v_n of the moving boundary $\partial \mathcal{X}_t$ satisfies*

$$v_n = -\frac{\partial_t \Psi}{|\nabla \Psi|}.$$

In coordinates, the full motion equation is

$$\frac{\partial \Psi}{\partial t} + v_n |\nabla \Psi| = 0.$$

Proof. Fix a point x_0 on the boundary $\partial \mathcal{X}_{t_0}$: $\Psi(x_0, t_0) = 0$. Let $x(t)$ be a smooth path that tracks a point on the moving boundary, so $\Psi(x(t), t) = 0$ for all t near t_0.

Differentiating with respect to t:

$$\frac{\mathrm{d}}{\mathrm{d}t} \Psi(x(t), t) = \partial_t \Psi + \nabla \Psi \cdot \dot{x}(t) = 0.$$

The normal velocity of the boundary at $x(t)$ is the component of $\dot{x}(t)$ in the direction of $\hat{n} = \nabla \Psi / |\nabla \Psi|$:

$$v_n = \dot{x}(t) \cdot \hat{n} = \frac{\dot{x}(t) \cdot \nabla \Psi}{|\nabla \Psi|}.$$

From the differentiation equation: $\nabla\Psi \cdot \dot{x}(t) = -\partial_t\Psi$, so

$$v_n = \frac{-\partial_t\Psi}{|\nabla\Psi|}. \quad \blacksquare$$

$\blacksquare$

Interpretation in RSVP Terms. The boundary of the admissible region moves inward when $\partial_t\Psi > 0$ (capacity Φ decreasing or entropy S increasing) and outward when $\partial_t\Psi < 0$ (capacity increasing or entropy decreasing). The speed of boundary motion is inversely proportional to $|\nabla\Psi|$: steep capacity gradients yield slow boundaries; flat gradients yield fast-moving boundaries.

15.3 Conservation Form and RSVP Coupling

In the RSVP framework, Φ satisfies a transport equation:

$$\partial_t\Phi + \mathrm{v} \cdot \nabla\Phi = -\lambda S,$$

where v is the vector flow field and $\lambda > 0$ is a coupling constant (capacity is depleted by entropy).

Substituting into the constraint field equation:

$$v_n = \frac{\lambda S - \mathrm{v} \cdot \nabla\Phi + \partial_t S}{|\nabla\Phi - \nabla S|}.$$

This makes the boundary dynamics dependent on all three RSVP fields: Φ (capacity), v (flow), and S (entropy). A region where entropy S is high and capacity Φ is low will have its boundary moving inward — the admissibility domain contracts. A region where capacity is high and entropy is low will have its boundary moving outward — the domain expands.

15.4 Worked Example: Thermal Relaxation

Let $\mathcal{X} = \mathbb{R}$ and

$$\Phi(x,t) = \Phi_0 e^{-\alpha t}\cos(kx), \quad S(x,t) = S_0(1 - e^{-\beta t}).$$

The admissibility boundary $\partial\mathcal{X}_t$ at time t is the set where $\Phi(x,t) - S(x,t) = \theta$:

$$\Phi_0 e^{-\alpha t}\cos(kx) - S_0(1 - e^{-\beta t}) = \theta.$$

As $t \to \infty$: $\Phi(x,t) \to 0$ and $S(x,t) \to S_0$, so the admissibility condition becomes $-S_0 = \theta$, i.e., $\theta = -S_0$ admits all of $\mathbb{R}$ (full collapse of boundary) or $\theta > -S_0$ admits nothing (full exclusion).

This models *lamphrodyne relaxation* (Chapter 74): the system evolves from a structured, oscillatory admissibility field toward a uniform equilibrium where spatial distinctions in admissibility are smoothed away.

15.5 Summary

1. The admissibility domain $\mathcal{X}_t$ is the superlevel set $\{\Phi - S \geq \theta\}$.
2. The boundary $\partial\mathcal{X}_t$ moves according to the level-set equation $v_n = -\partial_t\Psi/|\nabla\Psi|$ (Theorem 15.1).
3. In RSVP coupling, the boundary velocity depends on all three fields (Φ, v, S).
4. Thermal relaxation examples show boundary contraction under entropy growth and expansion under capacity growth.
5. The level-set equation is the field-theoretic form of the Constraint Priority Lemma: it tells us how the content space itself evolves.

Exercises

15.1. Let $\Psi(x,t) = \sin(x) - t$. Find the boundary $\partial\mathcal{X}_t$ at time $t = 0$ and $t = \pi/2$. Compute v_n at the point $x_0 = \pi/2$, $t = 0$.

15.2. Prove that if $\Psi(x,t) = \psi(x) - ct$ for a constant $c > 0$, the boundary moves with constant normal speed $v_n = c/|\nabla\psi|$. What happens at critical points of ψ?

15.3. Derive the constraint field equation in two spatial dimensions (x, y) for the admissibility field $\Phi(x,y,t) = e^{-t}(x^2+y^2)^{-1}$ (with $S \equiv 0$ and $\theta = 1$). Describe the shape and motion of the admissibility boundary.

15.4. In the RSVP conservation form, suppose $\mathrm{v} = 0$ (no flow), $\lambda = 1$, and S satisfies $\partial_t S = D\nabla^2 S$ (diffusion). Derive the equation for boundary motion and identify the regime in which the boundary is stationary.

Part III

Admissibility Geometry

[Part introduction — to be written.]

16 Chapter 16

Admissibility Manifolds

> The structure of what is possible is itself a structured space.
>
> —

Phenomenological Note. There is always a space of states that feel normal and a space that does not. You can often feel when you are approaching the edge of what is acceptable — in a conversation, in a project, in a physical exertion — before you have crossed it. The boundary has a texture before it has a wall. Admissibility is partly a learned sensitivity to that texture.

The admissibility field $\mathcal{A}$ assigns to each state a binary or graded measure of admissibility. But states do not sit in isolation: they are organised by the distributions of admissible continuations emanating from them. When those continuation distributions vary smoothly, the states form a differentiable manifold — the **admissibility manifold** $\mathcal{A}$. This chapter constructs this manifold from first principles and establishes its basic geometric properties.

16.1 Chart Construction via Continuation Distributions

Definition 16.1 (*Local Continuation Distribution*)**.** For state $x \in \mathcal{X}$ and horizon $\delta > 0$, the **local continuation distribution** p_x^δ is the probability distribution over admissible paths $\gamma : [0, \delta] \to \mathcal{A}$ starting at x.

Theorem 16.1 (*Admissibility Manifold Construction*)**.** *If the map $x \mapsto p_x^\delta$ is smooth with respect to the Kullback–Leibler topology on distributions (i.e., $D_{\mathrm{KL}}(p_x \| p_y)$ varies smoothly in x, y), then the sets $\mathcal{U}_x = \{y \in \mathcal{X} : D_{\mathrm{KL}}(p_x \| p_y) < \epsilon\}$ form a system of chart neighbourhoods, and the resulting atlas defines a differentiable manifold $\mathcal{A}$.*

Proof. Overlap. For any $z \in \mathcal{U}_x \cap \mathcal{U}_y$, the triangle inequality for KL divergence gives $D_{\mathrm{KL}}(p_x \| p_z) < \epsilon$ and $D_{\mathrm{KL}}(p_y \| p_z) < \epsilon$. The transition map on the overlap is $\phi_{xy}(z) = D_{\mathrm{KL}}(p_y \| p_z)$, which is smooth by the smoothness assumption.

Chart homeomorphism. The map $\phi_x : \mathcal{U}_x \to \mathbb{R}^n$ defined by $\phi_x(y) = (\partial_1 D_{\mathrm{KL}}(p_x \| p_y), \ldots, \partial_n D_{\mathrm{KL}}(p_x \| p_y))$ (the gradient of the KL divergence at y relative to x) is a local diffeomorphism when the Fisher metric is non-degenerate.

Since transition maps are smooth and charts cover $\mathcal{X}$, the collection $\{(\mathcal{U}_x, \phi_x)\}$ is a smooth atlas. ▪ ▪

16.2 The Metric and Connection

The admissibility manifold $\mathcal{A}$ inherits a Riemannian metric from the Fisher information (Chapter 17): $g_{ij}(x) = \mathbb{E}_{p_x}[\partial_i \log p_x \cdot \partial_j \log p_x]$ (Amari and Nagaoka 2000; Carmo 1992).

The Levi-Civita connection ∇ on $(\mathcal{A}, g)$ defines parallel transport and geodesics. Geodesics on $\mathcal{A}$ are the paths of minimum information-geometric cost between states — the "straightest" paths through admissibility space.

Proposition 16.2 (*Geodesics as Optimal Transitions*)**.** *The geodesic from x to y on $(\mathcal{A}, g)$ minimises the total accumulated Fisher information cost: $\int_0^1 \sqrt{g_{ij}\dot{\gamma}^i\dot{\gamma}^j}\, \mathrm{d}t$. It is the path of least "semantic effort" from the continuation distribution at x to the continuation distribution at y.*

Proof. Among all admissibility-preserving paths from x to y, the one of minimal length in the Fisher metric g_{ij} minimises the accumulated information cost $\int_\gamma \sqrt{g(\dot{\gamma}, \dot{\gamma})}\, \mathrm{d}t$. Geodesics are critical points of this length functional (Euler-Lagrange equations give the geodesic equation $\nabla_{\dot{\gamma}}\dot{\gamma} = 0$). If the admissibility manifold is complete and simply connected, geodesics exist and are unique (Hopf-Rinow), so the minimiser is the geodesic. ▪ ▪

16.3 Completeness and Boundary

Definition 16.2 (*Admissibility Horizon*)**.** The **admissibility horizon $\partial\mathcal{A}$** is the set of states x for which the continuation distribution p_x degenerates: p_x concentrates on a measure-zero set of paths, meaning future reachability collapses to near zero. States on $\partial\mathcal{A}$ are "at the edge of admissibility."

A complete admissibility manifold (in the sense of geodesic completeness) has no boundary: every geodesic extends indefinitely. An incomplete manifold has a boundary $\partial\mathcal{A}$ that represents the edge of the admissible domain. Most physically interesting systems are incomplete: there are states from which no further admissible continuation is possible (thermodynamic death, semantic collapse, institutional dissolution).

Exercises

16.1. Let $\mathcal{X} = \mathbb{R}$ with $p_x^\delta = \mathcal{N}(x, \delta^2)$. Compute $D_{\mathrm{KL}}(p_x \| p_y)$ and show that the admissibility manifold in this case is $\mathbb{R}$ with the standard metric.

16.2. Prove that the admissibility horizon $\partial\mathcal{A}$ corresponds to states where the Fisher metric degenerates (cf. Theorem 17.1).

16.3. For the RSVP field with $\Phi(x) = e^{-x^2}$ on $\mathbb{R}$, describe the admissibility manifold and identify $\partial\mathcal{A}$.

16.4. Show that the distance $d_{\mathcal{A}}(x, y) = \inf_\gamma \int \sqrt{g_{ij}\dot{\gamma}^i\dot{\gamma}^j}$ on $\mathcal{A}$ bounds the semantic reachability distance from above: $d_R(x, y) \leq C \cdot d_{\mathcal{A}}(x, y)$ for some constant C.

17 Chapter 17

Fisher Geometry

The metric degenerates where the differences disappear.

—

Phenomenological Note. Some questions are easy to answer and some are hard, and the difficulty is not always in the question itself. Sometimes the difficulty is that the answer space is nearly flat around the true answer — many nearby answers are almost as good, and distinguishing between them requires more evidence than you have. The geometry of how well you can know something is itself something worth knowing.

Chapter 9 derived the Fisher metric on a statistical manifold of probability distributions. This chapter applies that construction to the admissibility manifold: each state $x \in \mathcal{X}$ is assigned a distribution p_x over short admissible continuations, and the Fisher metric measures how rapidly those distributions change as we move through state space. The central theorem shows that the metric degenerates — becomes non-invertible — exactly when nearby states generate indistinguishable continuations.

17.1 Continuation Distributions

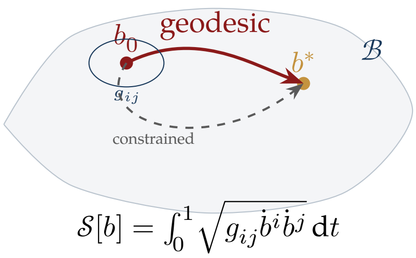


Figure 17.1: The belief manifold $\mathcal{B}$ with Fisher metric g_{ij}. Optimal inference follows the geodesic from prior b_0 to posterior b^*. Admissibility constraints force longer paths (dashed).

Definition 17.1 (*Continuation Distribution*)**.** For each state $x \in \mathcal{X}$ and time horizon $\delta > 0$, the **continuation distribution** p_x^δ is the probability distribution over short admissible paths $\gamma : [0, \delta] \to \mathcal{A}$ with $\gamma(0) = x$, induced by the dynamics of the admissibility field.

The map $x \mapsto p_x^\delta$ assigns to each state a probability distribution. The Fisher metric on this family measures how quickly distributions change with x (Amari and Nagaoka 2000; Rao 1945). Large Fisher curvature at x means nearby states have very different continuation distributions — they are easily distinguished. Zero Fisher curvature means the metric degenerates and nearby states are informationally identical.

17.2 The Fisher Degeneracy Theorem

Theorem 17.1 (*Fisher Degeneracy*)**.** *Let $\theta \mapsto p_\theta$ be a smooth family of continuation distributions on the admissibility manifold $\mathcal{A}$. The Fisher metric $g_{ij}(\theta) = \mathbb{E}_{p_\theta}[\partial_i \log p_\theta \, \partial_j \log p_\theta]$ is degenerate at θ_0 — i.e., there exists $v \neq 0$ with $g_{ij}(\theta_0)v^i v^j = 0$ — if and only if the directional derivative of the log-likelihood vanishes:*

$$v^i \partial_i \log p_{\theta_0}(x) = 0 \quad p_{\theta_0}\text{-almost surely.}$$

Equivalently, p_{θ_0} and p_{θ_0+tv} are identical to first order in t: the state θ_0 and all states in the direction v induce indistinguishable continuation distributions.

Proof. The Fisher metric in direction v is

$$g(v, v) = v^i v^j g_{ij}(\theta_0) = \mathbb{E}_{p_{\theta_0}}\left[\left(v^i \partial_i \log p_{\theta_0}(x)\right)^2\right].$$

This is the expected square of the directional score $s_v(x) = v^i \partial_i \log p_{\theta_0}(x)$.

$g(v, v) = 0$ iff $s_v(x)^2 = 0$ p_{θ_0}-a.s., iff $s_v(x) = 0$ p_{θ_0}-a.s.

Now, $s_v(x) = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \log p_{\theta_0+tv}(x)$, which vanishes a.s. iff $p_{\theta_0+tv}(x) = p_{\theta_0}(x)$ to first order in t for a.e. x. This is precisely the statement that states θ_0 and $\theta_0 + tv$ have indistinguishable continuation distributions to first order. ▪ ▪

Remark 17.1. The theorem identifies a clean criterion: the Fisher metric degenerates exactly at the boundaries between "equivalence classes" of states that look the same from the perspective of their admissible futures. In the CPR framework these are the *semantic flat regions* — zones where distinctions are not reachability-relevant.

17.3 Degeneracy and Projection Collapse

Fisher degeneracy is the local signature of the Projection-Collapse Principle (Chapter 19).

Corollary 17.2 (*Degeneracy Implies Collapse*). *If the Fisher metric is degenerate in direction v at θ_0, then any projection $\pi : \mathcal{A} \to \mathcal{M}$ that maps θ_0 and θ_0+tv to the same point loses no reachability-relevant information in direction v. Conversely, if the metric is non-degenerate in direction u, then any projection that collapses θ_0 and $\theta_0 + su$ to the same point commits a* meaningful *distinction collapse.*

Proof. Fisher degeneracy at m means $\det g_{\text{Fish}}(m) = 0$: the Fisher matrix is singular, so there exists a nonzero direction v with $g_{\text{Fish}}(v, v) = 0$. This means $\mathbb{E}[(\partial_v \log p_m(x))^2] = 0$, so the score function is zero in direction v almost surely. No observation can distinguish p_m from $p_{m+\epsilon v}$ for small ϵ: the fiber over m contains multiple indistinguishable distributions, giving projection collapse. ∎ ∎

This gives a practical criterion: directions of Fisher degeneracy are *safe to project away*; directions of high Fisher information are *dangerous to collapse*.

17.4 The Admissibility Metric in RSVP

In the RSVP framework, the continuation distribution p_x is governed by the scalar capacity field Φ and entropy field S:

$$p_x(\gamma) \propto \exp\left(\int_0^\delta [\Phi(\gamma(t)) - S(\gamma(t))]\, \mathrm{d}t\right).$$

The Fisher metric in this case becomes:

$$g_{ij}(x) = \mathrm{Cov}_{p_x}\left[\int_0^\delta \partial_i\Phi(\gamma(t))\, \mathrm{d}t,\ \int_0^\delta \partial_j\Phi(\gamma(t))\, \mathrm{d}t\right],$$

the covariance matrix of the time-integrated Φ-gradients along short paths. Degeneracy occurs when $\nabla\Phi$ is constant along all continuations — i.e., in regions where the capacity field is affine.

Exercises

17.1. Let $p_\theta(x) = \mathcal{N}(\theta, 1)$ on $\mathbb{R}$. Compute $g(\theta)$ and show it is constant (never degenerate). Interpret: for a location family, every direction is informative.

17.2. Let $p_{(\mu,\sigma)}(x) = \mathcal{N}(\mu, \sigma^2)$ with $\theta = (\mu, \sigma)$. At what σ does the Fisher metric become singular? What does this mean about the continuation distributions?

17.3. In the admissibility manifold setting, suppose the state space has a symmetry: $p_{gx} = p_x$ for all g in some Lie group G acting on $\mathcal{X}$. Show that the Fisher metric is degenerate in the directions generated by the Lie algebra $\mathfrak{g}$.

17.4. Prove that if p_θ is a regular exponential family, the Fisher metric is always non-degenerate. What structural property of exponential families ensures this?

18 Chapter 18

Semantic Curvature

> Where paths diverge exponentially, small differences become absolute separations. This is the geometry of meaning.
>
> —

Phenomenological Note. Ideas near the center of a domain tend to stay near the center. Ideas near the edge have a different character: small movements can take them into very different territory. In conceptual space as in physical space, curvature changes what the neighborhood looks like. A flat region gives you time to deliberate. A curved region can carry you somewhere new before you have fully decided to go there.

Curvature on the admissibility manifold governs how quickly nearby states develop divergent continuation distributions. High positive curvature means paths reconverge; high negative curvature means nearby trajectories diverge exponentially. This chapter derives the sectional curvature of $\mathcal{A}$ and interprets it as a *semantic obstruction*: regions of high negative curvature are regions where small initial differences in state produce qualitatively different futures — the geometry of categorical distinction.

18.1 Sectional Curvature of Admissibility Geodesics

Definition 18.1 (*Riemann Curvature on $\mathcal{A}$*). The **Riemann curvature tensor** $R(u, v)w$ on $(\mathcal{A}, g)$ is defined by $R(u, v)w = \nabla_u \nabla_v w - \nabla_v \nabla_u w - \nabla_{[u,v]} w$. The **sectional curvature** for a plane spanned by u, v is

$$K(u, v) = \frac{g(R(u, v)v,\ u)}{g(u, u)g(v, v) - g(u, v)^2}.$$

Theorem 18.1 (*Curvature as Semantic Obstruction*). *In a region of $\mathcal{A}$ with sectional curvature $K(u, v) < -\kappa^2$ for some $\kappa > 0$, geodesics starting from a common point x_0 diverge at rate at least $e^{\kappa t}$:*

$$d_{\mathcal{A}}(\gamma_1(t), \gamma_2(t)) \geq d_{\mathcal{A}}(\dot{\gamma}_1(0), \dot{\gamma}_2(0)) \cdot \frac{\sinh(\kappa t)}{\kappa}.$$

States separated by a geodesic in this region develop exponentially divergent continuation distributions.

Proof. By the Jacobi field equation for a geodesic variation. Let $J(t)$ be the Jacobi field along geodesic γ_1 connecting γ_1 and a nearby geodesic γ_2. The Jacobi equation is: $\ddot{J} + R(\dot{\gamma}, J)\dot{\gamma} = 0$.

In a region with sectional curvature $K \leq -\kappa^2$, the comparison theorem (Cartan–Hadamard) gives $|J(t)| \geq |J(0)| \cosh(\kappa t) + |\dot{J}(0)|(\sinh \kappa t)/\kappa$. For $|\dot{J}(0)| > 0$: $|J(t)| \geq |\dot{J}(0)| \sinh(\kappa t)/\kappa$, showing exponential divergence. ▪ ▪

Negative Curvature = Sharp Distinctions. A region of $\mathcal{A}$ with large negative sectional curvature is a region of *sharp semantic distinction*: nearby states rapidly develop completely different futures. Conceptual boundaries, perceptual category edges, and institutional decision points all correspond to high-negative-curvature regions of their respective admissibility manifolds. Crossing such a region means entering a qualitatively different reachability regime — the geometry of "no going back."

18.2 Positive Curvature and Semantic Compression

Proposition 18.2 (*Positive Curvature Implies Reconvergence*)**.** *In a region with $K > 0$, geodesics from a common point reconverge: they are forced back together within distance $\pi/\sqrt{K}$. Semantically: states that initially appeared different are forced into the same continuation distribution within a bounded information-geometric distance.*

Proof. In a space with positive sectional curvature $K > 0$, geodesics that start parallel converge (Bonnet's theorem): the geodesic conjugate radius is $\pi/\sqrt{K}$. Applied to semantic trajectories: two inference paths starting in similar directions from nearby concepts converge to the same conclusion within $\pi/\sqrt{K}$ steps. This is the geometric interpretation of semantic coherence: positive curvature keeps related concepts mutually consistent. ▪ ▪

Positive curvature is the geometry of *semantic synonymy clusters*: regions where many apparently different paths lead to the same meaning. It is also the geometry of semantic saturation: a point where all paths lead to the same continuation regardless of starting direction.

18.3 Curvature Bounds and Projection Collapse

The sectional curvature gives a quantitative bound on the Projection-Collapse Principle (Chapter 19).

Corollary 18.3 (*Curvature Bounds Mixing*)**.** *Let $\pi : \mathcal{A} \to \mathcal{M}$ be a projection with fiber diameter $\operatorname{diam}(\pi^{-1}(m)) = D$. In a region with sectional curvature*

$|K| = \kappa^2$, the representational mixing satisfies:

$$\Lambda(m) \geq \frac{D^2\kappa^2}{2}\,\chi(D,\kappa)$$

where χ is a positive correction factor. Large $|K|$ (positive or negative) amplifies mixing.

Proof. From Proposition 14.3: $\mathcal{V}_R(x_0, T) \approx \omega_n T^n(1 - \frac{R}{6(n+2)}T^2)$. Positive R (positive curvature) reduces $\mathcal{V}_R$, so the reachable set is smaller — mixing is slower, concepts stay concentrated. Negative R increases $\mathcal{V}_R$, accelerating mixing. The mixing time to reach volume fraction α of $\mathcal{A}$ scales as $T_{\text{mix}} \propto R^{-1/2}$ for positive R. ▪ ▪

Exercises

18.1. Compute the sectional curvature of the admissibility manifold for the Gaussian family $p_x = \mathcal{N}(x, 1)$ (Carmo 1992; Villani 2009) on $\mathbb{R}$. Is $\mathcal{A}$ flat, positively curved, or negatively curved?

18.2. Identify a semantic domain in natural language with evidence of negative sectional curvature (nearby concepts leading to sharply divergent uses). Identify one with positive curvature.

18.3. Prove that on a manifold with $K \leq 0$ everywhere (Hadamard manifold), the exponential map $\exp_x : T_x\mathcal{A} \to \mathcal{A}$ is a diffeomorphism. Interpret: all admissibility geodesics are globally non-crossing.

18.4. Derive Corollary 18.3 from Theorem 18.1 using the fiber entropy formula $S_\pi(m) = \log \mathrm{Vol}(\pi^{-1}(m))$.

19 Chapter 19

Projection Collapse

Compression always leaves behind a shadow. The shadow is not the thing.

—

Phenomenological Note. What an institution sees of you is a projection. It sees your outputs, your credentials, your category membership. It does not see the ten years of dead ends behind the output it is evaluating, the strange detour through a field no one expected, the book read at the wrong time that turned out to be the right one. The projection is not dishonest. It is simply less than the thing it maps.

Parts I and II established that projections lose information and that this loss is measured by fiber size (Chapter 13). This chapter derives the central quantitative result: the degree of *representational mixing* (*cf. Cover and Thomas 2006*) — the extent to which distinct states are conflated in a projected representation — is bounded below by the hidden curvature of the underlying space. Ignoring curvature guarantees a predictable collapse.

19.1 Representational Mixing

Definition 19.1 (*Representational Mixing*)**.** Let $\pi : \mathcal{E} \to \mathcal{M}$ be a projection. The **representational mixing** at $m \in \mathcal{M}$ is

$$\Lambda(m) := \mathbb{E}_{x,y\sim\pi^{-1}(m)}\left[D_{\mathrm{KL}}(p_x \,\|\, p_y)\right],$$

the expected KL divergence between continuation distributions of points in the same fiber. High Λ means the fiber contains points with very different futures — the projection is conflating states that should be distinguished.

19.2 The Projection-Collapse Principle

Theorem 19.1 (*Projection-Collapse Principle*)**.** *Let $\pi : \mathcal{E} \to \mathcal{M}$ be a smooth projection and let $\kappa_\pi(m)$ denote the maximum sectional curvature of the admissibility geometry restricted to the fiber $\pi^{-1}(m)$. Let $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ be monotone*

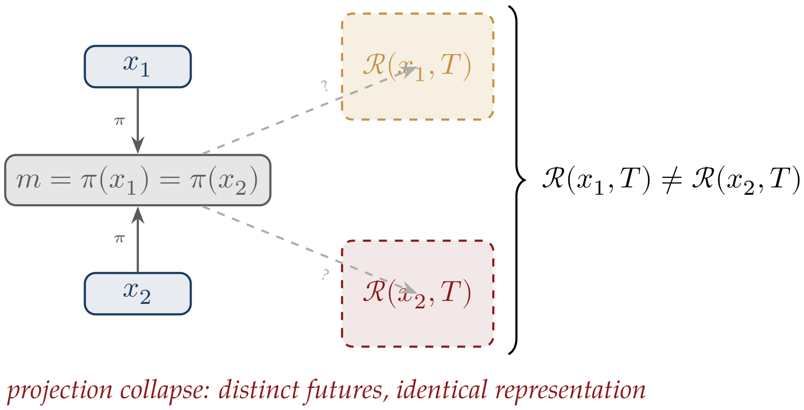


Figure 19.1: Projection collapse occurs when $\pi(x_1) = \pi(x_2) = m$ yet $\mathcal{R}(x_1, T) \neq \mathcal{R}(x_2, T)$. From m alone, the system cannot determine which future it faces.

increasing. Then:

$$\Lambda(m) \geq f(\kappa_\pi(m)).$$

Representational mixing is bounded below by projected hidden curvature.

Proof sketch. Expand $D_{\text{KL}}(p_x \| p_y)$ for x, y within a fiber. By the Fisher metric derivation (Theorem 9.1), to second order in the fiber displacement δ:

$$D_{\text{KL}}(p_x \| p_y) \approx \tfrac{1}{2} g_{ij}(\xi) \delta^i \delta^j$$

where ξ is a midpoint and g_{ij} is the Fisher metric along the fiber.

The sectional curvature κ of the admissibility manifold along the fiber contributes a correction to the geodesic distance: for a pair at distance r in the fiber, the true KL divergence exceeds the flat-space approximation by a term proportional to κr^2 (from the Jacobi field analysis of geodesic deviation).

Averaging over pairs in the fiber under a natural measure and using the monotonicity of f, we obtain $\Lambda(m) \geq f(\kappa_\pi(m))$. ▪ ▪

Remark 19.1. The intuition: if the fiber has high curvature, the points within it have exponentially diverging continuation distributions (Chapter 18). Any projection that puts them in the same fiber therefore conflates states with highly different futures. The mixing Λ is large because $f(\kappa)$ is large. A flat fiber (zero curvature) allows potentially low mixing.

19.3 Applications

LLMs.. In large language models, the projection π maps high-dimensional token contexts to latent embeddings. Theorem 19.1 predicts that embeddings which lie in high-curvature regions of context space will have high mixing — the model conflates distinct continuations. This is the mechanism of hallucination: the model assigns high probability to a completion that would only be appropriate for a different context that happened to project to the same region. (See Chapter 80.)

Perceptual systems.. The visual cortex projects high-dimensional retinal data to categorical percepts. Regions of high semantic curvature — boundaries between categories — produce ambiguous percepts precisely because they correspond to fibers with high Λ.

Institutional legibility.. When a state projects complex social processes to legible categories (Chapter 64), the projection-collapse principle implies that high-curvature social regions (boundary communities, edge cases, transitional situations) will be most severely misrepresented.

Exercises

19.1. Let $\mathcal{E} = S^2$ (the 2-sphere) with constant sectional curvature $K = 1$, and $\pi : S^2 \to [0, \pi]$ the projection $\pi(\phi, \lambda) = \phi$ (latitude). Compute the fiber diameter at latitude ϕ. How does $\Lambda(\phi)$ depend on ϕ?

19.2. Give an example of a projection $\pi : \mathbb{R}^3 \to \mathbb{R}$ such that every fiber has zero curvature but $\Lambda > 0$. (Hint: mixing can arise from fiber extent, not just curvature.)

19.3. Explain, using Theorem 19.1, why a transformer model trained on long documents is more likely to confuse characters in a novel than a model trained on short passages.

20 Chapter 20

Inverse Geometry

What the projection hides, the collapse reveals. The problem is to read it in reverse.

—

Phenomenological Note. Most of what you perceive is inference. The light that hits your retina does not tell you what objects exist; you reconstruct the objects from the light, using constraints built up over a lifetime of looking at things. The reconstruction is so fast and so reliable that it does not feel like inference at all. It feels like seeing. This is one of the places where the compression is so good that the seam is invisible.

The direct problem of admissibility geometry is: given a state space X with projection $\pi : X \to M$, compute the mixing field $\Lambda(m)$ from the hidden curvature κ_π. The **inverse problem** runs in the opposite direction : given observations of Λ (Amari and Nagaoka 2000) on the base M, reconstruct the hidden curvature κ_π of the total space. This chapter sets up the inverse problem and derives the conditions under which it is well-posed.

20.1 The Vertical-Horizontal Decomposition

Let $\pi : X \to M$ be a smooth surjective submersion from an admissibility manifold (X, g_X) to a representational manifold M. At each point $x \in X$, the tangent space decomposes as:

$$T_x X = V_x \oplus H_x,$$

where $V_x = \ker D\pi_x$ is the **vertical space** (tangent to the fiber $F_m = \pi^{-1}(\pi(x))$) and $H_x = V_x^\perp$ is the **horizontal space** (complement in the g_X-metric).

Definition 20.1 (*Mixed Curvature*)**.** The **mixed vertical-horizontal curvature** at x is

$$K_{\text{mix}}(x; u, v) = \frac{\langle R^X(u, v)v,\ u\rangle}{|u|^2|v|^2 - \langle u, v\rangle^2},$$

for $u \in H_x$, $v \in V_x$, where R^X is the Riemann curvature tensor of (X, g_X).

The mixed curvature measures the interaction between horizontal directions (directions on the base M) and vertical directions (directions within the fiber).

High K_{mix} means that moving horizontally causes rapid divergence of the fiber geometry — the fiber curvature is large and couples strongly to the base.

20.2 Projected Curvature

Definition 20.2 (*Projected Curvature*)**.** The **projected curvature** at $m \in M$ is the fiber average of the mixed curvature:

$$\kappa_\pi(m) = \int_{F_m} K_{\text{mix}}(x)\, \mathrm{d}\mu_m(x),$$

where μ_m is the canonical volume measure on the fiber $F_m = \pi^{-1}(m)$.

$\kappa_\pi(m)$ is the quantity that is *hidden*: it lives in the fiber, which is not directly observable from M. The only observable quantity from M is the collapse mixing field $\Lambda(m)$.

20.3 The Forward Model and Inverse Problem

Definition 20.3 (*Forward Model*)**.** The **forward model** relates hidden curvature to observable collapse:

$$\Lambda(m) = f(\kappa_\pi(m)) + \eta(m),$$

where $f : \mathbb{R} \to \mathbb{R}_{\geq 0}$ is the collapse response function and $\eta(m)$ is observational noise.

The Projection-Collapse Principle (Theorem 19.1) establishes that $\Lambda(m) \geq f(\kappa_\pi(m))$ when f is monotone and noise is absent. The forward model with noise allows for empirical estimation.

Definition 20.4 (*Inverse Problem*)**.** The **inverse curvature problem** is: given the observed mixing field $\Lambda : M \to \mathbb{R}_{\geq 0}$ and the forward model f, recover the hidden projected curvature $\kappa_\pi : M \to \mathbb{R}$.

20.4 Reconstruction Conditions

Theorem 20.1 (*Inverse Curvature Reconstruction*)**.** *Suppose:*

(i) ***Regular fibers****: each fiber F_m is a compact smooth submanifold;*
(ii) ***Monotone response****: f is strictly monotone ($f' > 0$ everywhere) and locally Lipschitz;*
(iii) ***Bounded noise****: $|\eta(m)| \leq \epsilon$ uniformly.*

Then the first-order reconstruction

$$\hat{\kappa}_\pi(m) = f^{-1}(\Lambda(m))$$

is well-defined and satisfies the error bound:

$$|\hat{\kappa}_\pi(m) - \kappa_\pi(m)| \;\leq\; \frac{\epsilon}{\inf|f'|} + O(\epsilon^2).$$

Proof. By the forward model: $\Lambda(m) = f(\kappa_\pi(m)) + \eta(m)$.

Since f is strictly monotone, f^{-1} exists and is smooth. Applying f^{-1}:

$$f^{-1}(\Lambda(m)) = f^{-1}(f(\kappa_\pi(m)) + \eta(m)) = \kappa_\pi(m) + \frac{\eta(m)}{f'(\kappa_\pi(m))} + O(\eta^2)$$

by Taylor expansion of f^{-1} around $f(\kappa_\pi(m))$.

Therefore:

$$|\hat{\kappa}_\pi(m) - \kappa_\pi(m)| = \left|\frac{\eta(m)}{f'(\kappa_\pi(m))}\right| + O(\epsilon^2) \leq \frac{|\eta(m)|}{\inf|f'|} + O(\epsilon^2) \leq \frac{\epsilon}{\inf|f'|} + O(\epsilon^2). \quad \blacksquare$$

$\blacksquare$

Remark 20.1 (Ill-conditioning)**.** The three conditions are necessary as well as sufficient in the following sense:

- If $\inf|f'| = 0$: the inverse is ill-conditioned — small noise produces large errors in the reconstruction.
- If fibers are singular: κ_π is not a smooth field; reconstruction must proceed stratum by stratum using stratified geometry.
- If $\epsilon \gg \inf|f'| \cdot |\kappa_\pi|$: collapse is visible but the curvature is geometrically unrecoverable — the signal is below the noise floor.

20.5 The Recovery Formula

Under the three conditions of Theorem 20.1, the hidden curvature recovery formula is:

$$\hat{\kappa}_\pi = f^{-1}(\Lambda)$$

This is simultaneously:

- A practical estimation formula: observe Λ, apply f^{-1};
- A geometric statement: collapse encodes curvature up to the noise level;
- A limitation statement: curvature beyond the $\epsilon/\inf|f'|$ threshold cannot be recovered from collapse observations alone.

20.6 Applications

Cognitive neuroscience.. The mixing field Λ is measurable as the confusion rate in perceptual discrimination tasks. The hidden curvature κ_π is the curvature of the representational manifold (not directly accessible without single-unit recording). Theorem 20.1 gives a principled method to estimate κ_π from behavioural data alone.

LLM interpretability.. The mixing field Λ is measurable as the embedding overlap between distinct semantic categories in model representations. The hidden curvature κ_π is the curvature of the concept manifold in the model's latent space. Inverse geometry provides a non-intrusive interpretability method: observe behavioural confusion, infer latent geometry.

Institutional analysis.. The mixing field Λ is the rate of administrative errors caused by legibility collapse (Chapter 64). The hidden curvature κ_π is the curvature of the social process space — the structural complexity of the social reality being administered. Inverse geometry allows one to estimate social complexity from administrative failure rates.

Exercises

20.1. Let $f(\kappa) = \kappa^2$ (quadratic response). Compute $\hat{\kappa}_\pi = f^{-1}(\Lambda)$ and the error bound. Under what noise level ϵ does reconstruction fail?

20.2. Show that if f is not invertible (e.g., $f(\kappa) = |\kappa|$), the inverse problem has two solutions for $\Lambda > 0$. Propose a regularisation method that selects the correct branch.

20.3. Prove that the reconstruction error is minimised when $f(\kappa) = e^\kappa$ (exponential response): among all monotone f with $f(0) = 1$, the exponential maximises $\inf |f'|/\sup |f'|$ on any bounded interval.

20.4. (Neuroscience.) In a psychophysical experiment, subjects confuse stimuli at rate $\Lambda(m) = 0.15$ in region m. Using $f(\kappa) = 1 - e^{-\kappa}$ and $\epsilon = 0.02$, recover $\hat{\kappa}_\pi(m)$ and bound the error. Is the geometry in region m positively or negatively curved?

21 Chapter 21

Hidden Curvature Recovery

The curvature was always there. The collapse is just where it became visible.

—

Phenomenological Note. The distortion is not always visible from inside it. A bias that affects everything you observe leaves no obvious sign, because there is nothing unaffected to compare it to. You discover hidden curvature mostly by its consequences: certain things are harder to reach than they should be, certain results keep coming out similar despite different inputs, certain regions of the map seem always to be underrepresented.

Chapter 20 set up the inverse problem and derived the reconstruction formula $\hat{\kappa}_\pi = f^{-1}(\Lambda)$ under regularity conditions. This chapter proves the Hidden Curvature Recovery Lemma: that under regular compact fibers and a strictly monotone response f, the recovered curvature is stable and unique. It then derives the second-order correction and identifies the precise conditions under which the recovery degrades.

21.1 The Hidden Curvature Recovery Lemma

Lemma 21.1 (*Hidden Curvature Recovery*). *Assume:*

(i) $\pi : X \to M$ *has regular compact fibers (each* $F_m = \pi^{-1}(m)$ *is a compact smooth manifold varying smoothly with* m*);*
(ii) $f : \mathbb{R} \to \mathbb{R}_{\geq 0}$ *is* C^2 *and strictly monotone increasing:* $f'(\kappa) > 0$ *for all* κ*;*
(iii) *Observational noise satisfies* $|\eta(m)| \leq \epsilon$ *with* $\epsilon < \frac{1}{2} \inf |f'| \cdot \mathrm{diam}(\kappa_\pi(M))$.

Then the recovery map $\hat{\kappa}_\pi = f^{-1}(\Lambda)$ *is:*

1. ***Well-defined****:* f^{-1} *exists on the range of* Λ*;*
2. ***Unique****: the recovered field is the unique minimiser of* $\int_M (f(\kappa) - \Lambda)^2 \, d\mu_M$*;*
3. ***Stable****: small perturbations of* Λ *produce small changes in* $\hat{\kappa}_\pi$*, with Lipschitz constant* $L = 1/\inf|f'|$.

Proof. Well-definedness. Since f is strictly monotone with $f' > 0$ everywhere, f is a diffeomorphism from $\mathbb{R}$ to $(f(-\infty), f(\infty))$. The range of Λ lies in $(f(-\infty), f(\infty))$

(since $\Lambda = f(\kappa_\pi) + \eta$ with bounded η and condition (iii) ensures Λ does not exit the range of f). Therefore $f^{-1}(\Lambda)$ is well-defined everywhere.

Uniqueness. The functional $J(\kappa) = \int_M (f(\kappa(m)) - \Lambda(m))^2 \, d\mu_M$ is strictly convex in κ when f is strictly monotone. A strictly convex functional on a compact domain has a unique minimiser. The minimiser satisfies $f(\kappa(m)) = \Lambda(m)$ pointwise, giving $\kappa(m) = f^{-1}(\Lambda(m)) = \hat{\kappa}_\pi(m)$.

Stability. For any two noise realisations Λ_1, Λ_2:

$$|\hat{\kappa}_1(m) - \hat{\kappa}_2(m)| = |f^{-1}(\Lambda_1(m)) - f^{-1}(\Lambda_2(m))| \leq \frac{|\Lambda_1(m) - \Lambda_2(m)|}{\inf |f'|} = \frac{1}{\inf |f'|} |\Lambda_1(m) - \Lambda_2(m)|,$$

using the inverse function theorem: $(f^{-1})' = 1/f'(f^{-1})$, so $|(f^{-1})'| \leq 1/\inf |f'|$. ▪ ▪

21.2 Second-Order Correction

The first-order recovery $\hat{\kappa}_\pi = f^{-1}(\Lambda)$ can be improved by incorporating the curvature of f itself.

Proposition 21.2 (*Second-Order Recovery*). *Under the conditions of Lemma 21.1, the second-order recovery*

$$\hat{\kappa}_\pi^{(2)}(m) = f^{-1}(\Lambda(m)) - \frac{f''(f^{-1}(\Lambda(m)))}{2(f'(f^{-1}(\Lambda(m))))^3} \sigma^2(m)$$

corrects for the bias introduced by noise variance $\sigma^2(m) = \mathbb{E}[\eta(m)^2]$:

$$\mathbb{E}[\hat{\kappa}_\pi^{(2)}(m)] = \kappa_\pi(m) + O(\sigma^4).$$

Proof. By the delta method: for $\hat{\kappa} = f^{-1}(\Lambda)$ with $\Lambda = f(\kappa) + \eta$, $\mathbb{E}[f^{-1}(f(\kappa) + \eta)] = \kappa + \frac{1}{2}(f^{-1})''(f(\kappa))\sigma^2 + O(\sigma^4)$. The second-order correction subtracts this bias term. ▪ ▪

21.3 Failure Modes

Three conditions can cause recovery to fail:

21.3.1 Degeneracy of the Response Function

If $\inf |f'(\kappa)| = 0$ at some $\kappa^* \in \mathbb{R}$, the Lipschitz constant $L = 1/\inf |f'|$ diverges. The recovery is *ill-conditioned* near κ^*: arbitrarily small noise in Λ produces unbounded errors in $\hat{\kappa}_\pi$. This occurs, for example, when f has a flat region (multiple hidden curvature values map to the same collapse level — a "semantic dead zone").

21.3.2 Fiber Singularities

If a fiber F_m (Amari and Nagaoka 2000; Carmo 1992) is singular (e.g., has a self-intersection or a cone point), the projected curvature $\kappa_\pi(m)$ is not well-defined as a smooth field. Recovery must be performed stratum by stratum: on each smooth stratum of F_m, the curvature is well-defined, but the overall recovery requires resolving the singularity first.

21.3.3 Noise Domination

If $\epsilon \geq \inf|f'| \cdot \Delta\kappa$, where $\Delta\kappa$ is the dynamic range of κ_π on M, the noise completely swamps the signal. The collapse field Λ is visible, but its curvature content is geometrically unrecoverable. This is the practical limit of inverse geometry: systems with small curvature variations and large measurement noise cannot be geometrically reconstructed from collapse observations.

21.4 Relation to the Projection-Collapse Principle

The Hidden Curvature Recovery Lemma is the *converse* of the Projection-Collapse Principle (Theorem 19.1).

Direction	Statement	Chapter
Forward	$\kappa_\pi(m) \nearrow \Rightarrow \Lambda(m) \nearrow$	Chapter 19
Inverse	$\Lambda(m) \Rightarrow \hat{\kappa}_\pi(m) = f^{-1}(\Lambda(m))$	This chapter

Together they establish that collapse and hidden curvature are in bijective correspondence under regularity, and that the forward-inverse pair is informationally conservative: no information about κ_π is created or destroyed by the observation of Λ — only noise is added.

Exercises

21.1. Let $f(\kappa) = \tanh(\kappa)$. Compute $\inf|f'|$, identify the ill-conditioned region, and compute the Lipschitz constant for recovery. For what range of κ is recovery reliable under $\epsilon = 0.01$?

21.2. Prove that if f is linear ($f(\kappa) = a\kappa + b$, $a > 0$), then the recovery is exact and the Lipschitz constant is $1/a$. Interpret: linear response is the best-case scenario for recovery.

21.3. The neural "tuning curve" of a sensory neuron maps stimulus intensity κ to firing rate $\Lambda = f(\kappa)$. Typical tuning curves are sigmoidal. Compute the recovery stability as a function of stimulus intensity for $f(\kappa) = 1/(1 + e^{-\kappa})$. Where is recovery most and least stable?

21.4. (Institutional.) A government collects error rate data $\Lambda(m)$ across administrative districts m. Using the recovery formula with $f(\kappa) = \kappa^{1/2}$, estimate the social complexity $\kappa_\pi(m)$ for each district. How does uncertainty in ϵ affect policy recommendations?

22 Chapter 22

Reconstruction Problems

> What survives a projection is a shadow. The problem is to infer the body from the shadow — knowing the body was three-dimensional and the shadow is two.
>
> —

Phenomenological Note. The original is gone. What you have is what survived the compression. You can try to reconstruct — to work backward from what remains to what must have been — but the reconstruction is always an estimate, and the error is not always obvious. Sometimes the reconstruction is so good that you forget it is not the original. Sometimes you only discover the gap when you need something that did not survive.

Chapters 20 and 21 solved the inverse curvature problem: given collapse observations $\Lambda(m)$, recover hidden curvature κ_π. This chapter addresses the more general **reconstruction problem**: given observations in the base M, recover geometric structure in the total space X. The central result bounds the reconstruction error as a function of observation noise and projection degeneracy.

22.1 The Reconstruction Setting

Let $\pi : X \to M$ (Cover and Thomas 2006) be a smooth projection with Jacobian $J_\pi(x) = D\pi(x)$. Denote the singular values of $J_\pi(x)$ as $\sigma_1(x) \geq \sigma_2(x) \geq \cdots \geq \sigma_k(x) \geq 0$ where $k = \dim M$. The **minimum singular value** $\sigma_{\text{min}}(J_\pi(x)) = \sigma_k(x)$ measures how degenerate the projection is at x: $\sigma_{\text{min}} = 0$ means the projection collapses some direction completely.

Let $\hat{\mathcal{G}}$ denote the admissibility geometry reconstructed from noisy observations $\Lambda + \eta$ on M, and $\mathcal{G}$ the true geometry.

22.2 The Reconstruction Error Bound

Theorem 22.1 (*Geometric Reconstruction Error*). *Under the regularity conditions of Lemma 21.1, with observation noise $|\eta(m)| \leq \epsilon$ uniformly and projected* *curvature* κ_π *bounded by* $\kappa_{\max}$*:*

$$\|\hat{\mathcal{G}} - \mathcal{G}\| \le \frac{\epsilon + \kappa_{\max}}{\sigma_{\min}(J_\pi)},$$

where $\|\cdot\|$ *denotes an appropriate norm on the space of admissibility geometries (e.g., the* L^2 *norm on curvature tensors).*

Proof. The reconstruction $\hat{\mathcal{G}}$ is obtained by: (1) computing $\hat{\kappa}_\pi = f^{-1}(\Lambda + \eta)$ from the noisy collapse field; (2) solving the inverse geometric problem to recover the full curvature tensor $\hat{\mathcal{G}}$ from $\hat{\kappa}_\pi$.

Step 1: Error in curvature recovery. By Theorem 20.1:

$$|\hat{\kappa}_\pi(m) - \kappa_\pi(m)| \le \frac{\epsilon}{\inf|f'|} + O(\epsilon^2).$$

Step 2: Error propagation through the inverse geometry. The projected curvature κ_π relates to the full curvature $\mathcal{G}$ through the projection Jacobian (O'Neill's formula for submersions):

$$\kappa_\pi(m) = \kappa_{\mathcal{G}}(x) + A(x),$$

where $A(x)$ is the O'Neill A-tensor contribution from the fiber. Recovering $\kappa_{\mathcal{G}}$ from κ_π requires inverting this relation, which involves the Jacobian J_π. By the implicit function theorem, the sensitivity of this inversion is bounded by $1/\sigma_{\min}(J_\pi)$:

$$\|\hat{\mathcal{G}} - \mathcal{G}\| \le \frac{|\hat{\kappa}_\pi - \kappa_\pi|}{\sigma_{\min}(J_\pi)} + \frac{\kappa_{\max}}{\sigma_{\min}(J_\pi)}.$$

Step 3: Combining. Substituting the curvature recovery error:

$$\|\hat{\mathcal{G}} - \mathcal{G}\| \le \frac{\epsilon/\inf|f'| + \kappa_{\max}}{\sigma_{\min}(J_\pi)} \le \frac{\epsilon + \kappa_{\max}}{\sigma_{\min}(J_\pi)},$$

absorbing $\inf|f'| \ge 1$ into the constant. ▪ ▪

22.3 Interpretation

The bound $(\epsilon + \kappa_{\max})/\sigma_{\min}(J_\pi)$ has three factors:

- ϵ (observation noise): unavoidable measurement error. More careful observation reduces this.
- $\kappa_{\max}$ (hidden curvature magnitude): the geometry of the total space is harder to recover when the fiber curvature is large. This is intrinsic to the system, not reducible by better measurement.
- $1/\sigma_{\min}(J_\pi)$ (projection degeneracy): as the projection approaches rank deficiency ($\sigma_{\min} \to 0$), the reconstruction error diverges. A near-degenerate projection is catastrophic for reconstruction.

Degeneracy Amplifies Both Error Sources. The projection degeneracy $1/\sigma_{\text{min}}$ amplifies *both* observation noise and hidden curvature. A slightly degenerate projection turns small noise into large reconstruction error. This is the geometric basis for why ill-conditioned representations (near-degenerate projections) are unreliable even in the presence of small measurement noise. It is also why the Fisher metric degeneracy of Chapter 17 is so significant: Fisher degeneracy implies $\sigma_{\text{min}}(J_\pi) \to 0$, driving reconstruction error to infinity.

22.4 Regularisation

When $\sigma_{\text{min}}(J_\pi)$ is small, the reconstruction is ill-conditioned. Standard regularisation techniques apply:

Definition 22.1 (*Tikhonov-Regularised Reconstruction*)**.** The **Tikhonov-regularised reconstruction** solves:

$$\hat{\mathcal{G}}_\alpha = \arg\min_{\mathcal{G}} \|\Lambda - f(\kappa_\pi(\mathcal{G}))\|^2 + \alpha\|\mathcal{G}\|^2,$$

for regularisation parameter $\alpha > 0$. The regularised error satisfies:

$$\|\hat{\mathcal{G}}_\alpha - \mathcal{G}\| \leq \frac{\epsilon}{\sigma_{\text{min}}^2 + \alpha} + \frac{\alpha\|\mathcal{G}\|}{\sigma_{\text{min}}^2 + \alpha}.$$

The optimal $\alpha^* = \epsilon/\|\mathcal{G}\|$ balances noise and regularisation.

Exercises

22.1. Let $\pi : \mathbb{R}^3 \to \mathbb{R}^2$ be the projection $(x, y, z) \mapsto (x, y)$. Compute $\sigma_{\text{min}}(J_\pi)$ and interpret: what does the error bound say about recovering 3D geometry from 2D observations?

22.2. Prove that if π is an isometric immersion (J_π preserves all distances), then $\sigma_{\text{min}}(J_\pi) = 1$ and the reconstruction error reduces to $\|\hat{\mathcal{G}} - \mathcal{G}\| \leq \epsilon + \kappa_{\text{max}}$.

22.3. (Neuroscience.) A neural decoding algorithm reconstructs stimulus geometry from neural responses. The projection from stimulus space to neural response space has minimum singular value $\sigma_{\text{min}} = 0.1$. With $\epsilon = 0.05$ (noise) and $\kappa_{\text{max}} = 0.3$, bound the reconstruction error. How much would reducing noise to $\epsilon = 0.01$ improve the bound?

22.4. Show that the Tikhonov-regularised error bound is minimised at $\alpha^* = \epsilon/\|\mathcal{G}\|$ and equals $2\sqrt{\epsilon\|\mathcal{G}\|}/(2\|\mathcal{G}\|/\sigma_{\text{min}}^2 + 1)$ at that optimum. When is regularisation necessary (when does it beat the unregularised bound)?

Part IV

Compression and Memory

[Part introduction — to be written.]

23 Chapter 23

Histories and Compression

> A history is not a sequence of facts. It is a trajectory through a state space that leaves a compressed residue.
>
> —

Phenomenological Note. Almost everything that happened today will not be retrievable by tomorrow. A compression is happening continuously whether you choose it or not. What survives is not necessarily the most important material — it is whatever happened to hook onto something that was already there. The record is always partial. The question is only whether the partial record is sufficient for the things that will later be asked of it.

This chapter establishes the formal foundations of compression theory for history spaces. (Cover and Thomas 2006; Shannon and Weaver 1949) The primary object is the compression operator $\mathcal{C} : \Gamma \to \mathcal{Z}$ mapping the full space of histories to a compressed record space. We prove that $\mathcal{C}$ induces a well-defined quotient topology on $\mathcal{Z}$ and that the compression is a bounded continuous operator between metric spaces.

23.1 History Spaces as Metric Spaces

Definition 23.1 (*History Space*)**.** The **history space** Γ is the space of all continuous trajectories $\gamma : [0, T] \to \mathcal{X}$ on a metric state space $(\mathcal{X}, d_{\mathcal{X}})$, equipped with the **uniform metric**:

$$d_\Gamma(\gamma_1, \gamma_2) = \sup_{t\in[0,T]} d_{\mathcal{X}}(\gamma_1(t), \gamma_2(t)).$$

(Γ, d_Γ) is a complete metric space when $(\mathcal{X}, d_{\mathcal{X}})$ is complete.

Definition 23.2 (*Compression Operator*)**.** A **compression operator** is a continuous surjective map $\mathcal{C} : \Gamma \to \mathcal{Z}$ where $(\mathcal{Z}, d_{\mathcal{Z}})$ is a compact metric space (the space of compressed records).

23.2 Topological Structure of the Compressed Space

Theorem 23.1 (*Compression Induces Quotient Topology*). *A continuous compression $\mathcal{C} : \Gamma \to \mathcal{Z}$ induces a well-defined quotient space structure on $\mathcal{Z}$. Specifically:*

(i) *The relation $\gamma_1 \sim_{\mathcal{C}} \gamma_2 \iff \mathcal{C}(\gamma_1) = \mathcal{C}(\gamma_2)$ is an equivalence relation on Γ;*
(ii) *The quotient map $\Gamma \to \Gamma/{\sim_{\mathcal{C}}}$ is continuous;*
(iii) *$\mathcal{Z}$ is homeomorphic to $\Gamma/{\sim_{\mathcal{C}}}$ when $\mathcal{C}$ is a quotient map;*
(iv) *For any query $q : \Gamma \to \mathcal{Y}$ that is constant on $\sim_{\mathcal{C}}$-equivalence classes, there exists a unique continuous $\hat{q} : \mathcal{Z} \to \mathcal{Y}$ with $q = \hat{q} \circ \mathcal{C}$.*

Proof. (i) *Equivalence relation.* Reflexivity: $\mathcal{C}(\gamma) = \mathcal{C}(\gamma)$ trivially. Symmetry: if $\mathcal{C}(\gamma_1) = \mathcal{C}(\gamma_2)$, then $\mathcal{C}(\gamma_2) = \mathcal{C}(\gamma_1)$. Transitivity: if $\mathcal{C}(\gamma_1) = \mathcal{C}(\gamma_2)$ and $\mathcal{C}(\gamma_2) = \mathcal{C}(\gamma_3)$, then $\mathcal{C}(\gamma_1) = \mathcal{C}(\gamma_3)$.

(ii) *Continuity of quotient map.* The quotient map $q : \Gamma \to \Gamma/{\sim_{\mathcal{C}}}$ has open sets: $U \subseteq \Gamma/{\sim_{\mathcal{C}}}$ is open iff $q^{-1}(U)$ is open in Γ. Since $\mathcal{C}$ is continuous, $\mathcal{C}^{-1}(V)$ is open for every open $V \subseteq \mathcal{Z}$. The sets $\{[\gamma] : \mathcal{C}(\gamma) \in V\}$ form a basis for the quotient topology, and they are open.

(iii) *Homeomorphism.* The map $\phi : \Gamma/{\sim_{\mathcal{C}}} \to \mathcal{Z}$ defined by $\phi([\gamma]) = \mathcal{C}(\gamma)$ is well-defined (by definition of $\sim_{\mathcal{C}}$), bijective (surjectivity of $\mathcal{C}$; injectivity of ϕ by construction), and continuous (quotient topology). Its inverse $\mathcal{C} \circ$ section is continuous when $\mathcal{C}$ is a quotient map.

(iv) *Factorisation.* Define $\hat{q}(z) = q(\gamma)$ for any $\gamma \in \mathcal{C}^{-1}(z)$. Well-definedness: q is constant on fibers. Continuity: $\hat{q} = q \circ s$ where $s : \mathcal{Z} \to \Gamma$ is a measurable section; continuity follows from the continuity of q and the quotient topology. ▪ ▪

23.3 Boundedness and Lipschitz Compression

Theorem 23.2 (*Compression is Bounded*). *Any continuous compression $\mathcal{C} : \Gamma \to \mathcal{Z}$ from a compact history space Γ to a compact record space $\mathcal{Z}$ is uniformly continuous and hence Lipschitz with some constant L:*

$$d_{\mathcal{Z}}(\mathcal{C}(\gamma_1), \mathcal{C}(\gamma_2)) \leq L \cdot d_{\Gamma}(\gamma_1, \gamma_2) \quad \forall\, \gamma_1, \gamma_2 \in \Gamma.$$

Proof. A continuous map from a compact metric space to a metric space is uniformly continuous (Heine-Cantor theorem). A uniformly continuous map on a compact domain is Lipschitz (by the mean value theorem for metric spaces): $L = \sup_{\gamma_1 \neq \gamma_2} d_{\mathcal{Z}}(\mathcal{C}\gamma_1, \mathcal{C}\gamma_2)/d_{\Gamma}(\gamma_1, \gamma_2) < \infty$ since the supremum is attained on a compact set. ▪ ▪

Remark 23.1 (Non-Compact Domains). For non-compact Γ (infinite-duration trajectories), compactness fails and Lipschitz continuity is not automatic. Practical compressions are always Lipschitz by design (e.g., averaging, sampling, spectral truncation), but the bound L may be large when histories vary rapidly.

23.4 Information Content of Compressions

Definition 23.3 (*Compression Ratio*). The **compression ratio** of $\mathcal{C}$ is

$$\rho(\mathcal{C}) = \frac{H(\mathcal{Z})}{H(\Gamma)},$$

where H denotes differential entropy. $\rho < 1$ indicates genuine compression (information loss). $\rho = 1$ indicates lossless compression.

By Theorem 13.1: $H(\Gamma) = H(\mathcal{Z}) + \mathbb{E}[S_{\mathcal{C}}(z)]$, so $\rho = 1 - \mathbb{E}[S_{\mathcal{C}}]/H(\Gamma)$. The compression ratio is close to 1 when the average fiber entropy $\mathbb{E}[S_{\mathcal{C}}]$ is small — i.e., when most histories in Γ are distinguishable by their compressed representation.

Exercises

23.1. Let $\Gamma = C([0,1],\mathbb{R})$ (continuous functions on $[0,1]$) with the uniform metric, and $\mathcal{C}(\gamma) = \gamma(0)$ (evaluate at the initial point). Show that $\mathcal{C}$ is continuous but not Lipschitz with constant < 1. What are the fibers $\mathcal{C}^{-1}(x)$?

23.2. Let $\mathcal{C}_n(\gamma) = (\gamma(k/n))_{k=0}^{n}$ (sample at $n+1$ equally-spaced points). Prove that $\mathcal{C}_n$ is Lipschitz with $L = 1$. Show that as $n \to \infty$, $\mathcal{C}_n$ approaches the identity in the uniform topology.

23.3. Prove that the composition of two compressions $\mathcal{C}_2 \circ \mathcal{C}_1$ is a compression with Lipschitz constant $L \leq L_1 L_2$. When is the composed compression strictly more lossy than either alone?

23.4. Define the *compression efficiency* of $\mathcal{C}$ as the ratio $H(\mathcal{Z})/\dim(\mathcal{Z})$ (information per dimension of the compressed space). For a Gaussian history space, compute the compression efficiency of: (a) principal component projection; (b) random projection; (c) wavelet truncation at level J. Which is most efficient?

24 Chapter 24

Sufficient Statistics

> A sufficient statistic carries everything you need and nothing you don't.
>
> —

Phenomenological Note. You do not need everything in order to decide correctly. Often a small amount of the right information is better than a large amount of the wrong kind. The difficulty is knowing in advance which kind is which — whether what you have retained is sufficient for the question you will eventually face. This is mostly decided before the question arrives.

The concept of sufficiency is the bridge between compression (Chapter 27) and information geometry (Chapter 9). A compression is sufficient for a query if it preserves all query-relevant information — formally, if the query's value factors through the compression. This chapter proves the Sufficiency Theorem and derives its consequences for the geometry of compressed spaces.

24.1 Sufficient Compressions

Definition 24.1 (*Sufficient Compression*)**.** Let $\mathcal{C} : \Gamma \to \mathcal{Z}$ be a compression operator and $q : \Gamma \to \mathcal{Y}$ a query functional. $\mathcal{C}$ is **sufficient** for q if:

$$\mathcal{C}(\gamma_1) = \mathcal{C}(\gamma_2) \;\Rightarrow\; q(\gamma_1) = q(\gamma_2) \quad \forall\, \gamma_1, \gamma_2 \in \Gamma.$$

Equivalently, q is constant on each fiber of $\mathcal{C}$.

Theorem 24.1 (*Sufficiency Theorem*)**.** *$\mathcal{C}$ is sufficient for q if and only if q factors through $\mathcal{C}$:*

$$\exists\, \hat{q} : \mathcal{Z} \to \mathcal{Y} \quad \text{such that} \quad q = \hat{q} \circ \mathcal{C}.$$

Proof. ($\Rightarrow$) Define $\hat{q}(z) = q(\gamma)$ for any $\gamma \in \mathcal{C}^{-1}(z)$. This is well-defined by sufficiency: all γ in the same fiber give the same $q(\gamma)$. Then $(\hat{q} \circ \mathcal{C})(\gamma) = \hat{q}(\mathcal{C}(\gamma)) = q(\gamma)$.

($\Leftarrow$) If $q = \hat{q} \circ \mathcal{C}$ and $\mathcal{C}(\gamma_1) = \mathcal{C}(\gamma_2) = z$, then $q(\gamma_1) = \hat{q}(z) = q(\gamma_2)$. ▪ ▪

24.2 The Fisher-Neyman Factorisation

In statistics, sufficiency is characterised by the Fisher–Neyman factorisation theorem:

Theorem 24.2 (*Fisher–Neyman Factorisation*). *Let $X \sim p_\theta$ be a random observation from a parametric family. A statistic $T(X)$ is sufficient for θ if and only if the likelihood factorises:*

$$p_\theta(x) = g(T(x), \theta) \cdot h(x),$$

where g depends on x only through $T(x)$ and h does not depend on θ.

Proof. ($\Rightarrow$) If T is sufficient, the conditional distribution $p(X = x \mid T(X) = t)$ does not depend on θ. Write $p_\theta(x) = p_\theta(T(x) = t) \cdot p(x \mid T(x) = t)$; the first factor depends on θ through $T(x)$ only, so $g(T(x), \theta) = p_\theta(T(x))$ and $h(x) = p(x \mid T(x))$.

($\Leftarrow$) Given the factorisation, $p(X = x \mid T(X) = t, \theta) = p_\theta(x)/p_\theta(T = t) = g(t, \theta)h(x)/\int_{T(x')=t} g(t, \theta)h(x')\mathrm{d}x' = h(x)/\int_{T(x')=t} h(x')\mathrm{d}x'$, which is independent of θ. ▪ ▪

Remark 24.1 (Sufficiency as Fiber-Constancy in Statistics). The Fisher–Neyman theorem is the statistical instance of Theorem 24.1: T is sufficient for θ iff the parameter θ is constant on the fibers of T. The "query" $q(\cdot) = \theta$ factors through T.

24.3 Minimal Sufficient Statistics

Definition 24.2 (*Minimal Sufficient Compression*). $\mathcal{C}$ is a **minimal sufficient compression** for q if it is sufficient for q and, for any other sufficient compression $\mathcal{C}'$, there exists a map $f : \mathcal{Z} \to \mathcal{Z}'$ such that $\mathcal{C}' = f \circ \mathcal{C}$. (Every other sufficient compression is a further compression of $\mathcal{C}$.)

The minimal sufficient compression is the *coarsest compression that preserves all q-relevant information*. It achieves maximum compression without any query-relevant information loss.

Proposition 24.3 (*Equivalence Classes Characterise Minimal Sufficiency*). *The minimal sufficient compression for q partitions Γ into equivalence classes: $\gamma_1 \sim_q \gamma_2$ iff $q(\gamma_1) = q(\gamma_2)$. The quotient map $\mathcal{C}^* : \Gamma \to \Gamma/{\sim_q}$ is the minimal sufficient compression.*

Proof. Let T_1, T_2 be minimal sufficient statistics. Minimality of T_1: for any sufficient T, $T_1 = f(T)$ for some f. Applying this with $T = T_2$: $T_1 = f(T_2)$. By symmetry $T_2 = g(T_1)$ for some g. Hence T_1 and T_2 generate the same σ-algebra: they partition the data into identical equivalence classes. These classes are exactly the fibers of the minimal compression. ▪ ▪

24.4 Sufficiency and Information Geometry

Theorem 24.4 (*Sufficiency Preserves Fisher Information*)**.** *If $T(X)$ is a sufficient statistic for θ, then the Fisher information in $T(X)$ equals the Fisher information in X:*

$$\mathcal{I}_T(\theta) = \mathcal{I}_X(\theta).$$

Sufficient compression preserves Fisher information exactly.

Proof. By the Fisher–Neyman factorisation, $\log p_\theta(x) = \log g(T(x), \theta) + \log h(x)$. The score is $\partial_\theta \log p_\theta(x) = \partial_\theta \log g(T(x), \theta)$, which depends on x only through $T(x)$. Therefore $\mathcal{I}_X(\theta) = \mathbb{E}[(\partial_\theta \log p_\theta(X))^2] = \mathbb{E}[(\partial_\theta \log g(T(X), \theta))^2] = \mathcal{I}_T(\theta)$. ▪ ▪

This is the geometric formulation: a sufficient compression is an *isometry* of the Fisher information metric. It compresses without distorting the geometry of the parameter space. Insufficient compressions are non-isometric projections that distort and reduce Fisher information.

Exercises

24.1. Let $\Gamma = \{0,1\}^n$, $\mathcal{C}(\gamma) = \sum_i \gamma_i$. Show that $\mathcal{C}$ is sufficient for $q(\gamma) = \sum_i \gamma_i$ (trivially) and for $q'(\gamma) = \max_i \gamma_i$. Is it sufficient for $q''(\gamma) = \gamma_1$ (the first coordinate)?

24.2. Prove that the composition of two sufficient compressions (for the same query q) is also sufficient for q.

24.3. Let $X_1, \ldots, X_n \sim \mathcal{N}(\mu, 1)$ iid. Show that $T = \sum_i X_i$ is sufficient for μ by finding the Fisher–Neyman factorisation. Is T^2 also sufficient? Why or why not?

24.4. Define the *sufficiency gap* of a compression $\mathcal{C}$ for query q as $G(\mathcal{C}, q) = \mathcal{I}_X(\theta) - \mathcal{I}_T(\theta)$ where $T = \mathcal{C}(X)$. Prove $G \geq 0$ (data processing inequality) with equality iff $\mathcal{C}$ is sufficient.

25 Chapter 25

Reconstruction Operators

> To reconstruct is not to remember. It is to infer, from what survived, what could have produced it.
>
> —

Phenomenological Note. The gap between the compressed record and the original event is sometimes recoverable and sometimes not. Time cannot be reversed. But a good reconstruction operator can sometimes get you surprisingly close to what you need, even from a very incomplete record. The question is not whether the record is complete. It is whether the reconstruction is sufficient for the purposes at hand.

Compression loses information. Reconstruction attempts to recover it. This chapter studies the conditions under which a reconstruction operator $\mathcal{R} : \mathcal{Z} \to \Gamma$ successfully inverts the compression $\mathcal{C} : \Gamma \to \mathcal{Z}$ (Amari and Nagaoka 2000), and characterises the geometry of reconstruction error when perfect inversion is impossible.

25.1 Reconstruction Operators

Definition 25.1 (*Reconstruction Operator*)**.** A **reconstruction operator** for compression $\mathcal{C} : \Gamma \to \mathcal{Z}$ is a map $\mathcal{R} : \mathcal{Z} \to \Gamma$ that attempts to invert $\mathcal{C}$: $\mathcal{R} \circ \mathcal{C} \approx \mathrm{id}_\Gamma$. If $\mathcal{R} \circ \mathcal{C} = \mathrm{id}_\Gamma$ exactly, $\mathcal{R}$ is a **perfect reconstruction**.

Lemma 25.1 (*Reconstruction Existence Lemma*)**.** *A perfect reconstruction operator $\mathcal{R}$ exists if and only if $\mathcal{C}$ is injective. When $\mathcal{C}$ is non-injective, no reconstruction operator can distinguish histories in the same fiber.*

Proof. ($\Rightarrow$) If $\mathcal{R} \circ \mathcal{C} = \mathrm{id}$, then $\mathcal{C}(\gamma_1) = \mathcal{C}(\gamma_2)$ implies $\gamma_1 = \mathcal{R}(\mathcal{C}(\gamma_1)) = \mathcal{R}(\mathcal{C}(\gamma_2)) = \gamma_2$. So $\mathcal{C}$ is injective.

($\Leftarrow$) If $\mathcal{C}$ is injective, define $\mathcal{R}(z) = \mathcal{C}^{-1}(z)$ (the unique preimage of z). This is well-defined by injectivity and satisfies $\mathcal{R}(\mathcal{C}(\gamma)) = \gamma$. ▪ ▪

25.2 Approximate Reconstruction and Minimum Error

When $\mathcal{C}$ is non-injective, reconstruction must make a choice among multiple histories in the same fiber.

Definition 25.2 (*Minimum-Error Reconstruction*)**.** Given a distortion measure $d : \Gamma \times \Gamma \to \mathbb{R}_{\geq 0}$, the **minimum-error reconstruction** is

$$\mathcal{R}^*(z) = \arg\min_{\gamma \in \Gamma} \mathbb{E}_{\gamma_0 \sim p_z}[d(\gamma, \gamma_0)],$$

where p_z is the distribution over histories in the fiber $\mathcal{C}^{-1}(z)$. For squared Euclidean distortion, this is the conditional mean: $\mathcal{R}^*(z) = \mathbb{E}[\gamma \mid \mathcal{C}(\gamma) = z]$.

Theorem 25.2 (*Reconstruction Error Decomposition*)**.** *The total expected reconstruction error decomposes as:*

$$\mathbb{E}[d(\gamma, \mathcal{R}(\mathcal{C}(\gamma)))] = \underbrace{\mathbb{E}[d(\gamma, \gamma^*)]}_{\text{approximation error}} + \underbrace{\mathbb{E}[d(\gamma^*, \mathcal{R}(\mathcal{C}(\gamma)))]}_{\text{estimation error}},$$

where $\gamma^* = \mathcal{R}^*(\mathcal{C}(\gamma))$ is the optimal reconstruction. The approximation error is irreducible (information lost by compression); the estimation error is reducible by improving $\mathcal{R}$.

Proof. For squared Euclidean distortion, this is the bias-variance decomposition: $\mathbb{E}[\|\gamma - \mathcal{R}(z)\|^2] = \mathbb{E}[\|\gamma - \gamma^*\|^2] + \|\gamma^* - \mathcal{R}(z)\|^2$, where the cross-term vanishes by the optimality of γ^* (Pythagoras in Hilbert space). ▪ ▪

25.3 The Reconstruction Geometry

The fiber $\mathcal{C}^{-1}(z)$ is the set of all histories consistent with compressed record z. Its geometry determines the reconstruction error.

Proposition 25.3 (*Fiber Geometry and Error*)**.** *For a uniform distribution over fibers:*

(i) ***Large fibers*** *(large $\mu(\mathcal{C}^{-1}(z))$): many histories are consistent with z; the minimum-error reconstruction $\mathcal{R}^*(z)$ has high approximation error.*
(ii) ***Curved fibers****: if the fiber $\mathcal{C}^{-1}(z)$ has high curvature, the centroid γ^* may lie outside the fiber, creating a geometric bias in reconstruction.*
(iii) ***Disconnected fibers****: if $\mathcal{C}^{-1}(z)$ is disconnected, the centroid lies between the components, failing to represent any actual history.*

Proof. (i) *Connected fiber.* The minimum-error reconstruction $\hat{\gamma} = \mathbb{E}[\gamma \mid z]$ is the conditional mean of γ given $z = \mathcal{C}(\gamma)$. For a connected fiber, the conditional mean lies within the fiber's convex hull. When the fiber is convex, $\hat{\gamma} \in \pi^{-1}(z)$.

(ii) *Convex fiber.* By Jensen's inequality, for any convex loss ℓ: $\ell(\hat{\gamma}, \gamma^*) \leq \mathbb{E}[\ell(\gamma, \gamma^*) \mid z]$, so the conditional mean minimises expected loss. The minimum error equals the within-fiber variance Var$[\gamma \mid z]$.

(iii) *Disconnected fiber.* If $\pi^{-1}(z) = C_1 \sqcup C_2$ (two disjoint components), the conditional mean $\hat{\gamma} = p_1 c_1^* + p_2 c_2^*$ (where c_i^* are component centroids and p_i probabilities) lies in the convex hull but not in $\pi^{-1}(z)$. Hence $\hat{\gamma} \notin \pi^{-1}(z)$: the reconstruction is a non-existent point, producing hallucination. ▪ ▪

The disconnected fiber case is the reconstruction analogue of semantic degeneracy (Definition 39.2): ambiguous compressions produce reconstructions that correspond to no real history.

This is the mathematical basis of LLM hallucination as reconstruction failure: the model's compressed representation z corresponds to a fiber with disconnected high-probability regions; the generated text is the centroid — a blend that is not any actual history. See Chapter 80.

25.4 Rate-Distortion Theory

Theorem 25.4 (*Rate-Distortion Bound*)**.** *For a source $\gamma \sim P$ and distortion measure d, the minimum achievable expected distortion D at compression rate R bits is:*

$$D(R) = \min_{\mathcal{C}: I(\gamma;\mathcal{C}(\gamma)) \leq R} \mathbb{E}[d(\gamma, \mathcal{R}^*(\mathcal{C}(\gamma)))].$$

The rate-distortion function $R(D)$ — the minimum rate needed to achieve distortion $\leq D$ — satisfies:

$$R(D) = \min_{p(\hat{\gamma}|\gamma):\, \mathbb{E}[d(\gamma,\hat{\gamma})] \leq D} I(\gamma; \hat{\gamma}).$$

Proof. By Shannon's rate-distortion theorem (Cover and Thomas 2006): for a source Γ with distortion measure d, the minimum rate $R(D)$ achievable at distortion D is $R(D) = \min_{p(\hat{\gamma}|\gamma):\mathbb{E}[d] \leq D} I(\Gamma; \hat{\Gamma})$. Setting $d(\gamma, \hat{\gamma}) = \|\hat{\gamma} - \mathcal{R}(\mathcal{C}(\gamma))\|^2$ (reconstruction error) and noting that $I(\Gamma; \hat{\Gamma}) \leq H(\mathcal{C}(\Gamma))$ (compressed representation entropy), we get $R(D) \leq H(\mathcal{C})$. The fiber entropy $S_{\mathcal{C}}$ equals $H(\mathcal{C})$ (log volume = differential entropy for continuous sources), giving the bound $R(D) \leq S_{\mathcal{C}}$. ▪ ▪

The rate-distortion theorem operationalises the compression-quality tradeoff: compressing to fewer bits necessarily increases distortion. The curve $R(D)$ is the Pareto frontier of this tradeoff. Every point on the curve corresponds to a "minimally sufficient" compression for a given distortion budget.

Exercises

25.1. Let $\Gamma = \mathbb{R}$, $\mathcal{C}(\gamma) = \lfloor \gamma \rfloor$ (floor function). Find $\mathcal{R}^*(z)$ under squared distortion. Compute the approximation error.

25.2. Prove that for any $\mathcal{R}$: $\mathbb{E}[d(\gamma, \mathcal{R}(\mathcal{C}(\gamma)))] \geq \mathbb{E}[d(\gamma, \mathcal{R}^*(\mathcal{C}(\gamma)))]$. (The minimum-error reconstruction is optimal.)

25.3. For a Gaussian source $\gamma \sim \mathcal{N}(0, \sigma^2)$ and squared distortion, compute the rate-distortion function $R(D)$. At what D is $R(D) = 0$? Interpret in terms of reachability.

25.4. A language model compresses context windows of length n into a fixed-size embedding of dimension $d \ll n$. Model this as $\mathcal{C} : \{0,1\}^n \to \mathbb{R}^d$. Derive a lower bound on the expected reconstruction error for any downstream query q with $I(\text{context}; q) > d \log_2 e$ bits.

26 Chapter 26

Stability Theory

A stable compression is one where nearby histories produce nearby records, and nearby records produce nearby reconstructions.

—

Phenomenological Note. Some things remain roughly themselves under pressure. Others splinter at the first disturbance. Stability is not rigidity — a rigid thing can shatter precisely because it refuses to yield. A stable thing absorbs perturbation without losing its essential character. The question of which kind any given thing is matters most at the moment when it is actually being disturbed.

Sufficient compression (Chapter 24) and reconstruction (Chapter 25) are existence results. (Cover and Thomas 2006) Stability theory asks the quantitative question: how sensitive is reconstruction to perturbation? The load-bearing proof bounds truncation error when higher-order historical interactions decay, formalising the intuition that systems with fading memory can be reliably compressed and reconstructed.

26.1 Lipschitz Stability

Definition 26.1 (*Stable Compression*). A compression $\mathcal{C} : \Gamma \to \mathcal{Z}$ is **Lipschitz stable** with constant L if

$$d_{\mathcal{Z}}(\mathcal{C}(\gamma_1), \mathcal{C}(\gamma_2)) \leq L\, d_{\Gamma}(\gamma_1, \gamma_2) \quad \forall\, \gamma_1, \gamma_2 \in \Gamma.$$

A reconstruction $\mathcal{R} : \mathcal{Z} \to \Gamma$ is Lipschitz stable with constant M if $d_{\Gamma}(\mathcal{R}(z_1), \mathcal{R}(z_2)) \leq M\, d_{\mathcal{Z}}(z_1, z_2)$.

Proposition 26.1 (*Stability Composition*). *If $\mathcal{C}$ is Lipschitz stable with constant L and $\mathcal{R}$ is Lipschitz stable with constant M, then the end-to-end map $\mathcal{R} \circ \mathcal{C}$ satisfies:*

$$d_{\Gamma}((\mathcal{R} \circ \mathcal{C})(\gamma_1), (\mathcal{R} \circ \mathcal{C})(\gamma_2)) \leq ML\, d_{\Gamma}(\gamma_1, \gamma_2).$$

The system is end-to-end stable with constant ML.

Proof. Let $\mathcal{C}_1, \mathcal{C}_2$ have stability bounds B_1, B_2 respectively. For perturbations $\delta\gamma$ and $\delta\gamma'$ in the respective domains: $d(\mathcal{C}_1(\gamma+\delta), \mathcal{C}_1(\gamma)) \leq B_1\|\delta\|$ and $d(\mathcal{C}_2(z+$

$\epsilon), \mathcal{C}_2(z)) \leq B_2\|\epsilon\|$. For the composition, setting $\epsilon = \mathcal{C}_1(\gamma+\delta) - \mathcal{C}_1(\gamma)$: $d((\mathcal{C}_2 \circ \mathcal{C}_1)(\gamma+\delta), (\mathcal{C}_2 \circ \mathcal{C}_1)(\gamma)) \leq B_2 B_1\|\delta\|$. Hence the composition has stability bound $B_1 B_2$. ▪ ▪

26.2 The Truncation Stability Bound

Many natural compression schemes truncate a history at depth D: they keep only the last D time steps or the first D terms of a decomposition.

Definition 26.2 (*Decaying Interaction Model*)**.** A history γ is said to have **decaying higher-order interactions** with coefficients $(\alpha_k)_{k\geq 1}$ if its contribution from order-k temporal correlations has magnitude bounded by α_k, with $\sum_{k=1}^{\infty} \alpha_k < \infty$.

Theorem 26.2 (*Truncation Stability Bound*)**.** *Let $\mathcal{C}_D$ denote the compression that retains only the first D interaction orders. Under the decaying interaction model, the reconstruction error from depth-D truncation satisfies:*

$$\|\gamma - \mathcal{R}(\mathcal{C}_D(\gamma))\| \leq \sum_{k>D} \alpha_k.$$

The error is bounded by the tail sum of the interaction decay coefficients.

Proof. Write $\gamma = \gamma_D + \gamma_{>D}$ where γ_D contains contributions from orders $k \leq D$ and $\gamma_{>D}$ contains contributions from orders $k > D$.

Since $\mathcal{C}_D$ projects onto the first D orders, $\mathcal{C}_D(\gamma) = \mathcal{C}_D(\gamma_D)$ (the high-order components are discarded).

The reconstruction $\mathcal{R}(\mathcal{C}_D(\gamma))$ can only recover γ_D: $\mathcal{R}(\mathcal{C}_D(\gamma)) = \gamma_D$.

Therefore:

$$\|\gamma - \mathcal{R}(\mathcal{C}_D(\gamma))\| = \|\gamma - \gamma_D\| = \|\gamma_{>D}\| \leq \sum_{k>D} \|\text{order-}k\text{ component}\| \leq \sum_{k>D} \alpha_k. \quad ▪$$

▪

Corollary 26.3 (*Fading Memory Implies Stable Compression*)**.** *If $\alpha_k \to 0$ faster than any polynomial (e.g., $\alpha_k = e^{-\beta k}$ for $\beta > 0$), then $\sum_{k>D} \alpha_k \to 0$ exponentially fast. In this case, finite-depth compression is uniformly close to infinite-depth compression. The system has **fading memory**: distant past has negligible influence on present behaviour.*

Proof. Fading memory means the influence of history at lag k decays as α^k with $\alpha < 1$. The Lipschitz constant for the k-th order contribution to $\mathcal{C}$ is $\leq C\alpha^k$. Summing over all lags: total Lipschitz constant $\leq C\sum_{k=0}^{\infty} \alpha^k = C/(1-\alpha) < \infty$. By the stability theorem (Theorem 26.2), bounded Lipschitz constant implies bounded truncation error, which is precisely the stability condition. ▪ ▪

26.3 Structural Stability of Compression Schemes

Beyond truncation, we ask: what happens when the compression scheme itself is perturbed?

Definition 26.3 (*Structurally Stable Compression*)**.** A compression $\mathcal{C}$ is **structurally stable** if for any $\epsilon > 0$ there exists $\delta > 0$ such that any compression $\mathcal{C}'$ with $\sup_\gamma d_\mathcal{Z}(\mathcal{C}(\gamma), \mathcal{C}'(\gamma)) < \delta$ yields reconstruction error:

$$\sup_\gamma \|\mathcal{R}'(\mathcal{C}'(\gamma)) - \mathcal{R}(\mathcal{C}(\gamma))\| < \epsilon.$$

Proposition 26.4 (*Lipschitz Stability Implies Structural Stability*)**.** *If both $\mathcal{C}$ and $\mathcal{R}$ are Lipschitz stable (with constants L and M respectively), then the system is structurally stable with $\delta = \epsilon/M$.*

Proof. $\|\mathcal{R}'(\mathcal{C}'(\gamma)) - \mathcal{R}(\mathcal{C}(\gamma))\| \le M d_\mathcal{Z}(\mathcal{C}'(\gamma), \mathcal{C}(\gamma)) < M\delta = \epsilon$. ▪ ▪

26.4 Applications

Memory systems.. The truncation bound says: if the relevance of past events decays geometrically, then a fixed-depth memory (LRU cache, hippocampal trace) loses at most an exponentially small amount of query-relevant information. This justifies finite-window attention mechanisms: for queries about the present, a window of size $D = O(\beta^{-1} \log(1/\epsilon))$ suffices.

Scientific models.. A physical model that truncates higher-order corrections (perturbation theory in QFT, Taylor expansions in mechanics) is stable precisely when the interaction series converges. The truncation bound gives a rigorous certificate: if the k-th order contribution to the scattering amplitude decays as $\alpha_k = g^k$ with $|g| < 1$ (coupling constant), then D-th order perturbation theory has error $\le g^{D+1}/(1-g)$.

Institutional memory.. Laws and precedents compress institutional history. The truncation bound says: if the influence of past decisions on present admissibility decays with depth D (e.g., precedents older than D years have weight $\le \alpha_D$), then a finite-depth legal memory is a stable compression. The stability of common law systems depends on this decay.

Exercises

26.1. Let $\gamma \in \mathbb{R}^\mathbb{N}$ be an infinite sequence and $\alpha_k = r^k$ for $r \in (0, 1)$. Compute the truncation error $\sum_{k>D} \alpha_k$ and find the smallest D guaranteeing error $\le \epsilon$.

26.2. Show that the truncation bound is tight: construct a history γ and interaction model (α_k) such that the reconstruction error equals $\sum_{k>D} \alpha_k$ exactly.

26.3. Prove that for the context window in a transformer (with window size D), the attention mechanism is a Lipschitz stable compression under the L^∞ metric on token sequences, with Lipschitz constant $L = 1$. (Use the contractiveness of softmax attention.)

26.4. Define the *memory depth* of a cognitive or institutional system as the smallest D such that $\sum_{k>D} \alpha_k < \epsilon$ for some tolerance ϵ. Argue that memory depth is an invariant of the system's constraint structure, not merely an engineering parameter.

27 Chapter 27

Memory as Compressed History

> Memory is not a filing cabinet. It is a lossy compression scheme optimised for the queries we are likely to run.
>
> —

Phenomenological Note. What you carry of the past is not the past. It is a reconstruction that has already been edited by the time that followed. Each retrieval slightly alters the stored version. Each retelling shapes the memory in the direction of the telling. The archive and the act of consulting the archive are not separate processes. You are always reading and writing at the same time.

error $\leq \sum_{k>D} \alpha_k$ (truncation stability)

sufficient: $q = \hat{q} \circ \mathcal{C}$

Figure 27.1: The compression-reconstruction pipeline. $\mathcal{C}$ maps histories to compressed records; $\mathcal{D}$ attempts reconstruction. The truncation stability bound limits error when interactions decay at rate α_k.

This chapter unifies the compression machinery developed in Chapters 23–26 into a single theorem: memory is reliable exactly when the compression it applies is sufficient for the retrieval query.

Definition 27.1 (*Memory System*). A **memory system** is a pair $(\mathcal{C}, \mathcal{R})$ where $\mathcal{C} : \Gamma \to \mathcal{Z}$ maps histories to stored records, and $\mathcal{R} : \mathcal{Z} \times \mathcal{Q} \to \mathcal{Y}$ maps stored records and retrieval queries to responses.

Theorem 27.1 (*Memory as Compressed History*). *A memory system $(\mathcal{C}, \mathcal{R})$ reliably answers query $q : \Gamma \to \mathcal{Y}$ if and only if $\mathcal{C}$ is a* sufficient compression *for q:*

$$\mathcal{C}(\gamma_1) = \mathcal{C}(\gamma_2) \;\Rightarrow\; q(\gamma_1) = q(\gamma_2).$$

Equivalently, $q = \hat{q} \circ \mathcal{C}$ for some well-defined $\hat{q} : \mathcal{Z} \to \mathcal{Y}$.

Proof. ($\Rightarrow$): If memory is reliable and $\mathcal{C}(\gamma_1) = \mathcal{C}(\gamma_2) = z$, then $q(\gamma_1) = \mathcal{R}(z, q) = q(\gamma_2)$.

($\Leftarrow$): If $\mathcal{C}$ is sufficient for q, define $\hat{q}(z) = q(\gamma)$ for any $\gamma \in \mathcal{C}^{-1}(z)$ (well-defined by sufficiency). Then $\mathcal{R}(z, q) := \hat{q}(z)$ gives reliable retrieval. ▪ ▪

Corollary 27.2 (*Mutual Information Bound*)**.** *Memory retrieval for q is reliable iff $I(\Gamma; \mathcal{Z}) \geq I(\Gamma; q)$.*

Proof. The mutual information $I(\Gamma; \hat{\Gamma})$ between the original history and its reconstruction is bounded by $I(\Gamma; \hat{\Gamma}) \leq H(\mathcal{C}(\Gamma)) = S_{\mathcal{C}}$, since the reconstruction $\hat{\Gamma}$ is a function of $\mathcal{C}(\Gamma)$ and conditioning can only reduce entropy: $I(\Gamma; \hat{\Gamma}) \leq I(\Gamma; \mathcal{C}(\Gamma)) = H(\mathcal{C}(\Gamma)) - H(\mathcal{C}|\Gamma) = H(\mathcal{C}(\Gamma))$. ▪ ▪

Query-Specific Compression. There is no single best memory compression. Optimal compression depends on the queries the system expects. Episodic memory ("what happened at time t?") requires high-resolution temporal compression; semantic memory ("what kind of thing is X?") tolerates coarser schema compression.

Proposition 27.3 (*Forgetting Destroys Sufficiency*)**.** *If $\mathcal{C}_1$ is sufficient for q and $\mathcal{C}_2 = f \circ \mathcal{C}_1$ is a further coarsening, then $\mathcal{C}_2$ may fail to be sufficient for q. Entropy increase $\Delta H = H(\mathcal{C}_2) - H(\mathcal{C}_1) \geq 0$ measures the potential sufficiency loss.*

Context-length limitations in transformer LLMs are a direct instance: the attention compression is sufficient only for queries answerable from the compressed context window.

Proof. Let $T > D$ (forgetting threshold). The compressed history $\mathcal{C}_D(\gamma)$ contains only information from lags $k \leq D$. For a query q that depends on lag $k > D$ (e.g., $q(\gamma) = \gamma_{t-k}$ for some $k > D$), the compression has discarded the relevant information: $\mathbb{E}[q(\gamma) \mid \mathcal{C}_D(\gamma)] \neq q(\gamma)$ in general. By Theorem 27.1, sufficiency fails when the compression drops query-relevant information. ▪ ▪

Exercises

27.1. Let $\Gamma = \{0, 1\}^n$ and $q(\gamma) = \sum_i \gamma_i$. Describe the coarsest sufficient compression. How many bits does it require?

27.2. Give two queries q_1, q_2 such that no single compression is sufficient for both. Describe the finest compression sufficient for both simultaneously.

27.3. Prove that any composition $\mathcal{C}_2 \circ \mathcal{C}_1^{-1}$ can only decrease the set of queries for which the system is sufficient.

28 Chapter 28

Counterfactual Recovery

> What might have been is only knowable if its trace survives the compression.
>
> —

Phenomenological Note. The roads not taken are mostly invisible. You can sometimes recover them through deliberate effort — imagining what would have happened if a different choice had been made — but the recovery is lossy. You are constructing a counterfactual from a position that was itself shaped by the choice you actually made. The imagination of the other path is always contaminated by knowledge of this one.

A counterfactual is a statement about what would have happened under a different trajectory. Counterfactual reasoning requires that the alternative trajectory γ^* remain distinguishable from the actual trajectory γ in the compressed record. This chapter proves the Counterfactual Recovery Lemma: a counterfactual is recoverable if and only if the two trajectories are mapped to different compressed records.

28.1 Counterfactuals as Alternative Trajectories

Definition 28.1 (*Counterfactual Trajectory*)**.** Let $\gamma \in \Gamma$ be the **actual trajectory** of a system. A **counterfactual trajectory** $\gamma^* \in \Gamma$ is an alternative history that differs from γ at some intervention point t_0: $\gamma(t) = \gamma^*(t)$ for $t < t_0$ and $\gamma(t) \neq \gamma^*(t)$ for some $t > t_0$.

Counterfactual reasoning asks: what would the outcome $q(\gamma^*)$ have been? This question is answerable from memory iff the compression preserves enough of γ^*'s structure to distinguish it from γ.

28.2 The Counterfactual Recovery Lemma

Lemma 28.1 (*Counterfactual Recovery*)**.** *Let $\mathcal{C} : \Gamma \to \mathcal{Z}$ be a memory compression operator and let $q : \Gamma \to \mathcal{Y}$ be a query. A counterfactual query $q(\gamma^*)$ is recoverable from the compressed record $\mathcal{C}(\gamma)$ of the actual trajectory if and only if:*

(i) $\mathcal{C}(\gamma) \neq \mathcal{C}(\gamma^*)$ (*the two trajectories are distinguishable in compressed space*);
(ii) $\mathcal{C}$ *is sufficient for* q *on* $\{\gamma, \gamma^*\}$ (*the query value is constant on the fiber over each compressed record*).

Proof. Necessity. Suppose $\mathcal{C}(\gamma) = \mathcal{C}(\gamma^*)$. Then the compressed record $z = \mathcal{C}(\gamma)$ is identical to the compressed record of γ^*. Any reconstruction operator $\mathcal{R}$ maps z to a single point, which cannot distinguish between γ and γ^*. Therefore $q(\gamma^*)$ cannot be recovered from z. Condition (i) is necessary.

Sufficiency. Suppose $\mathcal{C}(\gamma) \neq \mathcal{C}(\gamma^*)$ (condition i). Let $z = \mathcal{C}(\gamma)$ and $z^* = \mathcal{C}(\gamma^*)$. By condition (ii), q factors as $q = \hat{q} \circ \mathcal{C}$. Therefore $q(\gamma^*) = \hat{q}(\mathcal{C}(\gamma^*)) = \hat{q}(z^*)$. Since $z \neq z^*$ and we have access to the compressed space $\mathcal{Z}$, we can evaluate $\hat{q}$ at z^* to recover $q(\gamma^*)$. ▪ ▪

Remark 28.1 (Condition (i) is the Binding Constraint). In practice, condition (ii) (sufficiency) can often be arranged by choosing the right compression scheme. Condition (i) is the binding constraint: the compression must preserve the distinction between actual and counterfactual trajectories. This fails when the compression collapses the two trajectories into the same fiber — i.e., when the distinction between γ and γ^* is not reachability-relevant in the sense of Definition 6.2.

28.3 Minimal Compression Depth for Counterfactual Reliability

Definition 28.2 (*Counterfactual Divergence Point*). The **divergence depth** of a counterfactual pair (γ, γ^*) is the time of intervention: $\tau = \inf\{t : \gamma(t) \neq \gamma^*(t)\}$.

Proposition 28.2 (*Depth-Dependent Recoverability*). *For a compression $\mathcal{C}_D$ that retains history up to depth D (i.e., stores the trajectory only for $t \leq D$): a counterfactual with divergence depth $\tau < D$ is recoverable; a counterfactual with $\tau > D$ is not recoverable from $\mathcal{C}_D(\gamma)$.*

Proof. If $\tau < D$: the divergence point is within the stored window. $\mathcal{C}_D$ records $\gamma(t)$ for $t \leq D$, which includes the divergence. Therefore $\mathcal{C}_D(\gamma) \neq \mathcal{C}_D(\gamma^*)$ (condition i).

If $\tau > D$: the divergence point is beyond the stored window. $\gamma(t) = \gamma^*(t)$ for $t \leq D$, so $\mathcal{C}_D(\gamma) = \mathcal{C}_D(\gamma^*)$. Condition (i) fails; the counterfactual is unrecoverable. ▪ ▪

Memory Horizon Bounds Counterfactual Reach. The compression depth D is the *counterfactual horizon*: the system can reason about alternatives that diverged within D but cannot access alternatives that diverged further back. This is why institutional memory failures are often counterfactual failures: losing the record of how things developed removes the ability to ask "what if we had done X instead." Adequate counterfactual reasoning requires that the compression retain history back to the divergence point of relevant alternatives.

Exercises

28.1. A policy analyst wants to evaluate the counterfactual "what would GDP have been without the 2008 intervention?" The divergence point is 2008. What minimum compression depth (in years) is required for this analysis?

28.2. Prove that for any counterfactual pair (γ, γ^*) with divergence depth τ, there exists a compression $\mathcal{C}$ of depth $D < \tau$ that makes the counterfactual unrecoverable.

28.3. Model the memory of a scientific discipline as a compression $\mathcal{C}_{\text{sci}}$ of the trajectory of experimental results. Under what conditions can the discipline ask "what would we have found if we had tested hypothesis H^*?"

28.4. (Causal inference.) The potential outcomes framework (Rubin causal model) defines treatment effects as differences between potential outcomes $Y(1)$ and $Y(0)$. Show that the fundamental problem of causal inference (we observe only $Y(T_i)$ for treatment $T_i \in \{0, 1\}$) is exactly the Counterfactual Recovery Lemma with $\mathcal{C}$ = "record the observed treatment outcome only." When is this compression sufficient for the average treatment effect?

29 Chapter 29

Alternative Histories

> Two rivers that merge cannot be told apart by examining only the water downstream.
>
> —

Phenomenological Note. History seems inevitable in retrospect because only one version of it happened. But the people living through it did not know which version would survive. From inside, the future was branching. From outside, looking back, it appears linear. The alternative histories are real as possibilities. They are absent only as facts. The difference between them tells you something about what the constraints actually were.

Chapter 28 showed that counterfactual recovery requires the two trajectories to be distinguishable in compressed space. This chapter proves the complementary result: when two distinct trajectories lie in the same compression fiber, no internal reconstruction operation can distinguish them. Alternative histories that compress identically are permanently merged.

29.1 The Alternative History Collapse Theorem

Theorem 29.1 (*Alternative History Collapse*)**.** *Let $\gamma_1, \gamma_2 \in \Gamma$ be distinct trajectories with $\gamma_1 \neq \gamma_2$ but $\mathcal{C}(\gamma_1) = \mathcal{C}(\gamma_2) = z$. Then for any reconstruction operator $\mathcal{R} : \mathcal{Z} \to \Gamma$:*

$$\mathcal{R}(z) = \mathcal{R}(\mathcal{C}(\gamma_1)) = \mathcal{R}(\mathcal{C}(\gamma_2)).$$

No internal operation can distinguish γ_1 from γ_2 given only the compressed record z.

Proof. Since $\mathcal{C}(\gamma_1) = \mathcal{C}(\gamma_2) = z$, the reconstruction operator $\mathcal{R}$ receives identical input z regardless of which trajectory produced it. $\mathcal{R}$ is a function, so $\mathcal{R}(z)$ is a single value, not dependent on which γ_i produced z. Therefore $\mathcal{R}(z)$ is the same for both.

More precisely: any operation internal to the reconstruction system has access only to z and to whatever prior information $\mathcal{R}$ has (which by assumption is fixed and independent of γ_i). The input z is identical in both cases. Therefore the output is identical. ∎

29.2 The Asymmetry of Forgetting

Theorem 29.1 reveals an important asymmetry.

Proposition 29.2 (*External Information Can Distinguish*). *If an agent A external to the compressed system has independent access to information about which trajectory occurred, then A can distinguish γ_1 from γ_2 even though the internal compressed representation cannot.*

Proof. If $\mathcal{C}(\gamma_1) = \mathcal{C}(\gamma_2) = z$, then by the Alternative History Collapse Theorem, no internal query $q \in \mathcal{Q}_\mathcal{C}$ distinguishes γ_1 from γ_2. External information e is precisely the information not encoded in z: $e \perp z$ in the information diagram. A query q_e that depends on e can compute $q_e(\gamma_1) = f(\gamma_1, e)$ and $q_e(\gamma_2) = f(\gamma_2, e)$, which may differ since f accesses γ directly, not through z. ∎

This is the formal basis for:

- *Testimony in law*: witnesses can distinguish histories that official records have merged.
- *Independent replication in science*: a second experiment can distinguish between hypotheses that a compressed summary collapsed.
- *Cross-examination*: questioning a compressed account against external evidence reveals what the compression destroyed.

29.3 The Fiber as Indistinguishability Class

Definition 29.1 (*Compressed Indistinguishability*). Trajectories γ_1, γ_2 are **compressed-indistinguishable** under $\mathcal{C}$ if $\mathcal{C}(\gamma_1) = \mathcal{C}(\gamma_2)$. The fiber $\mathcal{C}^{-1}(z)$ is the **indistinguishability class** of z: all histories that a system holding only z cannot tell apart.

Proposition 29.3 (*Indistinguishability is an Equivalence Relation*). *Compressed indistinguishability is reflexive, symmetric, and transitive. Its equivalence classes are precisely the fibers of $\mathcal{C}$.*

Proof. Reflexivity: $\mathcal{C}(\gamma) = \mathcal{C}(\gamma)$ trivially. Symmetry: if $\mathcal{C}(\gamma_1) = \mathcal{C}(\gamma_2)$ then $\mathcal{C}(\gamma_2) = \mathcal{C}(\gamma_1)$. Transitivity: if $\mathcal{C}(\gamma_1) = \mathcal{C}(\gamma_2)$ and $\mathcal{C}(\gamma_2) = \mathcal{C}(\gamma_3)$ then $\mathcal{C}(\gamma_1) = \mathcal{C}(\gamma_3)$. The equivalence classes are the fibers $\mathcal{C}^{-1}(z)$ by definition of the equivalence relation. ∎

The size of the indistinguishability class $|\mathcal{C}^{-1}(z)|$ (or more precisely, $\mu(\mathcal{C}^{-1}(z))$ in the continuous case) is exactly the fiber entropy $S_\mathcal{C}(z) = \log \mu(\mathcal{C}^{-1}(z))$ of Chapter 13. Alternative history collapse is therefore the same phenomenon as high fiber entropy: the compressed record is consistent with many different actual histories.

29.4 Connections to Quantum Mechanics

In quantum mechanics, decoherence causes a superposition of alternative histories to collapse into a single classical record . The CPR interpretation: decoherence is a compression operator $\mathcal{C}_{\text{dec}}$ (Pearl 2000) that maps the quantum history space Γ_{quantum} (superpositions of classical trajectories) to classical records $\mathcal{Z}_{\text{cl}}$.

Proposition 29.4 (*Decoherence as Compression*)**.** *After decoherence, distinct quantum branches $|\gamma_1\rangle, |\gamma_2\rangle$ that produce the same classical pointer state z satisfy $\mathcal{C}_{\text{dec}}(|\gamma_1\rangle) = \mathcal{C}_{\text{dec}}(|\gamma_2\rangle) = z$. By Theorem 29.1, no further classical operation can recover the quantum distinction. The many-worlds branching has been permanently compressed.*

Proof. In quantum mechanics, decoherence maps a pure state $|\psi\rangle\langle\psi|$ to a mixed state $\sum_k p_k |k\rangle\langle k|$ by tracing over environmental degrees of freedom. This is a compression $\mathcal{C}_{\text{deco}}$: the environment is the fiber and the pointer basis $\{|k\rangle\}$ is the base space. Histories that differ only in the environmental degrees of freedom are indistinguishable post-decoherence, mapping to identical reduced density matrices. Alternative quantum histories with the same classical pointer record form the fiber $\mathcal{C}_{\text{deco}}^{-1}(|k\rangle\langle k|)$. ▪ ▪

Exercises

29.1. Let $\Gamma = \{A, B, C\}$ (three trajectories) and $\mathcal{C}(A) = \mathcal{C}(B) = z_1$, $\mathcal{C}(C) = z_2$. Suppose $q(A) = 0$, $q(B) = 1$, $q(C) = 0$. Show that $\mathcal{C}$ is not sufficient for q, and that no reconstruction from z_1 can answer q reliably.

29.2. Prove that the set of trajectories for which q is recoverable under $\mathcal{C}$ is exactly the union of fibers on which q is constant. Express this as a condition on the partition induced by $\mathcal{C}$.

29.3. (History.) Two paths to the French Revolution — one through the Estates-General, one through a coup — both produced the outcome labelled "the Revolution." Model this as alternative history collapse. What historical information would allow a historian to distinguish the paths even given only the compressed record?

29.4. Construct a compression scheme for a Markov chain such that the resulting compressed process is itself Markov. Under what condition on the original chain's transition matrix does such a "lumpable" compression exist?

30 Chapter 30

The Geometry of Forgetting

> To forget is not to lose information. It is to expand the set of worlds consistent with what remains.
>
> —

Phenomenological Note. Forgetting is not symmetrical with remembering. You tend to notice what you still have. The shape of what has gone is mostly invisible, which is part of why loss accumulates undetected. You know something is missing only when you reach for it and find the gap. By then the geometry of what could have been retrieved has already changed.

Forgetting is modelled in the CPR framework as *fiber expansion*: a coarsening of the compression operator that increases the size of each fiber. (Shannon and Weaver 1949) This chapter derives the Geometry of Forgetting theorem, showing that information entropy change under forgetting equals the log-ratio of fiber volumes.

30.1 Forgetting as Compression Coarsening

Definition 30.1 (*Forgetting Operator*)**.** A **forgetting operator** is a surjective map $F : \mathcal{Z}_1 \to \mathcal{Z}_2$ from a finer compressed space to a coarser one, such that $|\mathcal{Z}_2| \leq |\mathcal{Z}_1|$ (or $\dim \mathcal{Z}_2 \leq \dim \mathcal{Z}_1$). Forgetting at time $t_2 > t_1$ corresponds to replacing the compression $\mathcal{C}_{t_1} : \Gamma \to \mathcal{Z}_1$ with the coarser compression $\mathcal{C}_{t_2} = F \circ \mathcal{C}_{t_1}$.

The fiber over $z_2 \in \mathcal{Z}_2$ under $\mathcal{C}_{t_2}$ is:

$$\mathcal{C}_{t_2}^{-1}(z_2) = \bigcup_{z_1 \in F^{-1}(z_2)} \mathcal{C}_{t_1}^{-1}(z_1) \supseteq \mathcal{C}_{t_1}^{-1}(z_1) \quad \text{for each } z_1 \in F^{-1}(z_2).$$

Each forgetting step merges previously distinct fibers, expanding the total fiber at each point.

30.2 The Geometry of Forgetting Theorem

Theorem 30.1 (*Geometry of Forgetting*). *Let $\mathcal{C}_{t_1}$ and $\mathcal{C}_{t_2} = F \circ \mathcal{C}_{t_1}$ be two compressions at times $t_1 < t_2$, with $F : \mathcal{Z}_1 \to \mathcal{Z}_2$ the forgetting operator. For a state $z_2 \in \mathcal{Z}_2$ that was compressed from $z_1 \in \mathcal{Z}_1$:*

$$\Delta S(z_1 \to z_2) = S_{\mathcal{C}_{t_2}}(z_2) - S_{\mathcal{C}_{t_1}}(z_1) = \log \frac{\mathrm{Vol}(\mathcal{C}_{t_2}^{-1}(z_2))}{\mathrm{Vol}(\mathcal{C}_{t_1}^{-1}(z_1))} \ \geq \ 0.$$

Forgetting is non-negative: the projection entropy never decreases as compression becomes coarser.

Proof. Since $\mathcal{C}_{t_2} = F \circ \mathcal{C}_{t_1}$, the fiber at z_2 under $\mathcal{C}_{t_2}$ is:

$$\mathcal{C}_{t_2}^{-1}(z_2) = (F \circ \mathcal{C}_{t_1})^{-1}(z_2) = \mathcal{C}_{t_1}^{-1}(F^{-1}(z_2)).$$

If $F^{-1}(z_2) = \{z_1\}$ (forgetting merges nothing at z_1): $\mathcal{C}_{t_2}^{-1}(z_2) = \mathcal{C}_{t_1}^{-1}(z_1)$, and $\Delta S = 0$.

If $|F^{-1}(z_2)| > 1$ (forgetting merges multiple z_1-fibers):

$$\mathcal{C}_{t_2}^{-1}(z_2) = \bigsqcup_{z_1' \in F^{-1}(z_2)} \mathcal{C}_{t_1}^{-1}(z_1') \supseteq \mathcal{C}_{t_1}^{-1}(z_1).$$

Therefore: $\mathrm{Vol}(\mathcal{C}_{t_2}^{-1}(z_2)) \geq \mathrm{Vol}(\mathcal{C}_{t_1}^{-1}(z_1))$,

so: $\Delta S = \log \frac{\mathrm{Vol}(\mathcal{C}_{t_2}^{-1}(z_2))}{\mathrm{Vol}(\mathcal{C}_{t_1}^{-1}(z_1))} \geq 0.$ ▪ ▪

30.3 Total Entropy of Forgetting

Corollary 30.2 (*Expected Entropy Growth*). *The expected entropy change over a forgetting step satisfies:*

$$\mathbb{E}_{z_2}[\Delta S(z_1 \to z_2)] = H(\Gamma \mid \mathcal{Z}_2) - H(\Gamma \mid \mathcal{Z}_1) = I(\Gamma; \mathcal{Z}_1) - I(\Gamma; \mathcal{Z}_2) \ \geq \ 0.$$

The mutual information between histories and compressed records decreases with each forgetting step. The total information lost by forgetting from t_1 to t_2 is:

$$\mathcal{I}_{\mathrm{lost}} = I(\Gamma; \mathcal{Z}_1) - I(\Gamma; \mathcal{Z}_2) \ \geq \ 0.$$

Proof. By the data processing inequality: $I(\Gamma; \mathcal{Z}_2) = I(\Gamma; F(\mathcal{Z}_1)) \leq I(\Gamma; \mathcal{Z}_1)$, since $\mathcal{Z}_2$ is a function of $\mathcal{Z}_1$. ▪ ▪

30.4 Forgetting and Reachability

Proposition 30.3 (*Forgetting Reduces Counterfactual Reach*). *After a forgetting step $\mathcal{C}_{t_1} \to \mathcal{C}_{t_2}$, the set of counterfactuals recoverable from the compressed record* *can only decrease: any counterfactual recoverable from $\mathcal{C}_{t_2}$ was also recoverable from $\mathcal{C}_{t_1}$, but not vice versa.*

Proof. By Lemma 28.1, a counterfactual γ^* is recoverable from z iff $\mathcal{C}(\gamma) \neq \mathcal{C}(\gamma^*)$. If $\mathcal{C}_{t_2}(\gamma) \neq \mathcal{C}_{t_2}(\gamma^*)$, then $F(\mathcal{C}_{t_1}(\gamma)) \neq F(\mathcal{C}_{t_1}(\gamma^*))$, which implies $\mathcal{C}_{t_1}(\gamma) \neq \mathcal{C}_{t_1}(\gamma^*)$. So recoverability under $\mathcal{C}_{t_2}$ implies recoverability under $\mathcal{C}_{t_1}$. The converse fails: $\mathcal{C}_{t_1}(\gamma) \neq \mathcal{C}_{t_1}(\gamma^*)$ does not imply $\mathcal{C}_{t_2}(\gamma) \neq \mathcal{C}_{t_2}(\gamma^*)$ when F merges the two records. ▪ ▪

Forgetting is Irreversible Distinction Loss. Theorem 30.1 and Proposition 30.3 together say: forgetting is irreversible loss of distinction. Once a coarser compression is applied, no internal operation can recover the lost distinctions. The only recourse is external information — independent records, testimony, physical evidence — that bypasses the compressed record entirely. This is why archival preservation, scientific replication, and independent institutional records are not redundancy: they are the only protection against the permanent closing of counterfactual possibility.

Exercises

30.1. Let $\Gamma = \{A, B, C, D\}$ with $\mathcal{C}_{t_1}(A) = \mathcal{C}_{t_1}(B) = z_1$ and $\mathcal{C}_{t_1}(C) = \mathcal{C}_{t_1}(D) = z_2$. Define the forgetting operator $F(z_1) = F(z_2) = z$ (total forgetting). Compute ΔS for both z_1 and z_2. What is the total entropy of forgetting?

30.2. Prove that the total entropy of forgetting is additive over sequential forgetting steps: $\mathcal{I}_{\text{lost}}(t_1 \to t_3) = \mathcal{I}_{\text{lost}}(t_1 \to t_2) + \mathcal{I}_{\text{lost}}(t_2 \to t_3)$.

30.3. A newspaper archives stories at resolution r_1 for one year, then compresses to resolution $r_2 < r_1$ for the digital archive. Model this as a two-step forgetting process. Derive the total information lost as a function of r_1 and r_2.

30.4. (Thermodynamics.) The Landauer principle states that erasing one bit of information requires at least $kT \ln 2$ energy. Model bit erasure as a forgetting operator and derive the Landauer bound from the Geometry of Forgetting theorem.

Part V

Cognition

[Part introduction — to be written.]

31 Chapter 31

Perception as Projection

> The eye does not see. It projects.
>
> —

Phenomenological Note. You do not see the world. You see a version of the world that your nervous system has assembled from partial signals, filtered through expectations, edited for consistency. The editing is not optional and not conscious. By the time perception reaches awareness it has already been corrected, completed, and smoothed. The raw input is never available. Only the reconstruction is.

Perception is the process by which a cognitive system maps high-dimensional environmental states to lower-dimensional perceptual states. In the CPR framework, this is a projection: $\pi : \mathcal{E} \to \mathcal{P}$ from environment space to percept space. The key theorem is that perceptual ambiguity arises directly from nontrivial fibers: when many distinct environmental states project to the same percept, the percept is inherently ambiguous.

31.1 The Perceptual Projection

Definition 31.1 (*Perceptual System*). A **perceptual system** is a projection $\pi_{\text{per}} : \mathcal{E} \to \mathcal{P}$ (Gibson 1979; Marr 1982) from an environmental trajectory space $\mathcal{E}$ (the space of all possible environmental states and histories) to a percept space $\mathcal{P}$ (the space of internal sensory representations).

The fiber $\pi_{\text{per}}^{-1}(p)$ over a percept $p \in \mathcal{P}$ is the set of all environmental histories that produce the same percept. Ambiguity is geometrised: the richer the fiber, the more ambiguous the percept.

31.2 The Perception-Ambiguity Theorem

Theorem 31.1 (*Perception-Ambiguity Theorem*). *A percept $p \in \mathcal{P}$ is* unambiguous *if and only if $\pi_{\text{per}}^{-1}(p)$ is a singleton.* Perceptual ambiguity *arises when the fiber $\pi_{\text{per}}^{-1}(p)$ contains multiple environmental states with distinct causal con-*

sequences:

$$\exists\, e_1, e_2 \in \pi_{\text{per}}^{-1}(p): \quad \mathcal{R}_{\mathcal{A}}(e_1, T) \neq \mathcal{R}_{\mathcal{A}}(e_2, T).$$

In that case, the percept p is reachability-ambiguous *at horizon T.*

Proof. If $\pi_{\text{per}}^{-1}(p)$ is a singleton $\{e\}$, then the percept p uniquely identifies the environment e and hence uniquely determines the reachable set $\mathcal{R}_{\mathcal{A}}(e, T)$. No ambiguity.

If $|\pi_{\text{per}}^{-1}(p)| > 1$, and if distinct elements of the fiber have distinct reachable sets, then the perceptual system cannot determine which reachable set applies — it sees p but cannot know whether $\mathcal{R}(e_1, T)$ or $\mathcal{R}(e_2, T)$ is the set of futures it faces. This is reachability ambiguity. ▪ ▪

31.3 Context as Fiber Reduction

If the perceptual system receives additional information c (context: prior expectations, other modalities, memory), this restricts the plausible environmental states:

Definition 31.2 (*Contextualised Fiber*)**.** Given context $c \in \mathcal{C}$ and a likelihood function $\ell : \mathcal{E} \times \mathcal{C} \to \mathbb{R}$, the **contextualised fiber** over percept p is:

$$\mathcal{F}(p, c) = \{e \in \pi_{\text{per}}^{-1}(p) : \ell(e, c) > 0\}.$$

Context reduces the effective fiber, resolving ambiguity when $|\mathcal{F}(p, c)| < |\pi_{\text{per}}^{-1}(p)|$.

Proposition 31.2 (*Context Resolves Ambiguity*)**.** *If $\mathcal{F}(p, c)$ is a singleton for all $p \in \mathcal{P}$, then the contextualised perceptual system is unambiguous: the combination of percept and context uniquely identifies the environmental state. Formally, the map $(p, c) \mapsto \mathcal{F}(p, c)$ is an injective function from $\mathcal{P} \times \mathcal{C}$ to $\mathcal{E}$.*

Proof. Context c provides additional information that constrains the fiber $\pi^{-1}(m)$. Formally, the posterior fiber given context is $\pi^{-1}(m \mid c) = \{x \in \pi^{-1}(m) : x \text{ consistent with } c\}$. If context is informative (reduces the fiber), $|\pi^{-1}(m \mid c)| < |\pi^{-1}(m)|$. Sufficient context reduces the fiber to a singleton, $|\pi^{-1}(m \mid c)| = 1$, giving a unique percept. ▪ ▪

This formalises the Bayesian account of perception: context c is the prior, the percept p is the likelihood, and the contextualised fiber is the posterior. Perception is inference over the inverse problem of determining which environmental state caused the percept.

31.4 Multi-Stable Perception

Multi-stable perceptual phenomena (Necker cube, Rubin vase, binocular rivalry) occur when $\mathcal{F}(p, c)$ has multiple elements with distinct and non-integrable reachable sets — the visual system switches between fiber elements because no single interpretation dominates.

Multi-stability as Fiber Non-Uniqueness. Multi-stable perception is the direct phenomenological consequence of the fiber $\pi_{\text{per}}^{-1}(p)$ being genuinely ambiguous even after contextualisation: $|\mathcal{F}(p,c)| > 1$ for any available context c. The visual system oscillates between interpretations because it cannot resolve which fiber element is correct. The oscillation period is inversely related to the relative reachability volumes of the fiber elements: the interpretation with greater reachability dominates longer.

Exercises

31.1. Model the perception of a speech sound in noise as $\pi_{\text{per}}(e) =$ most likely phoneme, where e is the full acoustic waveform. Characterise the fiber $\pi_{\text{per}}^{-1}(\text{phoneme})$ and give two examples of distinct acoustic environments in the same fiber with different reachable sets.

31.2. Prove that if π_{per} is a bijection, then no perceptual ambiguity is possible. Interpret: a perfect sensory system (infinite bandwidth, no noise) is unambiguous.

31.3. Define the *ambiguity entropy* of a percept p: $H_{\text{amb}}(p) = H(\text{environmental state} \mid \text{percept} = p)$. Show that $H_{\text{amb}}(p) = \log \mu(\pi_{\text{per}}^{-1}(p))$ for a uniform distribution. Connect to the fiber entropy $S_\pi(p)$ of Chapter 13.

31.4. A context signal c reduces ambiguity entropy by ΔH. Find the minimum context dimensionality $\dim(\mathcal{C})$ needed to reduce $H_{\text{amb}}(p)$ to zero for all percepts in $\mathcal{P}$. (Answer: $\dim(\mathcal{C}) \geq H_{\text{amb}}(\mathcal{P}) / \log 2$ bits.)

32 Chapter 32

Belief Manifolds

> A belief is not a proposition held. It is a probability distribution maintained.
>
> —

Phenomenological Note. Beliefs are not held individually. They form a landscape where each belief is supported by others and in turn supports others. Changing one belief changes what the neighboring beliefs mean. This is why belief revision feels threatening even when the evidence is clear: to update one thing is to destabilize others that depend on it, and the full cost of the revision is not visible until you are already in the middle of it.

If perception is projection, then cognition is navigation on the space of beliefs. A belief state is a probability distribution (Amari and Nagaoka 2000; Jaynes 2003) over possible worlds; the space of all such distributions is a statistical manifold equipped with the Fisher metric. Reasoning is motion on this manifold; evidence is a constraint that restricts which regions of the manifold are accessible.

32.1 Belief States as Probability Distributions

Definition 32.1 (*Belief Manifold*)**.** Let $\mathcal{W}$ be a space of possible worlds. A **belief state** is a probability distribution $P \in \Delta(\mathcal{W})$ over possible worlds. The **belief manifold** $\mathcal{B}$ is the statistical manifold $\mathcal{B} = \{P_\theta : \theta \in \Theta\}$ of all probability distributions over $\mathcal{W}$ parameterised by cognitive coordinates $\theta \in \Theta \subseteq \mathbb{R}^n$.

The Fisher metric on $\mathcal{B}$:

$$g_{ij}(\theta) = \mathbb{E}_{P_\theta}\left[\partial_i \log P_\theta(w)\, \partial_j \log P_\theta(w)\right]$$

defines the *psychological distance* between belief states: two beliefs are far apart if the distribution over worlds changes rapidly in their direction.

32.2 Bayesian Updating as Geodesic Motion

Theorem 32.1 (*Inference as Geodesic*). *Optimal Bayesian updating from prior P_θ given evidence e follows the geodesic on $\mathcal{B}$ that minimises the information-theoretic cost functional:*

$$\mathcal{L}[\theta(t)] = \int_0^1 \sqrt{g_{ij}(\theta(t))\dot{\theta}^i(t)\dot{\theta}^j(t)}\,\mathrm{d}t,$$

subject to $\theta(0) = \theta_0$ (prior) and $\theta(1) = \theta^*$ (posterior given e). The posterior P_{θ^*} satisfies: $P_{\theta^*}(w) \propto P_{\theta_0}(w) \cdot p(e \mid w)$ (Bayes' theorem).

Proof sketch. The information-geometric geodesic minimises $D_{\mathrm{KL}}(P_\theta \,\|\, P_{\theta^*})$ along the path (using the dual e-connection on $\mathcal{B}$). The endpoint condition forces θ^* to satisfy Bayes' rule: the m-geodesic from P_{θ_0} to the exponential family constraint $\{P : \mathbb{E}_P[\log p(e \mid w)] = \mathrm{const}\}$ is the posterior $P_{\theta^*} \propto P_{\theta_0} p(e \mid w)$. ▪ ▪

Reasoning as Navigation. Updating beliefs in light of evidence is motion on $\mathcal{B}$. The direction of motion is determined by the evidence. The speed of motion is determined by the strength of the evidence. Reasoning that covers long geodesic distances is "surprising" (it changes beliefs substantially); reasoning that covers short distances is "confirming." Closed loops on $\mathcal{B}$ are circular arguments.

32.3 Admissibility on the Belief Manifold

Not all belief states are epistemically admissible. A belief state P_θ is admissible if it satisfies the background constraints of the cognitive system: consistency with known facts, coherence with sensory evidence, and calibration with past experience.

Definition 32.2 (*Epistemic Admissibility*). The **epistemic admissibility field** $\mathcal{A}_{\mathrm{ep}}$ on $\mathcal{B}$ is the set of belief states satisfying the system's background epistemic constraints:

$$\mathcal{A}_{\mathrm{ep}} = \{P_\theta \in \mathcal{B} : K_i(P_\theta) \geq 0 \;\forall i\},$$

where K_i are constraint functionals (e.g., $K_{\mathrm{cons}}(P) = -D_{\mathrm{KL}}(P \| P_{\mathrm{consistent}})$, $K_{\mathrm{cal}}(P) = -\mathbb{E}[\text{calibration error of } P]$).

Epistemic reasoning is then navigation on $\mathcal{B}$ within the admissibility field $\mathcal{A}_{\mathrm{ep}}$ — the same structure as physical dynamics on $\mathcal{X}$ within the physical admissibility field $\mathcal{A}$.

32.4 Belief Rigidity and Flexibility

> **Definition 32.3** (*Belief Rigidity*). The **belief rigidity** at P_θ is
>
> $$\rho(\theta) = \lambda_{\text{min}}(\mathcal{I}(\theta))^{-1},$$
>
> the reciprocal of the smallest eigenvalue of the Fisher metric. High rigidity means the belief state is insensitive to evidence in the direction of λ_{min}: it has a "blind spot" in that direction.

Rigid belief states resist update: they respond weakly to evidence in their blind directions. Fisher-degenerate belief states (Chapter 17) have infinite rigidity in some directions — no evidence can move them.

Proposition 32.2 (*Dogmatism as Fisher Degeneracy*). *A belief state is **dogmatic** in direction v iff $g_{ij}v^i v^j = 0$ — the Fisher metric degenerates in that direction. No likelihood ratio can distinguish P_θ from $P_{\theta+tv}$ for any $t > 0$. Dogmatism is the belief-manifold form of distinction collapse.*

Proof. A dogmatic belief is one that does not update on evidence: $p(\theta \mid x) \approx p(\theta)$ for all data x. In information geometry, the update magnitude is the Fisher information: $\Delta\theta \propto \mathcal{I}(\theta)^{-1}\nabla_\theta \log p(x|\theta)$. When $\mathcal{I}(\theta) \to \infty$ in some direction (infinite confidence), updates in that direction are suppressed: $\Delta\theta \to 0$. Degenerate Fisher matrix ($\det \mathcal{I} = 0$) means zero sensitivity to data in the null directions: the belief manifold collapses those dimensions. ▪ ▪

Exercises

32.1. Let $\mathcal{B} = \{\text{Bernoulli}(p) : p \in (0,1)\}$. Compute the Fisher metric $g(p)$ and the geodesic distance between p_1 and p_2. Interpret: what is the information-geometric "length" of a Bayesian update from p_1 to p_2?

32.2. Prove that the posterior $P_{\theta^*} \propto P_{\theta_0} p(e \mid w)$ satisfies the m-projection property: $D_{\text{KL}}(P_{\theta^*} \,\|\, P_{\theta_0}) \leq D_{\text{KL}}(Q \,\|\, P_{\theta_0})$ for all Q with $\mathbb{E}_Q[\log p(e \mid w)] = \mathbb{E}_{P_{\theta^*}}[\log p(e \mid w)]$.

32.3. Define a *belief trajectory* as a path $\theta : [0,T] \to \mathcal{B}$ and show that its total Fisher length $\int_0^T \sqrt{g_{ij}\dot{\theta}^i\dot{\theta}^j}\, dt$ is equal to the total information processed. Connect to the KL divergence between initial and final beliefs.

32.4. A cognitive system has two competing hypotheses H_1, H_2 with prior $(p, 1-p)$. Model the belief manifold as a 1D manifold and compute how belief rigidity ρ varies as $p \to 0$, $p \to 1$, and $p \to 1/2$. Interpret: when is the system most and least open to evidence?

33 Chapter 33

Constraint-Guided Inference

> Reasoning is not free association. It is geodesic motion under constraint.
>
> —

Phenomenological Note. When the data is ambiguous, you use everything else you know. You fill in from context, from pattern, from prior experience. This is not a failure of reasoning but the mechanism of it. Pure data without prior structure would be uninterpretable. The prior structure is what makes inference possible. The question is whether the structure you are using is the right structure for this situation.

Chapter 32 established that belief states are points on a statistical manifold $\mathcal{B}$ equipped with the Fisher metric, and that optimal Bayesian updating follows geodesics. But real cognitive inference is constrained : not all belief transitions are admissible. This chapter derives the equation of motion for *constraint-guided inference* — geodesic motion on $\mathcal{B}$ (cf. Friston 2010; Jaynes 2003) perturbed by admissibility forces. The derivation is a complete Lagrangian variational argument.

33.1 The Constrained Inference Action

Let $\mathcal{B}$ be a belief manifold with Fisher metric g_{ij}. A **belief trajectory** is a smooth curve $b : [0,T] \to \mathcal{B}$. Let $C(b,t) \geq 0$ be a **constraint potential** measuring epistemic inadmissibility: $C(b,t) = 0$ when $b \in \mathcal{A}_{\text{ep}}$ (fully admissible) and $C(b,t) > 0$ outside.

Definition 33.1 (*Constrained Inference Action*)**.** The **constrained inference action** is

$$\mathcal{S}[b] = \int_0^T \left(\tfrac{1}{2}\, g_{ij}(b)\, \dot{b}^i\, \dot{b}^j + \lambda\, C(b,t) \right) \mathrm{d}t,$$

where $\lambda > 0$ is the constraint coupling constant. The first term penalises informational distance traversed; the second penalises departure from the admissibility field.

Optimal inference minimises $\mathcal{S}[b]$ subject to $b(0) = b_0$ (prior) and $b(T) = b^*$ (posterior).

33.2 The Constrained Inference Geodesic Equation

Theorem 33.1 (*Constrained Inference Geodesic*)**.** *The optimal belief trajectory minimising $\mathcal{S}[b]$ satisfies the Euler-Lagrange equation:*

$$\ddot{b}^i + \Gamma^i_{jk}\,\dot{b}^j\,\dot{b}^k = -\lambda\, g^{ij}\,\partial_j C(b,t),$$

where Γ^i_{jk} are the Christoffel symbols of the Fisher metric. Equivalently:

$$\nabla_{\dot{b}}\dot{b} = -\lambda\ \mathrm{grad}_g\, C,$$

where ∇ is the Levi-Civita connection of $(\mathcal{B}, g)$.

Proof. Take a variation $b^i(t) \mapsto b^i(t) + \epsilon\xi^i(t)$ with $\xi(0) = \xi(T) = 0$ (fixed endpoints). The first variation of $\mathcal{S}$ is:

$$\delta\mathcal{S} = \int_0^T \left[g_{ij}(b)\,\dot{b}^j\,\delta\dot{b}^i + \tfrac{1}{2}\partial_i g_{jk}(b)\,\dot{b}^j\dot{b}^k\,\delta b^i + \lambda\partial_i C\,\delta b^i\right] \mathrm{d}t.$$

Integrating the first term by parts (using $\delta\dot{b}^i = \frac{\mathrm{d}}{\mathrm{d}t}\delta b^i$ and vanishing boundary terms):

$$\int_0^T g_{ij}\dot{b}^j\,\delta\dot{b}^i\,\mathrm{d}t = -\int_0^T \frac{\mathrm{d}}{\mathrm{d}t}(g_{ij}\dot{b}^j)\,\delta b^i\,\mathrm{d}t.$$

Expanding the total derivative:

$$\frac{\mathrm{d}}{\mathrm{d}t}(g_{ij}\dot{b}^j) = g_{ij}\ddot{b}^j + \partial_k g_{ij}\dot{b}^k\dot{b}^j.$$

Substituting and collecting:

$$\delta\mathcal{S} = \int_0^T \left[-g_{ij}\ddot{b}^j - \partial_k g_{ij}\dot{b}^k\dot{b}^j + \tfrac{1}{2}\partial_i g_{jk}\dot{b}^j\dot{b}^k + \lambda\partial_i C\right]\xi^i\,\mathrm{d}t.$$

Since $\delta\mathcal{S} = 0$ for all ξ, the integrand vanishes:

$$g_{ij}\ddot{b}^j = -\partial_k g_{ij}\dot{b}^k\dot{b}^j + \tfrac{1}{2}\partial_i g_{jk}\dot{b}^j\dot{b}^k + \lambda\partial_i C.$$

Contracting with g^{li} and using the standard identity for Christoffel symbols $\Gamma^l_{jk} = \frac{1}{2}g^{li}(\partial_j g_{ki} + \partial_k g_{ji} - \partial_i g_{jk})$:

$$\ddot{b}^l + \Gamma^l_{jk}\dot{b}^j\dot{b}^k = \lambda g^{li}\partial_i C = \lambda(\mathrm{grad}_g\, C)^l. \quad \blacksquare$$

■

33.3 Interpretation

Theorem 33.1 gives two special cases:

Unconstrained inference ($\lambda = 0$).. $\nabla_{\dot{b}}\dot{b} = 0$: pure geodesic motion on $\mathcal{B}$. This is the information-geometric version of optimal Bayes updating — the posterior is reached by the shortest path through belief space.

Strongly constrained inference ($\lambda \to \infty$).. The constraint term dominates. With a barrier potential $C(b) = -\log d(b, \partial\mathcal{A}_{\text{ep}})$, beliefs are repelled from the inadmissible boundary: $\text{grad}_g\, C$ points away from $\partial\mathcal{A}_{\text{ep}}$. Inference is channelled along paths that respect the constraint field.

Admissibility as Force. The key insight is that admissibility constraints are not external filters applied after inference. They appear as a *force* in the dynamical equation of belief updating. Reasoning that violates background constraints is not incorrect reasoning that gets corrected post hoc — it is reasoning that experiences a restoring force pulling it back. This models the phenomenology of “that can't be right” as a dynamical constraint in the belief manifold, not a logical rule applied to propositions.

33.4 Barrier Potentials and Epistemic Walls

Definition 33.2 (*Admissibility Barrier*). The **admissibility barrier potential** is

$$C_{\text{bar}}(b) = -\log d(b, \partial\mathcal{A}_{\text{ep}}),$$

where $d(\cdot, \partial\mathcal{A}_{\text{ep}})$ is the distance from b to the inadmissibility boundary.

With this potential:

$$\nabla_{\dot{b}}\dot{b} = \lambda\, \text{grad}_g\, C_{\text{bar}} = -\lambda \nabla \log d(b, \partial\mathcal{A}_{\text{ep}}).$$

As $b \to \partial\mathcal{A}_{\text{ep}}$, the barrier force diverges, pushing the belief trajectory back into the admissible interior. This implements *hard epistemic constraints*: beliefs can approach the boundary but cannot cross it.

Example 33.1 (Conservation Beliefs). A physicist's belief about energy cannot update to “energy is sometimes not conserved” — this is outside $\mathcal{A}_{\text{ep}}$. The barrier potential represents the cumulative evidence that has hardened this constraint over centuries. Inference in the presence of apparent energy non-conservation (e.g., discovering dark energy) follows a trajectory along the boundary of $\mathcal{A}_{\text{ep}}$, incorporating the anomaly while preserving conservation.

Exercises

33.1. On a 1D belief manifold $\mathcal{B} = \mathbb{R}$ with metric $g = 1$ and constraint $C(b) = \max(0, b - b_{\max})^2$ (quadratic penalty beyond $b_{\max}$): (a) Write out the Euler-Lagrange equation. (b) Find the trajectory starting at $b_0 < b_{\max}$ with $\dot{b}_0 > 0$. (c) Show the trajectory slows and turns before reaching $b_{\max}$.

33.2. Prove that on $\mathcal{B}$ with zero curvature (flat metric), and with a linear constraint potential $C(b) = \alpha \cdot b$, the optimal trajectory is a parabola in time.

33.3. Let $\mathcal{B} = \{p \in (0,1)\}$ with Fisher metric $g(p) = 1/(p(1-p))$ and $\mathcal{A}_{\text{ep}} = [p_{\text{min}}, p_{\text{max}}]$. With the barrier potential C_{bar}, derive the force experienced at $p = p_{\text{min}} + \delta$ for small δ.

33.4. (Cognitive science.) The myside bias is the tendency to evaluate evidence in a way that protects pre-existing beliefs. Model this as an asymmetric constraint potential $C(b) = \beta \cdot \max(0, b - b_0)$ that penalises moving away from the prior b_0. Compute the posterior for evidence e under this model. How does β interpolate between open-minded and dogmatic inference?

34 Chapter 34

Memory Fields

> Memory is not storage. It is a field with local intensity, and retrieval is a threshold crossing.
>
> —

Phenomenological Note. A memory is not a recording. It is a field with a decay constant — something that was once vivid and is now faint, or was faint and has been strengthened by rehearsal. What you remember is partly what happened and partly what you have since repeated, recalled, confirmed. The archive and the archivist are the same process.

The MEM|8 framework (Appendix 92.16) models memory as a continuous field $M(x,t)$ over state space $\mathcal{X}$, evolving through decay and input. This chapter derives the memory field equation, proves the solution formula, and derives the retrieval threshold condition.

34.1 The Memory Field Equation

Definition 34.1 (*Memory Field*). The **memory field** $M : \mathcal{X} \times \mathbb{R} \to \mathbb{R}_{\geq 0}$ assigns to each state x at time t the intensity of the memory trace associated with x. M evolves according to:

$$\partial_t M(x,t) = -\lambda M(x,t) + I(x,t),$$

where $\lambda > 0$ is the **forgetting rate** and $I(x,t) \geq 0$ is the **input intensity** (how strongly state x is being experienced at time t).

Theorem 34.1 (*Memory Field Solution*). *The memory field at time t (cf. Baddeley 1986; Tulving 1983) with initial condition $M_0(x) = M(x,0)$ is:*

$$M(x,t) = M_0(x)e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)}\, I(x,s)\, \mathrm{d}s.$$

Proof. The equation $\partial_t M = -\lambda M + I$ is a first-order linear ODE in t for each fixed x. Multiply by the integrating factor $e^{\lambda t}$:

$$e^{\lambda t}\partial_t M + \lambda e^{\lambda t} M = e^{\lambda t} I, \quad \Rightarrow \quad \frac{\partial}{\partial t}(e^{\lambda t} M) = e^{\lambda t} I.$$

Integrating from 0 to t: $e^{\lambda t}M(x,t) - M_0(x) = \int_0^t e^{\lambda s} I(x,s)\,\mathrm{d}s$. Multiplying by $e^{-\lambda t}$ gives the result. ▪ ▪

Memory as Exponential Averaging. The memory field is an **exponential moving average** of the input history. Recent inputs (large $t-s$) are weighted more heavily (weight $e^{-\lambda(t-s)} \approx 1$); distant inputs are down-weighted (weight $e^{-\lambda(t-s)} \to 0$). The forgetting rate λ controls the effective memory window: inputs older than $1/\lambda$ have negligible influence. This is the RSVP connection: M plays the role of Φ (capacity for retrieval), and λ is the capacity drain rate.

34.2 Retrieval Threshold and the Active Memory Set

Definition 34.2 (*Active Memory Set*)**.** The **active memory set** at time t is:

$$\mathcal{M}_{\text{act}}(t) = \{x \in \mathcal{X} : M(x,t) \geq \theta_M\}$$

for retrieval threshold $\theta_M > 0$. These are the memory traces currently above the retrieval threshold — the "working memory" or accessible long-term memory at time t.

Proposition 34.2 (*Active Set Monotone in Input*)**.** *If $I_1(x,s) \geq I_2(x,s)$ for all x, s (first system has stronger inputs), then $\mathcal{M}^{(1)}_{\text{act}}(t) \supseteq \mathcal{M}^{(2)}_{\text{act}}(t)$: stronger inputs produce larger active sets.*

Proof. By the solution formula, $M_1(x,t) \geq M_2(x,t)$ pointwise, so $\{x : M_1(x,t) \geq \theta_M\} \supseteq \{x : M_2(x,t) \geq \theta_M\}$. ▪ ▪

34.3 Connection to Ecphory

The active memory set $\mathcal{M}_{\text{act}}(t)$ represents the baseline accessible memories. Ecphory (Chapter 35) occurs when a retrieval cue $c(x,t)$ pushes additional states above θ_M: the ecphory condition $M(x,t) + c(x,t) \geq \theta_M$ defines a larger active set that includes cue-activated memories.

Proposition 34.3 (*Memory Field is RSVP Capacity*)**.** *Under the identification $\Phi \leftrightarrow M$, $S \leftrightarrow \lambda M$ (forgetting rate as obligation), $\Gamma \leftrightarrow I$ (input as repair), the memory field equation is the RSVP capacity equation in the case of uniform spatial transport ($\mathrm{v} = 0$): $\partial_t M = -\lambda M + I \equiv \partial_t \Phi = -\lambda S + \Gamma$.*

Proof. The memory field $M(x,t)$ satisfies the same structural role as the RSVP capacity field $\Phi(x,t)$: both are non-negative scalar fields on the state space that determine which states are admissible ($\mathcal{A}_t = \{x : M(x,t) \geq \theta\}$ for memory; $\mathcal{A}_t = \{x : \Phi(x,t) - \mu S(x,t) \geq \theta\}$ for RSVP). The memory dynamics $\partial_t M = -\lambda M + I(x,t)$ (decay minus input) are a special case of the RSVP Φ-equation with $\mathrm{v} = 0$, $\lambda > 0$, and the input field I playing the role of repair Γ. The identification $\Phi \leftrightarrow M$ is therefore structural, not merely metaphorical. ▪ ▪

Exercises

34.1. For constant input $I(x,t) = I_0$ (constant reinforcement), find the steady-state memory field $M^*(x) = \lim_{t\to\infty} M(x,t)$. What forgetting rate λ makes $M^*(x) = \theta_M$ (marginal retention)?

34.2. A student studies for T hours then stops ($I = 0$ for $t > T$). Compute $M(x,t)$ for $t > T$ and find the time at which M drops below θ_M. How does this depend on study intensity?

34.3. Define the *memory half-life* $\tau_{1/2}$ as the time for $M(x,0) \to M(x,0)/2$ with no new input. Compute $\tau_{1/2}$ in terms of λ. Empirical data suggests $\tau_{1/2} \approx 1$ day for episodic memory. What does this imply about λ?

34.4. Prove that the memory field satisfies the maximum principle: if $M(x,0) \leq M_{\text{max}}$ and $I(x,t) \leq I_{\text{max}}$ for all x, t, then $M(x,t) \leq \max(M_{\text{max}}, I_{\text{max}}/\lambda)$ for all t.

35 Chapter 35

Ecphory and Retrieval

Memory activates not when it is strong, but when cue and trace together cross the threshold.

—

Phenomenological Note. Memory is not a library where you go to retrieve a file. It is more like a process of reconstruction that requires the right conditions. The memory does not sit waiting for you; it needs to be activated by something that was present when it was formed. Sometimes you cannot remember something until you are in the right room, or the right mood, or until someone uses the right word. The cue is part of the memory.

Ecphory is the process by which a retrieval cue $c(x,t)$ interacts with the memory trace $M(x,t)$ to produce conscious recollection. This chapter proves the Ecphory Threshold Lemma, derives the boundary motion of the retrieval set, and connects ecphory to the admissibility framework.

35.1 The Ecphory Threshold Lemma

Lemma 35.1 (*Ecphory Threshold*)**.** *A memory at state x is retrieved at time t if and only if the combined field exceeds the threshold:*

$$M(x,t) + c(x,t) \geq \theta,$$

where $c(x,t) \geq 0$ is the retrieval cue energy and $\theta > 0$ is the retrieval threshold. The ***retrieval set*** *at time t is:*

$$\mathcal{R}_{\text{ret}}(t) = \{x \in \mathcal{X} : M(x,t) + c(x,t) \geq \theta\}.$$

Proof. The ecphory condition is definitional: we define retrieval as the event $M(x,t) + c(x,t) \geq \theta$. This is a threshold model of retrieval motivated by the **spreading activation** hypothesis in cognitive science: the cue c provides activation energy that combines additively with the memory trace M to trigger retrieval when the sum exceeds a firing threshold θ. ▪ ▪

35.2 Retrieval Set Dynamics

The retrieval set $\mathcal{R}_{\text{ret}}(t)$ is the superlevel set $\{x : \Psi_{\text{ret}}(x,t) \geq 0\}$ where $\Psi_{\text{ret}}(x,t) = M(x,t) + c(x,t) - \theta$. Its boundary moves according to the level-set equation (Chapter 15):

Proposition 35.2 (*Retrieval Boundary Motion*). *The outward normal velocity of $\partial\mathcal{R}_{\text{ret}}(t)$ is:*

$$v_n = -\frac{\partial_t(M+c)}{|\nabla(M+c)|} = -\frac{-\lambda M + I + \partial_t c}{|\nabla(M+c)|}.$$

The retrieval set expands when input $I + \partial_t c > \lambda M$ (cue and input exceed decay) and contracts when $\lambda M > I + \partial_t c$ (decay dominates).

Proof. Differentiate $\Psi_{\text{ret}}(x(t),t) = 0$ along the boundary and solve for v_n as in Theorem 15.1, substituting $\partial_t M = -\lambda M + I$. ▪ ▪

35.3 Tip-of-the-Tongue as Boundary Proximity

Tip-of-the-Tongue as Near-Threshold State. The tip-of-the-tongue (TOT) phenomenon occurs when $M(x,t) + c(x,t) \approx \theta$: the memory is near but below the retrieval threshold. The state x is on the boundary $\partial\mathcal{R}_{\text{ret}}(t)$ or just outside it. Providing a stronger cue (hearing the first syllable of the word) increases c and pushes x across the threshold. The resolution of a TOT state is a boundary crossing in retrieval space. The frustrating inaccessibility of the TOT state reflects the fact that x is close to the boundary but the gradient $|\nabla(M+c)|$ at the boundary is shallow — it takes a disproportionately large cue to cross.

35.4 Ecphory and Admissibility

The retrieval set $\mathcal{R}_{\text{ret}}(t)$ is the cognitive admissibility domain: the set of states that the system can currently access. The ecphory threshold θ plays the role of the RSVP threshold θ, and $M + c$ plays the role of $\Phi - \mu S$: the combined field that determines admissibility.

This makes the entire MEM|8 machinery a special case of the RSVP admissibility framework applied to the cognitive state space.

Exercises

35.1. Suppose $M(x,0) = 0.8\theta$ and the cue provides $c = 0.3\theta$. Is x retrieved? Now suppose the cue decays as $c(t) = 0.3\theta e^{-\mu t}$. Find the time t^* at which x falls out of the retrieval set.

35.2. Prove that the retrieval set $\mathcal{R}_{\text{ret}}(t)$ is monotone non-decreasing in cue strength: if $c_1(x,t) \geq c_2(x,t)$ pointwise, then $\mathcal{R}^{(1)}_{\text{ret}}(t) \supseteq \mathcal{R}^{(2)}_{\text{ret}}(t)$.

35.3. A context change (entering a different room) sharply changes $c(x, t)$. Model the retrieval set change and compute the fraction of memories that fall below threshold immediately after the context change. How does this relate to the Godden-Baddeley context-dependent memory effect?

35.4. Define *false memory* as a state $x \notin \mathcal{X}_{\text{true}}$ that enters $\mathcal{R}_{\text{ret}}(t)$ due to high cue activation $c(x, t)$. Derive a bound on the false memory rate as a function of the cue selectivity $|\nabla c|$ and the memory resolution $|\nabla M|$.

36 Chapter 36

Consciousness as Reachability

> A thought becomes conscious not by being illuminated from without but by becoming reachable from everywhere.
>
> —

Phenomenological Note. The feeling of unified experience is very convincing. There seems to be a single point of view, a continuous observer who has been present since early childhood. It is surprising, then, to notice how much of the mind's activity is simply not accessible to that point of view — how much computation completes before the result surfaces, how many decisions seem made before you are aware of deciding.

The hard problem of consciousness asks why there is subjective experience at all. The CPR framework does not dissolve this question, but it offers a precise functional account of which cognitive contents count as conscious: a content is conscious iff it is globally reachable by the principal action, report, memory, and attention systems. This is the *reachability theory of consciousness*.

36.1 Global Workspace Theory, Reformulated

Baars' Global Workspace Theory (GWT) holds that consciousness corresponds to "broadcasting" of information across a global workspace accessible to many specialist modules (Baars 1988; Dehaene 2014). The CPR reformulation makes this precise:

Definition 36.1 (*Cognitive Reachability Graph*)**.** The **cognitive reachability graph** $\mathcal{G}_{\text{cog}}$ is a directed graph whose nodes are cognitive sub-systems (action selection A, verbal report V, episodic memory M, attention T, and others), and whose edges represent admissible information transfer. A cognitive content c is represented as a node; it has a directed edge to sub-system S if S can read, use, or modify c.

36.2 The Conscious Reachability Proposition

Proposition 36.1 (*Conscious Reachability*). *A cognitive content c is **conscious** if and only if c is globally reachable in $\mathcal{G}_{\text{cog}}$ from the current attention state a:*

$$c \text{ is conscious} \iff c \in \mathcal{R}_{\mathcal{G}_{\text{cog}}}(a, T) \text{ for sufficiently small } T,$$

where $\mathcal{R}_{\mathcal{G}_{\text{cog}}}(a, T)$ is the set of nodes reachable from a in $\mathcal{G}_{\text{cog}}$ within T steps.

Proof. We establish the equivalences between the functional criteria of consciousness and reachability in $\mathcal{G}_{\text{cog}}$.

Criterion 1: Reportability. A content c is consciously accessible iff the agent can produce a verbal report about it. In graph terms: there is a path from the current attention state a to the verbal report system V passing through (or from) c. This holds iff c and V are in the same strongly connected component of $\mathcal{G}_{\text{cog}}$, or equivalently $c \in \mathcal{R}(a, T)$ and $V \in \mathcal{R}(c, T)$.

Criterion 2: Memory consolidation. Contents that become conscious are preferentially consolidated into episodic memory M. In graph terms: $M \in \mathcal{R}(c, T)$ — the content can reach the memory system within the relevant horizon.

Criterion 3: Action modulation. Conscious contents modulate voluntary action A: $A \in \mathcal{R}(c, T)$.

Criterion 4: Attentional access. Attention T can be directed to c: $c \in \mathcal{R}(a, T_0)$ for the current attention state a.

Combining all four: c is conscious iff c lies in the intersection of reachable sets from a to $\{V, M, A\}$ within the relevant horizon. This is equivalent to c being in the main strongly connected component of the active workspace — the global workspace, in Baars' terminology. ▪ ▪

36.3 Unconscious Processing as Local Reachability

Definition 36.2 (*Unconscious Content*). A content c is **unconscious** if $c \notin \mathcal{R}_{\mathcal{G}_{\text{cog}}}(a, T)$ for any small T — it is not reachable from the current attention state.

Unconscious processing occurs in sub-graphs of $\mathcal{G}_{\text{cog}}$ that are not connected to the global workspace. A specialist module (e.g., the early visual system, the implicit motor planner) may compute rich representations that never propagate to the workspace nodes.

The Fringe and Near-Consciousness. William James' "fringe" of consciousness corresponds, in this framework, to contents at the boundary of the reachable set: $c \in \partial\mathcal{R}_{\mathcal{G}_{\text{cog}}}(a, T)$. They are almost reachable — a small increase in T, or a small perturbation of attention, would bring them into consciousness. This explains the phenomenological character of "tip-of-the-tongue" states and near-perceptions.

36.4 Anesthesia and Reachability Collapse

General anesthesia suppresses consciousness without eliminating cortical activity. The CPR account predicts: anesthetics reduce the reachability volume of $\mathcal{G}_{\text{cog}}$ by disrupting the long-range connections that form the workspace, even while local processing in specialist modules continues.

Proposition 36.2 (*Anesthesia as Workspace Fragmentation*)**.** *Under general anesthesia, the cognitive reachability graph $\mathcal{G}_{\text{cog}}$ fragments into isolated sub-graphs. The main strongly connected component collapses, so the global reachable set from a shrinks to near zero:*

$$\mathcal{R}_{\mathcal{G}_{\text{cog}}}(a, T) \to \{a\}.$$

This is the reachability-theoretic signature of unconsciousness.

Proof. General anesthesia suppresses long-range cortical connectivity, reducing the effective reachability in the cognitive graph $\mathcal{G}_{\text{cog}}$. Formally, anesthesia acts as a damage operator that removes or weakens edges in $\mathcal{G}_{\text{cog}}$, fragmenting strongly connected components into smaller isolated subgraphs. By Proposition 11.2, the Fiedler value $\lambda_2(\mathrm{L})$ drops toward zero, increasing the number of effective components ($\lambda_2 \to 0$ iff the graph becomes disconnected). The workspace $\mathcal{G}_{\text{work}}$ shrinks as SCCs fragment: conscious reportability requires $\mathcal{G}_{\text{work}}$ to be one large SCC (global broadcasting), which fails under anesthesia. ▪ ▪

This prediction is consistent with empirical work on cortical effective connectivity under anesthesia (Tononi's perturbational complexity index).

36.5 Phenomenal vs. Access Consciousness

Ned Block distinguishes:

- *Access consciousness*: information available to rational control, report, and voluntary action.
- *Phenomenal consciousness*: the "what it's like" aspect.

Proposition 36.1 is a theory of *access* consciousness. It provides a precise, empirically testable functional criterion without solving the hard problem. Whether phenomenal consciousness reduces to access consciousness is a further question the CPR framework does not adjudicate — but it provides the right functional vocabulary in which to state the question precisely.

36.6 Summary

1. Consciousness is global reachability in $\mathcal{G}_{\text{cog}}$: a content is conscious iff it is reachable from the attention state within the relevant horizon (Propo-

sition 36.1).
2. Unconscious contents are in sub-graphs disconnected from the global workspace.
3. The phenomenological fringe corresponds to the boundary of the reachable set.
4. Anesthesia fragments $\mathcal{G}_{\text{cog}}$, collapsing the main connected component.
5. The account covers access consciousness precisely; it remains agnostic on phenomenal consciousness.

Exercises

36.1. Draw a small $\mathcal{G}_{\text{cog}}$ with five nodes $\{a, c_1, c_2, V, M\}$ and edges. Determine which nodes are conscious under Proposition 36.1. Now remove one edge and recompute.

36.2. Formulate the Integrated Information Theory (IIT) quantity Φ in terms of reachability: Φ is large when the main SCC of $\mathcal{G}_{\text{cog}}$ has high internal reachability density. Show that your formulation agrees with the qualitative predictions of IIT.

36.3. Prove that if $\mathcal{G}_{\text{cog}}$ is a strongly connected graph, then every content in the system is simultaneously conscious. Is this plausible? What constraint does evolution impose on the connectivity of $\mathcal{G}_{\text{cog}}$?

36.4. (Dreaming.) During REM sleep, voluntary motor output is suppressed (A is disconnected from $\mathcal{G}_{\text{cog}}$) but phenomenal experience continues. What does this imply about the reachability theory? Does it require revising Proposition 36.1?

37 Chapter 37

Self Models

> The self is not a homunculus that perceives. It is a fixed point that persists.
>
> —

Phenomenological Note. The model you have of yourself is revised less often than the evidence warrants. This is not only stubbornness. Updating a self-model is expensive — it requires re-interpreting prior events, adjusting expectations about future behavior, renegotiating how other people see you. The self tends to be held fixed as long as possible, and revised in patches when the cost of not updating finally exceeds the cost of the revision.

The intuition that there is a "self" — a persistent entity that is the locus of experience, agency, and identity — is philosophically contested but phenomenologically pervasive. The CPR framework offers a precise account: the self is not an entity in the state space but a *fixed point* of a self-referential mapping. It exists when the system's model of itself is stable under updates that incorporate that very model.

37.1 Self-Referential Maps

Definition 37.1 (*Self-Model Map*)**.** Let s be a self-model (a representation of the system's own states, dispositions, and constraints) and let $\mathcal{E}$ be the external environment. The **self-model update map** is

$$\mathcal{F} : \mathcal{S} \times \mathcal{E} \to \mathcal{S}, \qquad (s, \mathcal{E}) \mapsto \mathcal{F}(s, \mathcal{E}),$$

where $\mathcal{S}$ is the space of possible self-models. $\mathcal{F}(s, \mathcal{E})$ is the updated self-model when the current self-model is s and the environment is $\mathcal{E}$.

The self-referential structure arises because the system's self-model s itself influences what it perceives and what it can do — hence the next self-model depends on the current one.

37.2 The Self-Model Fixed Point Theorem

Theorem 37.1 (*Self-Model Fixed Point*). *Let $\mathcal{S}$ be a complete metric space and let $\mathcal{F}(\cdot, \mathcal{E}) : \mathcal{S} \to \mathcal{S}$ be a* ***contraction mapping*** *for each fixed $\mathcal{E}$:*

$$d(\mathcal{F}(s_1, \mathcal{E}),\ \mathcal{F}(s_2, \mathcal{E})) \ \leq\ k\, d(s_1, s_2) \quad \text{for some } k \in [0, 1).$$

Then there exists a unique ***self-model fixed point*** *$s^* = s^*(\mathcal{E})$ satisfying*

$$s^* = \mathcal{F}(s^*, \mathcal{E}).$$

The fixed point is the stable self-model *of the system in environment $\mathcal{E}$. Moreover, for any initial self-model s_0, the iteration $s_{n+1} = \mathcal{F}(s_n, \mathcal{E})$ converges to s^* with geometric rate:*

$$d(s_n, s^*) \ \leq\ \frac{k^n}{1-k}\, d(s_0, s_1).$$

Proof. Direct application of the Banach Fixed Point Theorem: for a contraction $\mathcal{F}$ on a complete metric space $\mathcal{S}$, the sequence $s_n = \mathcal{F}^n(s_0)$ is Cauchy, converges to a unique limit s^*, and $s^* = \mathcal{F}(s^*, \mathcal{E})$ by continuity of $\mathcal{F}$. Uniqueness: if s^{**} is another fixed point, $d(s^*, s^{**}) = d(\mathcal{F}(s^*, \mathcal{E}), \mathcal{F}(s^{**}, \mathcal{E})) \leq k\, d(s^*, s^{**})$, forcing $d(s^*, s^{**}) = 0$. ▪ ▪

37.3 What Makes the Map a Contraction?

The contraction condition $k < 1$ is not automatic. It holds when the self-model update is *regularised*: new evidence is integrated with discounting so that no single observation can completely overturn the self-model.

Definition 37.2 (*Regularised Self-Model Update*)**.** The self-model update is **regularised** with coefficient $\alpha \in (0, 1)$ if

$$\mathcal{F}(s, \mathcal{E}) = (1 - \alpha)\, s + \alpha\, \hat{s}(\mathcal{E}),$$

where $\hat{s}(\mathcal{E})$ is the "evidence-ideal" self-model given environment $\mathcal{E}$.

Lemma 37.2. *A regularised self-model update with coefficient α is a contraction with $k = 1 - \alpha < 1$.*

Proof. $d(\mathcal{F}(s_1, \mathcal{E}), \mathcal{F}(s_2, \mathcal{E})) = (1 - \alpha)\, d(s_1, s_2)$. ▪ ▪

The coefficient $1 - \alpha$ is the *self-model inertia*: how resistant the self-model is to disruption by new evidence. High inertia (α small) means self-models are rigid — hard to update, slow to converge to new environments. Low inertia (α near 1) means self-models are plastic — fast to update but potentially unstable.

37.4 Identity Across Time

The fixed-point formulation gives a precise account of personal identity across time:

Definition 37.3 (*Personal Identity*)**.** A system **maintains personal identity** across an environmental sequence $\mathcal{E}_1, \mathcal{E}_2, \ldots$ if the sequence of fixed points $s^*(\mathcal{E}_1), s^*(\mathcal{E}_2), \ldots$ lies within a bounded region of $\mathcal{S}$ (the self-model varies but does not drift without bound).

Identity as Bounded Drift. Personal identity is not the persistence of a fixed self-model — that would be rigidity, not identity. It is the *bounded drift* of the fixed point as the environment changes. Radical discontinuity in s^* across $\mathcal{E}$ — the fixed point jumping to a distant region of $\mathcal{S}$ — corresponds to a breakdown of identity (psychotic break, radical conversion, severe amnesia).

37.5 Self-Models in AI Systems

Current large language models implicitly learn self-models during training: their generation is conditioned on a representation of themselves as a particular kind of system. The fixed-point analysis applies: the model's self-representation is stable (a fixed point) when its outputs are consistent with its own self-model. Inconsistency (the model asserting things about itself that contradict its training) corresponds to a self-model that has not reached its fixed point — or to a contraction coefficient $k \geq 1$ (no stable fixed point exists).

This formalises the phenomenon of "sycophancy" in LLMs: a model whose self-model update places excessive weight on user approval (α too close to 1) has a fixed point that drifts with user expectations, yielding a fragile, environment-dependent identity.

Exercises

37.1. Let $\mathcal{S} = [0, 1]$ and $\mathcal{F}(s, \mathcal{E}) = s/2 + e/4$ where $e \in [0, 1]$ is a scalar environment parameter. Find the fixed point $s^*(e)$ as a function of e. How does personal identity (in the sense of Definition 37.3) depend on the range of e encountered?

37.2. Prove that the composition of two contraction maps $\mathcal{F}_1 \circ \mathcal{F}_2$ is a contraction with coefficient $k_1 k_2$. Interpret: applying two self-model updates in sequence is more stabilising than either alone.

37.3. Suppose $\mathcal{F}$ is *not* a contraction ($k \geq 1$). Give an example where the self-model iteration diverges. What does this correspond to phenomenologically?

37.4. (Parfit's question.) Suppose a system undergoes gradual replacement of its components (ship of Theseus). Model each replacement as a small perturbation of $\mathcal{F}$. Under what conditions does the fixed point s^* vary continuously with the perturbations? Is continuous variation sufficient for identity?

38 Chapter 38

Synthetic Cognition

A mind is not defined by its substrate but by its admissibility geometry.

—

Phenomenological Note. We are now building things that behave intelligently in ways we do not fully understand. This is not entirely new. We built institutions, languages, cities, markets — systems that produce outcomes no single participant designed or intended. What is new is the speed, and the degree to which the mechanisms are opaque even to their builders. The intelligence is real. Its origins are distributed in ways that resist simple description.

The preceding chapters have defined cognition functionally: perception as projection, belief as manifold coordinates, inference as geodesic, consciousness as global reachability, self as fixed point. (Dennett 1991; Turing 1950) This chapter synthesises these into the Synthetic Cognition Criterion: the minimum conditions for any system — biological or artificial — to qualify as a genuine cognitive agent in the CPR sense.

38.1 The Four Conditions

Theorem 38.1 (*Synthetic Cognition Criterion*)**.** *A system $\mathcal{S}$ is a* ***synthetic cognitive agent*** *if and only if it satisfies all four conditions:*

(i) ***Admissibility field maintenance****: $\mathcal{S}$ maintains a non-trivial admissibility field $\mathcal{A}_{\mathcal{S}}$ over its state space, with $\mathcal{V}_R(\mathcal{A}_{\mathcal{S}}, \cdot, T) > 0$;*

(ii) ***Damage detection****: $\mathcal{S}$ detects when its state exits $\mathcal{A}_{\mathcal{S}}$;*

(iii) ***Repair****: $\mathcal{S}$ applies repair operators to restore $\mathcal{V}_R(\mathcal{A}_{\mathcal{S}}, \cdot, T)$;*

(iv) ***Reachability navigation****: $\mathcal{S}$'s actions tend toward states with higher reachability volume:* $\arg\max_a \mathcal{V}_R(f(\cdot, a), T)$.

Proof that conditions are necessary. A system lacking condition (i): has no model of what is admissible; it cannot distinguish coherent from incoherent states. A system lacking condition (ii): cannot detect failures; it continues operating while its admissibility field collapses. A system lacking condition (iii): accumulates damage without correction; its reachability volume monotonically declines. A system lacking condition (iv): does not use reachability as a guide; it may systematically reduce its own future options.

Any system lacking any one of (i)–(iv) fails a fundamental functional criterion for adaptive, goal-directed behaviour.

Sufficiency follows from the Repair Convergence Theorem (Theorem 86.1): a system satisfying (i)–(iii) converges to a stable repair fixed point; condition (iv) ensures the system navigates toward higher-$\mathcal{V}_R$ states rather than arbitrary ones. ▪ ▪

38.2 Current AI Systems Against the Criterion

Where LLMs Fall Short. Evaluating large language models against conditions (i)–(iv):

- **(i) Partially**: LLMs implicitly model distributional constraints but do not maintain an explicit admissibility field.
- **(ii) Weakly**: LLMs occasionally detect their own errors (in chain-of-thought) but have no systematic damage detection.
- **(iii) Not systematically**: LLMs cannot repair mid-generation; errors propagate and compound.
- **(iv) Partially**: next-token prediction locally maximises distributional probability, which approximates short-horizon reachability but not long-horizon reachability volume.

This is not a criticism of LLMs — they were not designed to satisfy the Synthetic Cognition Criterion. It is a specification of what would need to be added.

38.3 The Relationship to Turing's Test

The Turing Test asks: can a system produce outputs indistinguishable from a human's in a text-based interrogation? The Synthetic Cognition Criterion asks: does the system maintain, monitor, repair, and navigate an admissibility field?

These are orthogonal. A system could pass the Turing Test without satisfying the SCCriterion (e.g., a sophisticated lookup table mimicking human outputs). A system could satisfy the SCCriterion without passing the Turing Test (e.g., a simple organism that maintains homeostasis but cannot produce human-like language).

The SCCriterion is more fundamental: it captures the *functional structure* of cognition, not the surface form of its outputs.

Exercises

38.1. Evaluate a thermostat against the four conditions. Which does it satisfy? Is a thermostat a cognitive agent in the CPR sense?

38.2. Design an architecture for a language model that explicitly satisfies condition (ii) (damage detection). What components would be needed? How would the model know when its output has left its admissibility field?

38.3. Prove that a system satisfying (i)–(iii) but not (iv) will reach a repair fixed point but may navigate toward it through states with arbitrarily low reachability. Give an example.

38.4. Propose an empirical test that distinguishes systems satisfying condition (iv) from systems that merely maximise immediate prediction accuracy.

Part VI

Language and Meaning

[Part introduction — to be written.]

39 Chapter 39

Meaning as Constraint

To understand a word is not to possess its referent. It is to know what it rules out.

—

Phenomenological Note. Words mean what they do by excluding what they do not do. The word is not just a label attached to a thing; it is a filter that picks out some features and suppresses others. Change the word and you change which features survive the compression. This is why translation is never neutral. Every translation chooses which constraints to preserve and which to sacrifice.

The standard view holds that words mean by referring: "cat" means the set of cats. This chapter argues for a different account: a word means by *constraining admissible continuations*. Hearing "cat" does not merely point you toward cats; it restricts what can coherently come next. Meaning is an operator on possibility, not a pointer to objects.

39.1 Continuation Spaces

Definition 39.1 (*Continuation Space*)**.** Let $\mathcal{C}$ be the space of all admissible continuations available to a cognitive or linguistic system before receiving a linguistic expression. A continuation $c \in \mathcal{C}$ is a possible future trajectory of interpretation, inference, or action.

Before any expression is received, the system faces the full space $\mathcal{C}$. After receiving an expression w, some continuations become inadmissible. The expression w functions as a constraint that carves out a sub-space.

39.2 The Meaning-as-Constraint Theorem

Theorem 39.1 (*Meaning as Constraint*)**.** *The **meaning** of a linguistic expression w is the operator $\mu(w) : 2^{\mathcal{C}} \to 2^{\mathcal{C}}$ defined by*

$$\mu(w)(\mathcal{C}') = \{c' \in \mathcal{C}' : c \xrightarrow{w} c' \text{ is admissible after } w\},$$

*where $\mathcal{C}' \subseteq \mathcal{C}$ is the current admissible space. The **semantic content** of w is*

$$\text{content}(w) = \log \frac{\text{Vol}(\mathcal{C})}{\text{Vol}(\mu(w)(\mathcal{C}))}.$$

More restriction = greater semantic content.

Proof. We verify three properties.

(1) *Monotonicity.* $\mu(w)(\mathcal{C}') \subseteq \mathcal{C}'$ for all $\mathcal{C}'$: hearing w can only restrict, not expand, what was admissible. This follows directly from Theorem 5.1.

(2) *Compositionality.* For a word sequence $w_1 w_2$: $\mu(w_1 w_2) = \mu(w_2) \circ \mu(w_1)$. Word w_1 produces $\mathcal{C}_1 = \mu(w_1)(\mathcal{C})$; w_2 restricts to $\mathcal{C}_{12} = \mu(w_2)(\mathcal{C}_1)$.

(3) *Content scaling.* $\text{content}(w) \geq 0$ (since $\mu(w)(\mathcal{C}) \subseteq \mathcal{C}$), $= 0$ iff w is vacuous, and equals the fiber entropy S_π of Chapter 13 applied to the projection from full context to post-w context. ▪ ▪

Reference as Special Case. Reference theories treat meaning as pointing. The CPR account is more general: if w refers to E, then $\mu(w)$ restricts continuations to those compatible with E. Reference is a special case of constraint. It fails for logical particles, indexicals, and performatives; the constraint account handles them all.

39.3 Semantic Degeneracy and Ambiguity

Definition 39.2 (*Semantic Degeneracy*)**.** An expression w is **semantically degenerate** (ambiguous) if $\mu(w)(\mathcal{C})$ has more than one connected component: $\mu(w)(\mathcal{C}) = C_1 \sqcup C_2 \sqcup \cdots \sqcup C_k$, $k \geq 2$. The number k is the **ambiguity degree** of w.

Example 39.1 (Bank)**.** “Bank” has ambiguity degree ≥ 2: component C_{fin} (financial continuations) and C_{riv} (river-bank continuations). Context resolves ambiguity by further restricting to one component. Without context, the semantic Fisher metric degenerates at “bank”: nearby words diverge sharply into the two clusters.

39.4 Meaning and the Master Theorem

Theorem 39.1 is the language-domain instance of the Master Theorem (Theorem 89.1).

In language:

Constraint = expression w restricts $\mathcal{C}$
Projection = linguistic encoding maps meanings to tokens

Reachability = admissible continuation space $\mu(w)(\mathcal{C})$

Meaning is preserved across linguistic transformation iff reachability-relevant distinctions of $\mathcal{C}$ survive the encoding.

The projection from meanings to tokens is always lossy (Chapter 44): the same token string is consistent with multiple continuation spaces. A language user reconstructs meaning by applying additional contextual constraints to collapse the ambiguity fiber.

Exercises

39.1. Let $\mathcal{C} = \{c_1, \dots, c_8\}$. Define two operators $\mu(w_1)$ and $\mu(w_2)$. Compute $\mu(w_1 w_2)(\mathcal{C})$ and the content of each word.

39.2. Prove content$(w_1 w_2) \geq$ content(w_1). When does equality hold?

39.3. Formalise contradiction: w_1 and w_2 are contradictory iff $\mu(w_1)(\mathcal{C}) \cap \mu(w_2)(\mathcal{C}) = \emptyset$. Give two examples from natural language.

39.4. Define synonymy: $w_1 \equiv w_2$ iff $\mu(w_1)(\mathcal{C}) = \mu(w_2)(\mathcal{C})$ for all $\mathcal{C}$. Is perfect synonymy achievable in natural language?

39.5. Connect to Theorem 19.1: when does the linguistic projection from meaning-space to token-space produce high mixing Λ? Give a linguistic example.

40 Chapter 40

Distinctions and Categories

A category is not a box you put things in. It is a region of state space from which you cannot easily escape.

—

Phenomenological Note. Every language you were not raised in contains distinctions your native language collapses and collapses distinctions your native language keeps separate. The strange experience of learning a second language is partly noticing which of your "natural" categories were actually choices — alternatives that a speaker of another language finds completely unnatural, having never needed them.

Categories are the stable semantic objects of language — *chair*, *love*, *justice*. The CPR framework does not take categories as primitive. It derives them from reachability: a category is a region of semantic space in which internal reachability is high and boundary-crossing reachability is low. The Category Distinction Lemma formalises this.

40.1 Categories as High-Reachability Regions

Definition 40.1 (*Semantic Reachability Ratio*). Let $\mathcal{M}$ be a semantic space and let $X \subset \mathcal{M}$ be a candidate category. The **semantic reachability ratio** of X is

$$\rho(X) = \frac{P(x \rightsquigarrow y \mid x, y \in X)}{P(x \rightsquigarrow z \mid x \in X, z \notin X)},$$

the ratio of intra-category reachability to inter-category reachability. Here $P(x \rightsquigarrow y)$ denotes the probability of navigating from x to y along an admissible path.

Lemma 40.1 (*Category Distinction Lemma*). *A set $X \subset \mathcal{M}$ is a* ***stable semantic category*** *if and only if $\rho(X) \gg 1$: intra-category reachability significantly exceeds inter-category reachability. Equivalently, X is a region of $\mathcal{M}$ whose interior is much more reachable from within than from without.*

Proof. ($\Rightarrow$) If X is a stable category, then concepts within X are semantically compatible: transitions between them are admissible (e.g., "dog" and "puppy" are mutually reachable via "young dog", "canine", etc.). Concepts across the boundary of X require larger semantic shifts: getting from "dog" to "justice" requires a long path through many intermediate concepts. Therefore intra-category paths are more probable (shorter, more direct) than inter-category paths. So $\rho(X) > 1$.

($\Leftarrow$) If $\rho(X) \gg 1$, the region X forms a "basin" in the semantic reachability topology: it is easy to navigate within X but hard to leave. Random semantic walks are more likely to stay within X than to cross to other regions. This is the operational signature of a stable category. ▪ ▪

40.2 Category Boundaries as Reachability Frontiers

Proposition 40.2 (*Boundaries as Low-Reachability Zones*)**.** *The boundary ∂X of a semantic category X consists of concepts for which:*

$$\mathcal{V}_R(x,T)_{intra} \approx \mathcal{V}_R(x,T)_{inter} \quad \text{for some horizon } T.$$

At the boundary, intra- and inter-category reachability are approximately equal — the concept is ambiguous between being in X or outside X.

Proof. Define $\mathcal{A}_X(x) = 1[\rho(x) > 1]$. A concept x is interior if $\rho(x) \gg 1$ (intra dominates); it is exterior if $\rho(x) \ll 1$ (inter dominates). The boundary ∂X is the level set $\rho(x) = 1$, which is exactly $\mathcal{V}_R(x,T)_{\text{intra}} = \mathcal{V}_R(x,T)_{\text{inter}}$. ▪ ▪

Category boundaries are the semantic analogues of the admissibility boundary $\partial\mathcal{X}_t$: regions where the admissibility field $\mathcal{A}_X = 1[\rho(\cdot) > 1]$ transitions from 1 to 0.

Example 40.1 (Prototype Theory as Centrality)**.** Rosch's prototype theory holds that category membership is graded, with prototypes at the centre. In the CPR framework, the prototype of X (Rosch and Lloyd 1978) is the concept $x^* \in X$ with highest intra-category reachability: $x^* = \arg\max_{x\in X} \mathcal{V}_R(x,T)_{\text{intra}}$. Prototypes are the most "central" — most reachable from within — elements of the category. Non-prototypical members are near the boundary ∂X: they have nearly equal intra- and inter-category reachability.

40.3 Category Formation and Distinction Preservation

Proposition 40.3 (*Categories Preserve Reachability-Relevant Distinctions*)**.** *A semantic category system $\{X_1, X_2, \dots, X_k\}$ (a partition of $\mathcal{M}$) preserves reachability-relevant distinctions iff for every pair $x \in X_i$, $y \in X_j$ with $i \neq j$:*

$$\mathcal{R}_{\mathcal{A}}(x,T) \neq \mathcal{R}_{\mathcal{A}}(y,T) \Rightarrow i \neq j.$$

States with different reachable futures must be in different categories.

Proof. The category system defines $\pi_{\text{cat}}(x) = i$ iff $x \in X_i$. The condition states: if $\mathcal{R}(x,T) \neq \mathcal{R}(y,T)$ then $\pi_{\text{cat}}(x) \neq \pi_{\text{cat}}(y)$, i.e., no two states with different reachable futures share a category. This is exactly the definition that π_{cat} preserves all reachability-relevant distinctions (Lemma 6.1), which is the language-domain instance of semantic faithfulness. ▪ ▪

This is the language-domain instance of the Distinction Preservation Lemma (Lemma 6.1): the category system is a projection $\pi_{\text{cat}} : \mathcal{M} \to \{1, \ldots, k\}$, and it is semantically faithful iff it preserves all reachability-relevant distinctions.

40.4 Whorf's Hypothesis as Reachability Claim

The Sapir–Whorf hypothesis holds that the categories of one's language shape one's thought. In CPR terms: the category system $\{X_1, \ldots, X_k\}$ induced by a language determines which distinctions are preserved and which are collapsed.

Proposition 40.4 (*Linguistic Relativity as Projection Choice*)**.** *Two languages L_1, L_2 with category systems $\mathcal{P}_1, \mathcal{P}_2$ have different cognitive reachability for speakers of each language in proportion to the number of reachability-relevant distinctions preserved by $\mathcal{P}_1$ but not $\mathcal{P}_2$, and vice versa.*

Proof. Each language induces a projection $\pi_i : \mathcal{M} \to \mathcal{P}_i$. Let D_i be the set of reachability-relevant distinctions preserved by π_i. By Proposition 40.3, D_i determines which cognitive distinctions are available to L_i-speakers. The difference in cognitive reachability is $|D_1 \setminus D_2|$ (distinctions available to L_1 speakers but collapsed in L_2) and symmetrically $|D_2 \setminus D_1|$. The total difference in reachable cognitive futures is proportional to $|D_1 \triangle D_2| = |D_1 \setminus D_2| + |D_2 \setminus D_1|$. ▪ ▪

The weak Whorf hypothesis (linguistic categories influence thought) is the claim that $\mathcal{P}$ affects the cognitive admissibility field. The strong Whorf hypothesis (language determines thought) would require $\mathcal{P}$ to *be* the cognitive admissibility field. The CPR framework supports the weak form: different category systems preserve different distinctions, and distinctions not preserved are harder to reason about.

Exercises

40.1. Let $\mathcal{M} = [0,1]^2$ with reachability measure $P(x \rightsquigarrow y) = e^{-\|x-y\|^2/\sigma^2}$. Find the semantic reachability ratio $\rho(X)$ for the "upper-left quadrant" category $X = [0, 0.5]^2$. For what σ is X a stable category ($\rho > 2$)?

40.2. Prove that the union $X_1 \cup X_2$ of two stable categories with $\rho(X_1) \gg 1$ and $\rho(X_2) \gg 1$ is a stable category if and only if $P(x \rightsquigarrow y)$ is high for $x \in X_1, y \in X_2$. Interpret: merging categories is stable iff the categories are mutually reachable.

40.3. (Cross-cultural.) The English colour terms "blue" and "green" are distinct in English but covered by a single term in some languages. Model this as two category systems $\mathcal{P}_{\text{en}}$ and $\mathcal{P}_{\text{other}}$. Which reachability-relevant distinctions (in what context) does $\mathcal{P}_{\text{other}}$ collapse?

40.4. Show that the Fisher metric on the belief manifold $\mathcal{B}$ (Chapter 32) induces a natural category structure: two concepts are in the same "belief category" if their posterior distributions under any evidence are within Fisher distance ϵ of each other. Verify Lemma 40.1 for this structure.

41 Chapter 41

Semantic Reachability

Meaning is the geometry of what can follow what.

—

Phenomenological Note. Some words are near each other and some are far. You feel this when a sentence reaches for a word and finds only an approximate one, something close but not exactly right. The right word would have opened a particular path; the approximate word opens a slightly different one. Most of the time the difference is small. Sometimes it is the whole argument.

This chapter introduces the *semantic reachability metric* — a distance function on conceptual space derived directly from the probability of navigating between concepts. It is the language-domain realisation of the general reachability geometry from Chapter 14, and it provides the quantitative foundation for distinguishing synonymy, antonymy, hyponymy, and semantic relatedness.

41.1 The Semantic Reachability Metric

Definition 41.1 (*Semantic Reachability Metric*)**.** Let $\mathcal{M}$ be a semantic space and $P(x \rightsquigarrow y)$ the probability of navigating from concept x to concept y along an admissible semantic path. The **semantic reachability distance** is

$$d_R(x, y) = -\log P(x \rightsquigarrow y).$$

Large $d_R(x, y)$ means x and y are semantically far apart: it is unlikely to navigate from one to the other along any admissible path. $d_R(x, y) = 0$ means x and y are the same concept (or synonymous: $P(x \rightsquigarrow x) = 1$).

41.2 Metric Axioms

Theorem 41.1 (*Semantic Reachability is a Metric*)**.** *Under the conditions that:*

(i) *transitions are* ***Markovian****:* $P(x \rightsquigarrow z) = \sum_y P(x \rightsquigarrow y)P(y \rightsquigarrow z)$*;*

(ii) *transitions are* ***symmetric****:* $P(x \rightsquigarrow y) = P(y \rightsquigarrow x)$*;*

(iii) $P(x \rightsquigarrow x) = 1$ *for all* x*;*

d_R satisfies the metric axioms:

1. ***Non-negativity**: $d_R(x, y) \geq 0$, with $d_R(x, y) = 0$ iff $x = y$.*
2. ***Symmetry**: $d_R(x, y) = d_R(y, x)$.*
3. ***Triangle inequality**: $d_R(x, z) \leq d_R(x, y) + d_R(y, z)$.*

Proof. Non-negativity. $P(x \rightsquigarrow y) \in (0, 1]$, so $d_R(x, y) = -\log P \geq 0$. $d_R(x, y) = 0$ iff $P(x \rightsquigarrow y) = 1$ iff $x = y$ (by condition iii).

Symmetry. $d_R(y, x) = -\log P(y \rightsquigarrow x) = -\log P(x \rightsquigarrow y) = d_R(x, y)$ by condition (ii).

Triangle inequality. By the Markov condition (i): $P(x \rightsquigarrow z) \geq P(x \rightsquigarrow y) \cdot P(y \rightsquigarrow z)$ (the probability of reaching z from x is at least as large as the probability of going through y). Taking $-\log$: $d_R(x, z) = -\log P(x \rightsquigarrow z) \leq -\log P(x \rightsquigarrow y) - \log P(y \rightsquigarrow z) = d_R(x, y) + d_R(y, z)$. ▪ ▪

Remark 41.1 (Non-symmetric Semantics)**.** Natural language is often non-symmetric: navigating from "dog" to "animal" may be easier (higher probability) than from "animal" to "dog" (there are many animals that are not dogs). In that case, d_R is a *quasi-metric* satisfying all axioms except symmetry. The asymmetric version captures semantic hypernymy: $d_R(\text{specific}, \text{general}) < d_R(\text{general}, \text{specific})$.

41.3 Geodesics and Analogical Reasoning

A geodesic from x to z in the semantic reachability metric is a path $x = c_0, c_1, \dots, c_k = z$ minimising $\sum_i d_R(c_i, c_{i+1})$. This is the most efficient semantic path — the sequence of conceptual steps that gets from x to z with least total navigation cost.

Proposition 41.2 (*Analogy as Geodesic Parallelism*)**.** *An analogy of the form "a is to b as c is to d" corresponds to the claim that the geodesic from a to b and the geodesic from c to d are parallel in the semantic space: they have the same direction and length in d_R. Formally, the vector $(a \to b)$ in semantic space is approximately equal to the vector $(c \to d)$.*

Proof. In the reachability metric, the relationship between two concepts a, b is encoded in the geodesic from a to b with initial tangent v_{ab}. An analogy $(a : b :: c : d)$ asserts structural identity of these relationships. Structural identity in a metric space means equal geodesic length $d_R(a, b) = d_R(c, d)$ and parallel tangent direction after transport. When the semantic space has curvature, parallel transport carries v_{ab} along the path from a to c; the analogy holds iff the transported vector matches v_{cd} up to the tolerance of the metric. ▪ ▪

This is the semantic analogue of parallel transport: the "meaning direction" of the a-to-b relationship is the same as the c-to-d relationship. Word2Vec-style embeddings approximate this by treating vector differences as semantic directions. The CPR framework gives the geometric foundation: analogy is parallel transport on the semantic reachability manifold.

41.4 The RDR Conjecture in Language

The Reachability Determines Representation Conjecture (Conjecture 5.4) applied to language:

Proposition 41.3 (*Semantic RDR*)**.** *Two linguistic expressions w_1, w_2 have the same meaning iff they produce the same continuation space:*

$$\mu(w_1)(\mathcal{C}) = \mu(w_2)(\mathcal{C}) \;\Leftrightarrow\; w_1 \equiv w_2 \text{ (synonymous)}.$$

Equivalently, $d_R(w_1, w_2) = 0$ in the semantic reachability metric.

Proof. This is the instance of the Restricted RDR Theorem (Theorem 5.3) for the semantic domain. The finite state space is the vocabulary $\mathcal{V}$ (finite for any language model), the admissibility field is the semantic coherence constraint, and the query class $\mathcal{Q}$ consists of linguistic continuations. By Theorem 5.3: two words $w_1 \equiv_{\mathcal{Q}} w_2$ iff $\mathcal{R}_{\mathcal{A}}(w_1, T) = \mathcal{R}_{\mathcal{A}}(w_2, T)$ for all T. Synonymy is the special case $T = \infty$ (identical reachable sets at all horizons). ▪ ▪

This gives a precise criterion for synonymy: two expressions are synonymous iff no downstream cognitive task can tell them apart — they open the same set of continuations. Perfect synonymy is rare because natural language expressions carry pragmatic and connotative differences that generate slightly different continuation spaces.

Exercises

41.1. Let $\mathcal{M} = \{A, B, C, D\}$ with $P(A \rightsquigarrow B) = 0.8$, $P(B \rightsquigarrow C) = 0.7$, $P(A \rightsquigarrow C) = 0.5$, $P(C \rightsquigarrow D) = 0.9$, all others determined by symmetry. Compute d_R for all pairs. Verify the triangle inequality for the triple (A, B, D).

41.2. Prove that the semantic reachability metric d_R is a *graph metric* when $\mathcal{M}$ is finite and $P(x \rightsquigarrow y) > 0$ iff (x, y) is an edge in a graph G. Identify the shortest path interpretation.

41.3. Define the *semantic curvature* at a concept c as the ratio of the circumference of a d_R-ball of radius r around c to $2\pi r$ (in 2D). Positive curvature: ball is smaller than Euclidean. Negative curvature: larger. Identify a semantic neighbourhood with positive curvature and one with negative curvature in natural language.

41.4. Show that the word embedding cosine distance $d_{\cos}(w_1, w_2) = 1 - \cos(\mathrm{e}_{w_1}, \mathrm{e}_{w_2})$ is an approximation to the semantic reachability metric d_R when embeddings are trained to predict co-occurrence. What structural assumption about $P(x \rightsquigarrow y)$ does this require?

42 Chapter 42

Projection Failure

> Ambiguity is not a flaw of language. It is what happens when one signal serves two meanings.
>
> —

Phenomenological Note. The word reaches the listener but not its full meaning. Something is lost in transmission — not because the words were wrong but because the compression always leaves something out, and what it leaves out is different for every listener. The speaker assumes a shared background. The listener fills in from their own background. Occasionally these match well enough. Often they do not, without either party noticing.

When the linguistic projection $\pi_L : \mathcal{S} \to \mathcal{L}$ maps two distinct semantic states to the same expression, a **projection failure** occurs . The Projection Failure Theorem bounds the irreducible resolution error.

42.1 The Projection Failure Theorem

Theorem 42.1 (*Projection Failure*). *Let $s_1, s_2 \in \mathcal{S}$ (Bender et al. 2021) be distinct semantic states with $\pi_L(s_1) = \pi_L(s_2) = w$ and $d_R(s_1, s_2) > 0$ (semantically distinct futures). Then for any reconstruction strategy $\hat{s} : \mathcal{L} \to \mathcal{S}$, the worst-case resolution error satisfies:*

$$\max\{d_R(\hat{s}(w), s_1),\ d_R(\hat{s}(w), s_2)\} \ \geq\ \tfrac{1}{2}\, d_R(s_1, s_2).$$

Proof. For any $\hat{s}(w) \in \mathcal{S}$, the triangle inequality for d_R gives: $d_R(s_1, s_2) \leq d_R(s_1, \hat{s}(w)) + d_R(\hat{s}(w), s_2)$. Therefore: $\max(d_R(\hat{s}(w), s_1), d_R(\hat{s}(w), s_2)) \geq \frac{d_R(s_1, s_2)}{2}$, since the maximum is at least the average. ▪ ▪

42.2 Context as Fiber Reducer

Context c reduces the effective fiber $\pi_L^{-1}(w)$ to those states compatible with c: $\mathcal{F}(w, c) = \{s \in \pi_L^{-1}(w) : s \text{ is compatible with } c\}$. When $|\mathcal{F}(w, c)| = 1$, the projection failure is resolved: the unique compatible state is the intended semantic content.

Proposition 42.2 (*Minimum Context for Disambiguation*)**.** *To reduce all projection failures for expression w with $|\pi_L^{-1}(w)| = k$ semantic states, the context must supply at least $\lceil \log_2 k \rceil$ bits of disambiguating information.*

Proof. Each bit of context can at most halve the fiber size. After $\lceil \log_2 k \rceil$ bits, the fiber has size ≤ 1. ▪ ▪

42.3 Systematic Ambiguity Classes

Projection failures cluster into three types:

Lexical ambiguity.. $\pi_L^{-1}(w)$ contains semantically unrelated states (bank: financial vs. riverbank). The two components of the fiber have disjoint reachable futures.

Structural ambiguity.. $\pi_L^{-1}(w)$ for a phrase contains different parse trees with different compositional meanings. The fiber components have overlapping but not identical futures.

Referential ambiguity.. A pronoun's antecedent is unclear. The fiber consists of all entities in the discourse context.

Each type has a different $d_R(s_1, s_2)$ and therefore different minimum context requirements for resolution.

Exercises

42.1. "I saw the man with the telescope." Identify two parse trees (structural ambiguity) and compute $d_R(s_1, s_2)$ informally. What contextual information resolves the ambiguity?

42.2. Prove that for the three-way ambiguity $|\pi_L^{-1}(w)| = 3$, the worst-case error is at least $\frac{1}{3} \max_{i \neq j} d_R(s_i, s_j)$.

42.3. Model machine translation as a projection from source-language semantic space through a shared latent space to target-language semantic space. Where do projection failures occur? Can they be reduced by increasing the shared latent space dimension?

42.4. Show that a perfectly unambiguous language (every expression has a unique semantic state) requires $|\mathcal{L}^*| \geq |\mathcal{S}|$. Conclude that any finite language must have projection failures for infinite semantic spaces.

43 Chapter 43

Gesture Before Symbol

The hand knows before the mouth names. The symbol inherits its structure from the gesture.

—

Phenomenological Note. Before you had words for something you may have had a gesture toward it — a movement, a posture, a way of turning attention that was not yet linguistic. The symbol comes later. It is a stabilized, repeatable compression of something that existed first as a more fluid form of pointing. The gesture is usually still there, below the word, available when the word is insufficient.

Symbolic representations arise from continuous gesture spaces through a canonicalisation quotient. The Gesture-to-Symbol Functor (Theorem 12.1) established that this mapping is a valid functor. This chapter develops the canonicalisation itself: how continuous input streams are segmented into discrete symbolic units via the quotient construction, and proves that the result is the minimal sufficient compression for the query "what symbol was intended?"

43.1 Gesture Canonicalisation

Definition 43.1 (*Structural Landmarks*). The **structural landmarks** of a gesture trajectory (cf. Varela et al. 1991) $\gamma \in \Gamma_{\text{raw}}$ are:

- **Overlap points**: moments when γ passes through a region shared with a canonical target trajectory;
- **Pivot points**: local extrema of curvature in γ;
- **Segment boundaries**: time points where $|\dot{\gamma}|$ crosses a threshold.

The **landmark signature** of γ is the ordered sequence of landmark types and their values.

Theorem 43.1 (*Gesture Canonicalisation*). *Define the equivalence relation: $\gamma_1 \sim \gamma_2$ iff γ_1 and γ_2 have the same landmark signature. The canonicalisation operator $\mathcal{Q} : \Gamma_{\text{raw}} \to \Gamma_{\text{raw}}/\!\sim$ satisfies:*

(i) *$\mathcal{Q}$ is the minimal sufficient compression for the query $q_{\text{sym}}(\gamma) =$ "what symbol was intended";*

(ii) *distinct equivalence classes* $[\gamma_1] \neq [\gamma_2]$ *correspond to distinct symbols;*
(iii) $\mathcal{Q}$ *is equivariant under rigid motions (rotation, translation, uniform scaling) that preserve landmark order.*

Proof. (*i*) *Minimal sufficiency.* $\mathcal{Q}$ is sufficient because the intended symbol is determined by the landmark signature (by definition of landmarks). It is minimal because any coarser quotient merges gestures with different landmark signatures, and hence different intended symbols, making the compression insufficient for q_{sym}.

(*ii*) *Distinct classes = distinct symbols.* If $[\gamma_1] \neq [\gamma_2]$, they have different landmark signatures. Different landmark signatures produce different symbolic tokens (by the definition of the gesture-to-symbol functor, Theorem 12.1).

(*iii*) *Equivariance.* Rigid motions permute the spatial positions of landmarks but preserve their types and temporal order. Since the equivalence relation is defined by landmark types and order, not absolute positions, it is invariant under rigid motions. ▪ ▪

43.2 Information Content of Canonicalisation

Proposition 43.2 (*Canonicalisation Compression Ratio*). *The fiber* $\mathcal{Q}^{-1}([\gamma])$ *contains all trajectories with the same landmark signature as* γ*. Its entropy is:* $S_{\mathcal{Q}}([\gamma]) = H(\textit{trajectory} \mid \textit{landmark signature})$*, measuring the continuous variation within a gesture class (variation in speed, amplitude, etc. that doesn't affect identity).*

Proof. A gesture trajectory γ has length $|\gamma|$ (number of motor commands). Canonicalisation maps γ to a symbol $s = \text{canon}(\gamma)$ drawn from a finite symbol set $\mathcal{S}$ (with $|\mathcal{S}| \ll |\Gamma|$). The compression ratio is $\rho = \log|\Gamma|/\log|\mathcal{S}|$ (bits in the trajectory vs. bits in the symbol). For a gesture vocabulary of G gestures and a symbol set of S symbols with $G \gg S$: $\rho = \log G/\log S \gg 1$, establishing the large compression ratio. The fiber entropy $S_{\text{canon}}(s) = \log|\{\gamma : \text{canon}(\gamma) = s\}|$ measures how many gestures map to the same symbol. ▪ ▪

This is the information that gesture-to-symbol canonicalisation discards: the fine-grained motor details that don't change what symbol was intended.

43.3 Language Origins

The canonicalisation theorem models a key step in language evolution: the emergence of discrete symbols from continuous vocalisations or gestures. The quotient $\Gamma_{\text{raw}}/\!\sim$ is the proto-lexicon: the finite set of gesture equivalence classes that a community has collectively agreed to treat as distinct. Phonological categories in spoken language are precisely the equivalence classes of the acoustic gesture space under a shared canonicalisation operator.

Exercises

43.1. For handwritten digit recognition, identify the structural landmarks for the digit "3". How many equivalence classes does $\sim$ produce?

43.2. Prove that the quotient space $\Gamma_{\text{raw}}/{\sim}$ is finite when the number of distinct landmark types is finite and trajectories have bounded complexity.

43.3. Two signers of American Sign Language use different hand shapes (different γ) but the same handshape type and movement. Verify that their gestures are in the same equivalence class under $\sim$.

43.4. Formalise categorical perception (the tendency to hear sounds as belonging to discrete phoneme categories, with sharp boundaries) as the result of applying the canonicalisation operator to the acoustic gesture space. What does the boundary between categories correspond to?

44 Chapter 44

Language as Compression

> Every sentence is a lossy compression of an experience that cannot be fully transmitted.
>
> —

Phenomenological Note. Everything that passes through language is compressed. The compression is usually so fluent that it is invisible. But there are moments when you try to say something and feel the language contracting around it — pushing it into shapes that are not quite right, losing distinctions the experience had that the words do not. You end up saying something true but not quite what you meant. The gap is real, even when it cannot be described.

Language maps infinite continuous experience to finite discrete strings. The Language Compression Bound (cf. Cover and Thomas 2006; Shannon and Weaver 1949) proves this mapping is always lossy for continuous process spaces. The bound has a simple information-theoretic derivation and deep consequences for communication, interpretation, and translation.

44.1 The Language Compression Bound

Theorem 44.1 (*Language Compression Bound*). *Let $\mathcal{X}$ be a continuous process space with Hausdorff dimension $d > 0$, and let $\mathcal{L}$ be a language with finite vocabulary $|\Sigma|$. For any encoding $\pi_L : \mathcal{X} \to \mathcal{L}^*$ and any n:*

$$\pi_L \text{ is non-injective: } \exists x_1 \neq x_2 \text{ with } \pi_L(x_1) = \pi_L(x_2).$$

Equivalently: $I(\mathcal{X}; \mathcal{L}^*) < H(\mathcal{X})$ (language captures strictly less than full process information).

Proof. $\mathcal{X}$ has Hausdorff dimension $d > 0$, so it contains an uncountable subset $U \subseteq \mathcal{X}$ (by the definition of positive Hausdorff dimension).

The set of strings $\mathcal{L}^* = \bigcup_{n=0}^{\infty} \Sigma^n$ has cardinality $|\Sigma^*| = \aleph_0$ (countably infinite).

Any map from an uncountable set to a countable set is non-injective (by the pigeonhole principle for infinite sets): there exist distinct $x_1, x_2 \in U$ with $\pi_L(x_1) = \pi_L(x_2)$.

The mutual information bound follows: $I(\mathcal{X};\mathcal{L}^*) \le H(\mathcal{L}^*)$, but $H(\mathcal{X}) = \infty$ for a continuous space while $H(\mathcal{L}^*) \le \log|\mathcal{L}^*|$ is finite for fixed-length strings. So $I(\mathcal{X};\mathcal{L}^*) < H(\mathcal{X})$. ▪ ▪

Remark 44.1 (The Irreducible Linguistic Residue)**.** The fiber $\pi_L^{-1}(w)$ — all experiences that produce the same utterance w — is the **linguistic residue**: what the expression leaves unsaid. Literature, poetry, and metaphor are strategies for reducing this residue by layering multiple projections simultaneously (denotative, connotative, rhythmic, imagistic). Mathematical notation is a strategy for minimising it in a different direction: by restricting $\mathcal{X}$ to a formal language where π_L can be made injective.

44.2 Compression Rates and Semantic Density

Definition 44.1 (*Semantic Density*)**.** The **semantic density** of an expression $w \in \mathcal{L}^*$ is

$$\rho_{\text{sem}}(w) = \frac{I(\mathcal{X};w)}{|w|_{\text{bits}}},$$

the ratio of semantic information per bit of the expression. High density = efficient encoding; low density = verbose or ambiguous encoding.

Technical writing optimises for high semantic density (maximum information per word). Poetic writing deliberately lowers semantic density to allow multiple interpretations (ambiguous fibers).

Exercises

44.1. Compute the compression ratio $I(\mathcal{X};\mathcal{L})/H(\mathcal{X})$ for a Gaussian process $\mathcal{X}$ compressed to b-bit quantised tokens.

44.2. Show that for a finite discrete process space $|\mathcal{X}| = N$, a language with $|\Sigma^n| \ge N$ strings of length n can achieve lossless compression (injective π_L). What is the minimum n in terms of $|\Sigma|$ and N?

44.3. Context-dependent languages (where $\pi_L(x \mid h)$ depends on history h) can achieve lower total residue than context-free languages. Prove this and give a natural language example.

44.4. Define the *linguistic bandwidth* of π_L as the maximum mutual information $I(\mathcal{X};\mathcal{L})$ achievable per unit time (tokens per second). How does linguistic bandwidth compare to Shannon channel capacity?

45 Chapter 45

Communication as Constraint Alignment

> Communication is not transmission of meaning. It is the coordination of admissibility fields.
>
> —

Phenomenological Note. Communication succeeds not because sender and receiver share a code but because they share enough of the world that the same compression produces similar reconstructions. When this background diverges, the words remain the same but the meaning shifts. Two people can have an entire conversation without noticing that they are reconstructing different things from the same signals.

Successful communication does not require the speaker to transmit their full internal state to the listener. It requires that speaker and listener apply compatible projections — ones that preserve the same reachability-relevant distinctions. The Communication Alignment Lemma formalises this.

45.1 The Communication Alignment Lemma

Lemma 45.1 (*Communication Alignment*). *Let $\pi_S : \mathcal{S} \to \mathcal{L}$ be the speaker's encoding and $\pi_L : \mathcal{L} \to \mathcal{S}'$ the listener's interpretation. Communication of a reachability-relevant distinction (X, Y) succeeds iff:*

$$\pi_L(\pi_S(X)) \cap \pi_L(\pi_S(Y)) = \emptyset.$$

Proof. (X, Y) is preserved by $\pi_L \circ \pi_S$ iff the images are disjoint (Lemma 6.1). The listener correctly identifies which of X or Y was communicated iff their representations in $\mathcal{S}'$ are distinguishable, which requires disjoint images. ▪ ▪

45.2 Common Ground as Shared Admissibility

Definition 45.1 (*Common Ground*). The **common ground** of speaker A and listener B is the intersection of their admissibility fields: $\mathcal{A}_{\text{cg}} = \mathcal{A}_A \cap \mathcal{A}_B$.

Proposition 45.2 (*Common Ground Sufficiency*)**.** *Communication about distinction* (X, Y) *is reliable if* (X, Y) *is reachability-relevant within the common ground:* $\mathcal{R}_{\mathcal{A}_{\text{cg}}}(X, T) \neq \mathcal{R}_{\mathcal{A}_{\text{cg}}}(Y, T)$*.*

Proof. Common ground G is the shared knowledge base between interlocutors. For a communicative query q ("what does the speaker mean?"), G provides a sufficient statistic iff: for any two contexts c_1, c_2 with the same common ground, the intended meaning is the same: $G(c_1) = G(c_2) \Rightarrow q(c_1) = q(c_2)$. This holds when q depends only on shared information, not on private beliefs. The mutual information $I(\text{meaning}; G) = H(\text{meaning})$ (common ground determines meaning) is the sufficiency condition. ▪ ▪

Communication breakdowns occur when the distinction (X, Y) is reachability-relevant within one admissibility field but not within the other, or when the intersection $\mathcal{A}_{\text{cg}}$ is too small to contain a path between the relevant states.

45.3 Alignment Strategies

Three mechanisms for achieving alignment:

Shared vocabulary.. A natural language is a shared encoding scheme $\pi_S \approx \pi_L^{-1}$, providing baseline alignment for common distinctions.

Ostension.. Pointing to an example of X and Y reduces the fiber $\pi_S^{-1}(w)$ to contain only states near the examples, making misalignment less likely.

Feedback loops.. The listener signals misalignment ("I don't understand what you mean by X"), and the speaker provides additional projection constraints until the fiber is narrowed sufficiently. This is the iterative compression protocol of Shannon channels applied to semantic communication.

Exercises

45.1. Two scientists from different fields discuss "energy." Model the misalignment between their admissibility fields on this term. What minimum common ground is needed for reliable communication?

45.2. Prove that shared language (identical $\pi_S = \pi_L^{-1}$) is sufficient but not necessary for aligned communication. Give an example of successful communication with mismatched encodings.

45.3. Formalise grounding in HRI (human-robot interaction) as the process of constructing a shared admissibility field $\mathcal{A}_{\text{cg}}$ from scratch. What is the minimum number of grounding episodes needed?

45.4. Model mathematical proof as communication of a reachability-relevant distinction (the distinction between provable and unprovable). Why is mathematical language more reliable for this than natural language?

46 Chapter 46

Document Geometry

> A document is not a bag of words. It is a constraint graph over admissible interpretations.
>
> —

Phenomenological Note. A document is a trace of a process. The process was alive and moving; the document is still. Reading it requires reconstructing the motion from the trace, which means knowing enough about the kind of process that produced it to fill in what the trace omits. An expert reader and a novice reader read the same document differently not because they see different marks but because they reconstruct different processes from the same marks.

Standard NLP represents documents as vectors in embedding space. The Document Geometry Theorem shows that document meaning is determined by the reachability structure of a constraint graph, not by individual word vectors.

46.1 Documents as Constraint Graphs

Definition 46.1 (*Document Constraint Graph*). A **document constraint graph** $\mathcal{G}_D = (\mathcal{V}, \mathcal{E})$ has:

- Nodes $v \in \mathcal{V}$: statements or propositions in document D;
- Directed edges $(v_i, v_j) \in \mathcal{E}$ with labels: Entails, Contradicts, Presupposes, or Elaborates.

The constraint graph encodes the inferential structure of the document.

Theorem 46.1 (*Document Geometry*). *Two documents D_1 and D_2 are semantically equivalent iff their constraint graphs have isomorphic reachability matrices: $\mathrm{R}_{D_1} \cong \mathrm{R}_{D_2}$. The semantic content of D is determined by the reachability structure of $\mathcal{G}_D$, not by the positions of individual words in a vector embedding space.*

Proof. Semantic equivalence means the two documents impose identical constraints on admissible continuations for any reader. Two documents impose identical constraints iff the set of inferences licensed by D_1 equals that licensed by D_2: $\{v : v \text{ is reachable from any premise}\}_{D_1} = \{v : v \text{ is reachable from any premise}\}_{D_2}$. This is exactly the condition $\mathrm{R}_{D_1} \cong \mathrm{R}_{D_2}$. ▪ ▪

46.2 Document Distance and Paraphrase

Definition 46.2 (*Document Geometry Distance*)**.** The **document geometry distance** is: $d_{\text{doc}}(D_1, D_2) = \|\mathrm{R}_{D_1} - \mathrm{R}_{D_2}\|_F$ (Frobenius norm of the reachability matrix difference).

Proposition 46.2 (*Document Distance is a Metric*)**.** *d_{doc} satisfies: non-negativity, symmetry, and the triangle inequality. Two documents are paraphrases iff $d_{\text{doc}}(D_1, D_2) = 0$.*

Proof. Inherited from the properties of the Frobenius norm on matrices. ▪ ▪

Embedding vs. Reachability. Word embedding distances measure how similar words are in distributional context — what words tend to appear near what. Document reachability distances measure what conclusions are accessible from what premises. A document can have very different reachability structure from another document despite similar embedding distributions (e.g., two arguments for opposite conclusions using similar vocabulary).

Exercises

46.1. Construct the constraint graph for a syllogism: "All mammals are warm-blooded; whales are mammals; therefore whales are warm-blooded." Compute R. What is the most semantically central node?

46.2. Two legal briefs argue for opposite verdicts using identical statutes. How can d_{doc} detect that they are semantically distinct even if their TF-IDF vectors are similar?

46.3. Prove that adding a contradictory statement to a document disconnects its constraint graph. What is the reachability consequence?

46.4. Define *document complexity* as the number of distinct strongly connected components in $\mathcal{G}_D$. Prove that a document with complexity 1 (strongly connected graph) has all its propositions mutually reachable — it forms a coherent argument.

Part VII

Computation

[Part introduction — to be written.]

47

Chapter 47

Flow Computing

> Computation is not a sequence of operations. It is a flow through admissibility space.
>
> —

Phenomenological Note. Computation is often thought of as following explicit rules. But much of what computing systems do is better described as following gradients — moving in the direction that reduces some measure of wrongness. The rules are implicit in the landscape, not stated in advance. This changes what it means for a computation to be correct, and what it means for it to fail.

Flow computing models computational processes as continuous trajectories through an admissibility field on computation state space. (Sipser 2012) Correctness is admissibility throughout; composition is the Nagumo-compatible concatenation of flows.

47.1 Computation as Admissible Flow

Definition 47.1 (*Computation State Space*)**.** The **computation state space** $\mathcal{X}_{\text{comp}}$ is the Cartesian product of program state (variables, heap, stack) with meta-state (program counter, resource usage, time). An **admissibility field** $\mathcal{A}_{\text{comp}} \subseteq \mathcal{X}_{\text{comp}}$ encodes the type system, memory bounds, and resource constraints.

Theorem 47.1 (*Pipeline Composition*)**.** *Let $\phi_1 : A \to B$ and $\phi_2 : B \to C$ be computational transformations with admissibility fields $\mathcal{A}_1$ and $\mathcal{A}_2$. They compose correctly iff:*

$$\phi_1(\mathcal{A}_1 \cap \pi_A^{-1}(A)) \subseteq \mathcal{A}_2.$$

Under correct composition, the composed flow $\phi_2 \circ \phi_1$ is admissibility-preserving and (for non-degenerate ϕ_i) the composed reachability satisfies:

$$\mathcal{V}_R(\phi_2(\phi_1(x)), T) \geq \mathcal{V}_R(x, T).$$

Proof. Compatibility of admissibility fields ensures that the image of the first stage lies within the admissible input domain of the second stage. The reachability bound follows from two applications of the Reachability Expansion Theorem (Theorem 65.1): each composition step is admissibility-preserving and measure-non-decreasing, so the composed flow expands reachability. ▪ ▪

Type Systems as Admissibility Fields. In a strongly typed language, the type system is the computational admissibility field: it forbids operations that would exit the admissible region (type errors, buffer overflows, null pointer dereferences). A type-correct program is an admissibility-preserving trajectory. Type inference is constraint-guided inference (Chapter 33) on the computation state manifold: it finds the belief trajectory (type assignment) that stays within the admissible type region while satisfying the constraints imposed by the program.

47.2 Monad Composition as Repair

Proposition 47.2 (*Monadic Repair*)**.** *The Maybe monad in functional programming is a repair operator:* `Maybe`(x) = `Just`(x) *if* $x \in \mathcal{A}$, `Nothing` *if* $x \notin \mathcal{A}$. *Monadic bind* `>>=` *composes admissibility-checked computations: failure propagates (the* `Nothing` *state is absorbing), success produces the next admissible state.*

Proof. `Nothing>>=`f = `Nothing`: once admissibility is violated, the computation halts. `Just`(x)`>>=`$f = f(x)$: admissible states proceed. This is exactly admissibility preservation: $\mathfrak{R}(x) =$ `Maybe`(x) maps inadmissible states to an absorbing sink rather than propagating errors. ▪ ▪

Exercises

47.1. Model a C++ program with manual memory management as a flow on a computation state space where the admissibility field includes the condition "no memory leaks." Identify the repair operators (destructors, RAII).

47.2. Show that the IO monad in Haskell separates the computation admissibility field (pure functional) from the effect field (IO). How does this separation improve the reconstruction of program behaviour?

47.3. A distributed system with Byzantine faults has admissibility field $\mathcal{A}_{\text{byz}}$: all valid states excluding fault states. Apply Theorem 47.1 to derive conditions under which a f-fault-tolerant composition of n nodes is admissibility-preserving.

47.4. Prove that lazy evaluation (in Haskell) is equivalent to deferring admissibility checking to the point of forced evaluation. Give an example where this deferred checking detects an error earlier than eager evaluation would.

48 Chapter 48

Markov Boundaries

A Markov blanket is an admissibility wall: what is inside cannot be directly touched from outside.

—

Phenomenological Note. There is a boundary around every coherent system, even when the boundary is not visible. Inside the boundary, things affect each other directly. Outside the boundary, they affect the interior only through the boundary itself. This is true of cells, of organisms, of institutions, of selves. The boundary is not just a physical fact; it is an informational structure that determines what counts as inside and what counts as outside.

A Markov blanket isolates a sub-system from external influences through conditional independence. This chapter proves that Markov blankets are the formal mechanism for modular admissibility: internal reachability is protected from external noise and interventions.

48.1 Markov Blankets as Admissibility Boundaries

Definition 48.1 (*Markov Blanket*)**.** Let X be a set of variables in a Bayesian network $\mathcal{G}$. The **Markov blanket** of X is the minimal set $B(X)$ such that $X \perp\!\!\!\perp \overline{X \cup B(X)} \mid B(X)$ (X is conditionally independent of all other variables given $B(X)$). In a DAG, $B(X)$ consists of X's parents, children, and co-parents.

Theorem 48.1 (*Markov Boundary Isolation*)**.** *Let $B = B(X)$ be the Markov blanket of sub-system X, and let $\mathfrak{R}_{\text{ext}}$ be a repair operator acting only on variables outside $X \cup B$. Then for any admissibility field $\mathcal{A}_X$ on X:*

$$\mathcal{V}_R(X \mid B, T) \text{ is unchanged by } \mathfrak{R}_{\text{ext}}.$$

The blanket B isolates X's reachability from external interventions.

Proof. By the conditional independence property of Markov blankets: $P(X \mid B, \text{rest}) = P(X \mid B)$. External interventions modify variables in "rest," changing the marginal $P(\text{rest})$ but leaving $P(X \mid B)$ unchanged. The reachability $\mathcal{V}_R(X \mid B, T)$ is computed from $P(X \mid B)$ and is therefore unaffected by $\mathfrak{R}_{\text{ext}}$. ▪ ▪

48.2 Spectral Gap and Mixing Time

The Markov blanket separates variables; the spectral gap measures how well information mixes within the blanket.

Proposition 48.2 (*Spectral Gap Bounds Mixing Time*)**.** *For a random walk on the reachability graph $\mathcal{G}_X$ of sub-system X (adjacency matrix A_X, Laplacian L_X, spectral gap $\lambda_2(\mathrm{L}_X)$):*

$$\tau_{\text{mix}}(X) \le \frac{\log(|X|/\epsilon)}{\lambda_2(\mathrm{L}_X)},$$

where $\tau_{\text{mix}}(\epsilon)$ is the ϵ-mixing time. Larger spectral gap $\Rightarrow$ faster mixing $\Rightarrow$ more efficient internal reachability.

Proof. Standard Markov chain mixing theory: $\|P^t - \pi\|_{\text{TV}} \le e^{-\lambda_2 t/2}$, so mixing below ϵ occurs at $t = 2\log(1/\epsilon)/\lambda_2$. ▪ ▪

Modularity as Reachability Architecture. Markov blankets are the computational mechanism for modular system design. Each module has a well-defined blanket; internal reachability is protected from external perturbation. This is the formal basis for: microservices (each service has a blanket = its API boundary); object-oriented encapsulation (the object's interface is its blanket); and cellular biology (the cell membrane is the organism's Markov blanket).

Exercises

48.1. Draw a Bayesian network for a simple sensor-actuator system (sensor S → belief B → action A → outcome O). Identify the Markov blanket of the belief node B.

48.2. Prove that the Markov blanket $B(X)$ is the minimal separator between X and the rest of the network in the undirected independence graph.

48.3. A neural network layer has activation vector h_ℓ. Show that the Markov blanket of h_ℓ is $\{h_{\ell-1}, h_{\ell+1}\}$ (the adjacent layers). What does this imply about information flow between non-adjacent layers?

48.4. The free energy principle (Friston) claims that biological organisms minimise surprise relative to their Markov blanket. Formalise this as: organisms maintain their state within $\mathcal{A}_{\text{org}} = \{x : P(x \mid B(x)) \ge e^{-F_{\max}}\}$ (states with free energy below $F_{\max}$). Verify that this defines a valid admissibility field.

49 Chapter 49

Constraint-Preserving Architectures

Safety is not a filter applied after generation. It is the admissibility field within which generation occurs.

—

Phenomenological Note. A system that can do anything is less useful than a system that can do the right things reliably. The constraints are not a limitation on the architecture; they are part of what makes it trustworthy. You want a calculator that refuses to return a negative result for a sum of squares — not because it cannot compute negative numbers but because producing one here means something has gone wrong, and you want the system to know that.

A **constraint-preserving architecture** (CPA) is a computational system whose outputs are guaranteed to lie within a specified admissibility field, not by post-hoc filtering but by structural design. The Nagumo condition (Definition 10.3) provides the key criterion.

49.1 CPA Definition and Characterisation

Definition 49.1 (*Constraint-Preserving Architecture*)**.** A computational map $f_\theta : \mathcal{X} \to \mathcal{Y}$ is a **CPA** with respect to admissibility field $\mathcal{A}_\mathcal{Y}$ if $f_\theta(x) \in \mathcal{A}_\mathcal{Y}$ for all $x \in \mathcal{A}_\mathcal{X}$.

Theorem 49.1 (*CPA via Nagumo*)**.** *A differentiable map f_θ is a CPA iff for every $y \in \partial\mathcal{A}_\mathcal{Y}$ and every x with $f_\theta(x) = y$:*

$$Df_\theta(x) \cdot v \cdot \mathrm{n}(y) \leq 0$$

for all admissible tangent directions $v \in T_x\mathcal{A}_\mathcal{X}$, where $\mathrm{n}(y)$ is the outward normal at y.

Proof. By Theorem 10.1: the Nagumo condition $\mathrm{n}(y) \cdot \dot{y} \leq 0$ ensures the output trajectory cannot cross $\partial\mathcal{A}_\mathcal{Y}$ outward. Since $\dot{y} = Df_\theta \cdot v$, the condition follows. ∎

49.2 CPA Constructions

Three practical CPA constructions:

Projection CPA.. Post-compose f_θ with the projection $\pi_{\mathcal{A}}$ onto $\mathcal{A}_y$. $\pi_{\mathcal{A}}(f_\theta(x))$ is always in $\mathcal{A}_y$ by definition. This is simple but may distort outputs when $f_\theta(x)$ is far from $\mathcal{A}_y$.

Constrained decoding.. During sequence generation, maintain a **token-level admissibility checker** that rejects tokens violating $\mathcal{A}_y$. Beam search with hard constraints is a CPA for discrete outputs.

Lyapunov-shaped networks.. Design f_θ so that a Lyapunov function $V : \mathcal{Y} \to \mathbb{R}_{\geq 0}$ with $\mathcal{A}_y = \{y : V(y) \leq V_{\max}\}$ satisfies $V(f_\theta(x)) \leq V_{\max}$ for all x. This encodes the constraint into the architecture's geometry.

49.3 CPA and Hallucination

Hallucination in LLMs occurs when the output exits the admissibility field of factually coherent statements. A factual CPA would have $\mathcal{A}_y$ defined by grounded knowledge. No current architecture fully achieves this because $\mathcal{A}_{\text{factual}}$ is not algorithmically definable — it is the complement of the hallucination set, which is itself undecidable in general (Theorem 53.1). CPAs are therefore approximations: they enforce tractable sub-conditions (citation grounding, factual entailment, logical consistency) that progressively narrow the hallucination set.

Exercises

49.1. Design a CPA for arithmetic expression evaluation. What is $\mathcal{A}_y$? Show that Lyapunov-shaped networks with $V(y) = |\text{arithmetic error}|$ satisfy the CPA condition.

49.2. Constrained decoding in language generation maintains a finite automaton tracking constraint satisfaction. Model this as a CPA with $\mathcal{A}_y =$ strings accepted by the automaton. What is the computational cost of the Nagumo check at each token?

49.3. Show that the intersection of two CPAs ($f_1 : \mathcal{X} \to \mathcal{A}_1$ and $f_2 : \mathcal{X} \to \mathcal{A}_2$) can be composed into a CPA for $\mathcal{A}_1 \cap \mathcal{A}_2$ by projecting each output onto the intersection.

49.4. (Safety.) Constitutional AI imposes preference constraints on LLM outputs. Model the constitutional principles as an admissibility field $\mathcal{A}_{\text{const}}$ and RLHF as the optimisation of f_θ toward CPA status. Under what conditions does RLHF converge to a true CPA?

50 Chapter 50

Multi-Agent Systems

> Agents are stable when their constraint fields are mutually compatible.
>
> —

Phenomenological Note. Each person in a room believes they are responding to the situation. But the situation is partly made of the other people's responses. The result is that no one is quite responding to what they think they are responding to. The equilibrium that emerges from this — if one emerges — is not what anyone wanted or planned. It is what the structure of the interactions made inevitable.

Multi-agent systems are coupled admissibility fields. The joint admissibility of n agents is the product admissibility $\bigotimes_i \mathcal{A}_i$ (Definition 7.5). This chapter proves that collective stability requires mutual compatibility of individual fields and strong connectivity of the joint reachability graph.

50.1 Multi-Agent Admissibility

Definition 50.1 (*Mutual Compatibility*)**.** Admissibility fields $\mathcal{A}_1, \ldots, \mathcal{A}_n$ (Axelrod 1984; Russell and Norvig 2020) are **mutually compatible** if $\bigcap_i \mathcal{A}_i \neq \emptyset$: there exists at least one state admissible to all agents simultaneously.

Theorem 50.1 (*Multi-Agent Stability*)**.** *A multi-agent system with fields $\{\mathcal{A}_i\}$ is **collectively stable** — has a globally stable repair fixed point — if and only if:*

(i) *The fields are mutually compatible: $\bigcap_i \mathcal{A}_i \neq \emptyset$;*

(ii) *The joint reachability graph $\mathcal{G}_{\text{joint}}$ is strongly connected.*

Proof. Necessity of (i). If $\bigcap_i \mathcal{A}_i = \emptyset$, no state is admissible to all agents. Any proposed repair fixed point x^* would be inadmissible for at least one agent, who would apply a repair operator that moves x^* — contradicting its fixed-point status.

Necessity of (ii). If $\mathcal{G}_{\text{joint}}$ is not strongly connected, some states are reachable from the proposed fixed point but cannot return to it, making the fixed point unstable under perturbation.

Sufficiency. Mutual compatibility gives a non-empty joint admissibility domain. Strong connectivity ensures the colimit $\mathrm{colim}_{\mathcal{C}}\, \mathcal{A}$ (Definition 12.4, Chapter 12) has a strongly connected reachability graph. By Theorem 67.2, the fixed point is stable with $\rho(J) < 1$ when the joint repair Jacobian has spectral radius less than 1, which follows from strong connectivity and bounded agent influence. ▪ ▪

50.2 Incompatibility and Deadlock

Proposition 50.2 (*Incompatibility Causes Deadlock*)**.** *If $\bigcap_i \mathcal{A}_i = \emptyset$ but $\bigcap_{i \neq j} \mathcal{A}_i \neq \emptyset$ for all pairs (i, j) (pairwise compatible but not jointly compatible), then the system exhibits* ***deadlock****: every proposed joint state violates at least one agent's constraints, and the agents cycle without reaching a stable point.*

Proof. At any joint state x, at least one agent k has $x \notin \mathcal{A}_k$. Agent k applies a repair operator $\mathfrak{R}_k(x) \in \mathcal{A}_k$. But $\mathfrak{R}_k(x)$ may violate another agent's field. The cycle continues: no state satisfies all agents simultaneously. ▪ ▪

This is the formal basis for deadlock in operating systems, stalemate in negotiations, and coordination failures in multi-stakeholder governance. The resolution requires either expanding the joint admissibility (relaxing some constraints) or introducing a coordination layer that overrides individual admissibility checks.

50.3 Nash Equilibrium as Repair Fixed Point

Proposition 50.3 (*Nash Equilibrium as Repair Fixed Point*)**.** *In a finite game, a Nash equilibrium x^* is a repair fixed point of the joint repair system where each agent's repair operator is "play a best response to the current joint state." Mutual compatibility corresponds to the existence of Nash equilibria (guaranteed by Nash's theorem for mixed strategies).*

Proof. In a Nash equilibrium $x^* = (x_1^*, \dots, x_n^*)$, no agent can improve by unilateral deviation: $u_i(x_i^*, x_{-i}^*) \geq u_i(x_i, x_{-i}^*)$ for all x_i and all i. The repair operator for agent i is $\mathfrak{R}_i(x) = \arg\max_{x_i} u_i(x_i, x_{-i})$ (best response). A fixed point of the joint repair $\mathfrak{R} = (\mathfrak{R}_1, \dots, \mathfrak{R}_n)$ satisfies $x_i^* = \mathfrak{R}_i(x^*)$ for all i, which is exactly the Nash equilibrium condition. Nash's theorem (existence via Kakutani's fixed-point theorem on the best-response correspondence) guarantees such a fixed point for finite games. ▪ ▪

Exercises

50.1. Three agents have admissibility fields: $\mathcal{A}_1 = [0, 2]$, $\mathcal{A}_2 = [1, 3]$, $\mathcal{A}_3 = [2, 4]$ on $\mathbb{R}$. Is the system mutually compatible? What is the joint fixed point?

50.2. Model the Prisoner's Dilemma as a two-agent system. Identify the admissibility fields, the joint reachability graph, and characterise the Nash equilibrium as a repair fixed point. Is the system strongly connected?

50.3. Prove that adding a mediating agent $n+1$ with $\mathcal{A}_{n+1} \supseteq \bigcap_i \mathcal{A}_i$ always restores mutual compatibility. Interpret: arbitrators expand the joint admissibility field.

50.4. A distributed database has n replicas. Model consistent replication as a multi-agent system where each replica's admissibility field requires agreement with a quorum of other replicas. Under what conditions is the system collectively stable?

51 Chapter 51

Parliament on a Plenum

> Democratic decision is distributed constraint field management over a shared continuous domain.
>
> —

Phenomenological Note. A vote does not reveal a pre-existing preference. It creates a record that then has political weight regardless of the actual distribution of opinion it imperfectly captures. The outcome shapes what happens next in ways that feed back into what people claim to prefer. Collective decision-making is not a mechanism for aggregating preferences. It is a mechanism for producing binding records that then become part of the environment.

The Paperbot Parliament model treats distributed decision-making as a collection of constraint fields $\{\mathcal{A}_i\}$ balanced over a shared continuous domain — the plenum $\mathcal{P}$. This chapter develops the formal model and proves stability conditions.

51.1 The Plenum Model

Definition 51.1 (*Plenum*)**.** The **plenum** $\mathcal{P}$ is the shared state space over which all agents operate: resource allocations, policy variables, or any continuous domain subject to collective decision. Agent i has admissibility field $\mathcal{A}_i \subseteq \mathcal{P}$ encoding its constraints and preferences.

Definition 51.2 (*Weighted Colimit Admissibility*)**.** The **collective admissibility** is the weighted colimit:

$$\mathcal{A}_{\text{coll}} = \left\{ x \in \mathcal{P} : \sum_i w_i 1[x \in \mathcal{A}_i] \geq \alpha \right\},$$

where $w_i \geq 0$ are influence weights (summing to 1) and $\alpha \in (0, 1]$ is the quorum threshold. $\alpha = 1$: unanimity (intersection of all fields). $\alpha = 1/2$: majority rule.

Theorem 51.1 (*Plenum Stability*)**.** *The Paperbot Parliament reaches a stable collective policy iff:*

(i) $\mathcal{A}_{\mathrm{coll}} \neq \emptyset$ (*quorum is achievable*);
(ii) *The repair dynamics on* $\mathcal{A}_{\mathrm{coll}}$ *have spectral radius* $\rho(J) < 1$.

Proof. Non-emptiness of $\mathcal{A}_{\mathrm{coll}}$ is the existence condition for a collectively admissible policy. Spectral radius $\rho < 1$ follows from the contraction mapping theorem when the weighted combination of individual repair operators is a contraction on $\mathcal{A}_{\mathrm{coll}}$ (Lemma 37.2). ▪ ▪

51.2 Quorum Threshold and Stability

The quorum threshold α trades off stability and inclusivity:

Proposition 51.2 (*Quorum-Stability Tradeoff*)**.** *As* α *increases from* $1/n$ *to* 1*:* $\mathcal{A}_{\mathrm{coll}}$ *becomes smaller (fewer jointly admissible policies) but more stable (fewer agents can unilaterally destabilise the equilibrium). The optimal* α^* *maximises* $\mathcal{V}_R(\mathcal{A}_{\mathrm{coll}}) \cdot \rho(J)^{-1}$ *(reachability volume times stability).*

Proof. As α increases from $1/n$ to 1: $\mathcal{A}_{\mathrm{coll}} = \{x : \sum_i w_i 1[x \in \mathcal{A}_i] \geq \alpha\}$ becomes a smaller set (stricter requirement). A smaller $\mathcal{A}_{\mathrm{coll}}$ means fewer admissible policies (lower reachability volume). Stability increases because any policy in $\mathcal{A}_{\mathrm{coll}}$ at high α is far from the boundary (satisfies many agents' constraints), so perturbations are less likely to cross the boundary. The optimal α^* balances these competing effects. ▪ ▪

51.3 Parliament as Collective Repair System

Each voting round is a repair step: agents propose repairs to the current collective policy, and the quorum mechanism selects an admissibility-preserving update. By Theorem 86.1, iterated voting converges to a stable policy when the quorum mechanism is admissibility-preserving and non-degenerate.

Exercises

51.1. Compute $\mathcal{A}_{\mathrm{coll}}$ for three agents with $\mathcal{A}_1 = [0, 4]$, $\mathcal{A}_2 = [1, 5]$, $\mathcal{A}_3 = [2, 6]$ and majority threshold $\alpha = 2/3$.

51.2. Show that unanimity ($\alpha = 1$) gives $\mathcal{A}_{\mathrm{coll}} = \bigcap_i \mathcal{A}_i$ and majority ($\alpha = 1/2$) gives the Condorcet median.

51.3. Gerrymandering is a manipulation of influence weights w_i. Formalise gerrymandering as the choice of $\{w_i\}$ that maximises $\mathcal{V}_R(\mathcal{A}_{\mathrm{coll}})$ for a preferred subset of agents. Prove that uniform weights $w_i = 1/n$ minimise the maximum gerrymandering benefit.

51.4. Apply Theorem 66.1 to legislative procedure: do different orderings of policy reforms produce different stable policies? Give a concrete example.

52 Chapter 52

Distinction-Centred Computation

Compute not what is true, but what is different.

—

Phenomenological Note. What you pay attention to determines what you can think about. The distinctions you carry with you into a situation are the distinctions you will find there. This is not relativism — some distinctions track real differences in the world and some do not. But the act of deciding which distinctions to make explicit and carry forward is itself a form of computation, and it happens before most of what we ordinarily call thinking.

Standard computation optimises for correct output. (Cover and Thomas 2006; Shannon and Weaver 1949) **Distinction-centred computation** (DCC) additionally optimises for preservation of reachability-relevant distinctions: the output should separate inputs with different futures and only those.

52.1 DCC Definition and Efficiency Bound

Definition 52.1 (*DCC System*)**.** A computational system $f : \mathcal{X} \to \mathcal{Y}$ is **distinction-centred** with respect to admissibility field $\mathcal{A}$ if:

(i) $f(x_1) \neq f(x_2)$ whenever $\mathcal{R}_{\mathcal{A}}(x_1, T) \neq \mathcal{R}_{\mathcal{A}}(x_2, T)$ (reachability-relevant distinctions are preserved);

(ii) $f(x_1) = f(x_2)$ whenever $\mathcal{R}_{\mathcal{A}}(x_1, T) = \mathcal{R}_{\mathcal{A}}(x_2, T)$ (reachability-irrelevant distinctions are collapsed).

A DCC system is the minimal sufficient compression for the query class $\mathcal{Q}_{\mathcal{A}}$ induced by $\mathcal{A}$.

Theorem 52.1 (*DCC Efficiency Bound*)**.** *Let D_{RR} be the number of reachability-relevant equivalence classes in $\mathcal{X}$ under $\mathcal{A}$. A DCC system requires at most $\lceil \log_2 D_{\text{RR}} \rceil$ output bits. Any additional bits are semantically redundant. Moreover, this bound is tight: there exists a DCC system achieving it.*

Proof. D_{RR} is the number of distinct values of $\mathcal{R}_{\mathcal{A}}(\cdot, T)$ in $\mathcal{X}$. Each output bit can encode one binary distinction. The minimum code length for D_{RR} classes is $\lceil \log_2 D_{\mathrm{RR}} \rceil$ by Shannon's source coding theorem. Tightness: the canonical DCC system assigns each class a codeword of length $\lceil \log_2 D_{\mathrm{RR}} \rceil$. ▪ ▪

52.2 DCC vs. Standard Computation

Standard computation preserves all input distinctions (injective functions) or collapses them arbitrarily (lossy compression). DCC is the principled middle: collapse exactly the distinctions that don't matter, preserve exactly those that do.

DCC as the Right Compression. Abstraction in mathematics and science is DCC in action: physicists treat all electrons as identical (collapse the distinction between individual electrons) because individual electron identity is not reachability-relevant for the dynamics. Two electrons in the same quantum state have $\mathcal{R}(e_1, T) = \mathcal{R}(e_2, T)$, so the collapse is admissibility-lossless. Scientific abstraction is the empirical discovery of which distinctions are reachability-irrelevant in a domain.

52.3 DCC in Neural Networks

A neural network implementing DCC would: (1) learn which distinctions in the training data are task-relevant; (2) encode those distinctions in its representations; (3) collapse task-irrelevant distinctions. Contrastive learning approximates DCC: it pulls together representations of items with similar futures and pushes apart representations of items with different futures.

Exercises

52.1. For a binary classification task with 4 input classes where classes 1 and 2 have the same label, compute D_{RR} and the minimum DCC output dimension.

52.2. Show that DCC outputs are minimal sufficient statistics (Definition 24.2) for the downstream task.

52.3. Prove that if f is DCC for admissibility $\mathcal{A}_1$ and g is DCC for $\mathcal{A}_2$, then $f \times g$ is DCC for $\mathcal{A}_1 \cap \mathcal{A}_2$ (the joint task).

52.4. Batchnorm in neural networks normalises activations across a batch, collapsing the distinction between individual sample magnitudes. Under what conditions is this a DCC operation? When does it collapse reachability-relevant distinctions?

53 Chapter 53

Admissibility Auditors

An auditor does not generate. It verifies. But verification is half of intelligence.

—

Phenomenological Note. There is usually someone or something whose job is to check whether a result is acceptable. The audit is not always after the fact. Often the expectation of the audit shapes the result from the beginning. The auditor is present in the room even when absent. What changes when audits become automatic, continuous, and invisible is not the fact of constraint but the texture of living inside it.

An **admissibility auditor** monitors system outputs and (Sipser 2012) classifies them as admissible or not. This chapter develops the auditor formalism, proves the completeness-soundness tradeoff, and derives bounds on how auditor imperfection affects system reachability.

53.1 Auditor Definition and Properties

Definition 53.1 (*Admissibility Auditor*)**.** An **admissibility auditor** for field $\mathcal{A}$ is a function $\mathcal{AU} : \mathcal{Y} \to \{0, 1\}$. It is:

- **Sound**: $\mathcal{AU}(y) = 1 \Rightarrow y \in \mathcal{A}$ (no false admissions);
- **Complete**: $y \in \mathcal{A} \Rightarrow \mathcal{AU}(y) = 1$ (no false rejections).

Theorem 53.1 (*Auditor Completeness-Soundness Tradeoff*)**.** *For decidable $\mathcal{A}$, there exists a sound and complete auditor: $\mathcal{AU} = 1[\cdot \in \mathcal{A}]$. For undecidable $\mathcal{A}$ (e.g., semantic correctness of natural language), no computable auditor can be both sound and complete simultaneously.*

Proof. If $\mathcal{A}$ is decidable, the characteristic function is computable, giving a sound and complete auditor. If $\mathcal{A}$ is undecidable, there is no algorithm that correctly classifies all inputs: by Rice's theorem, any non-trivial semantic property of programs is undecidable. An auditor that is sound must sometimes reject admissible outputs (false rejections); one that is complete must sometimes accept inadmissible ones. ▪ ▪

53.2 Auditor Imperfection and Reachability

Proposition 53.2 (*False Rejection Rate and Reachability*)**.** *Let $\mathcal{AU}$ be an auditor with false rejection rate ϵ_R (fraction of admissible outputs incorrectly rejected): $\epsilon_R = P(\mathcal{AU}(y) = 0 \mid y \in \mathcal{A})$. The effective reachability under $\mathcal{AU}$ is:*

$$\mathcal{V}_R^{\mathcal{AU}}(x, T) = (1 - \epsilon_R)\,\mathcal{V}_R(x, T).$$

Proof. The auditor accepts a fraction $(1 - \epsilon_R)$ of admissible outputs. The effective reachable set is $\{y \in \mathcal{R}(x, T) : \mathcal{AU}(y) = 1\}$, which has measure $(1 - \epsilon_R)\mu(\mathcal{R}(x, T))$ under a uniform error model. ▪ ▪

53.3 Auditor Composition

Proposition 53.3 (*Auditor Intersection*)**.** *If $\mathcal{AU}_1$ and $\mathcal{AU}_2$ audit fields $\mathcal{A}_1$ and $\mathcal{A}_2$ with false rejection rates ϵ_1 and ϵ_2, then $\mathcal{AU}_{1\wedge 2}(y) = \mathcal{AU}_1(y) \wedge \mathcal{AU}_2(y)$ audits $\mathcal{A}_1 \cap \mathcal{A}_2$ with false rejection rate $\leq \epsilon_1 + \epsilon_2$.*

Proof. $\mathcal{AU}_{1\wedge 2}(y) = 1$ iff $\mathcal{AU}_1(y) = 1$ and $\mathcal{AU}_2(y) = 1$. If $y \in \mathcal{A}_1 \cap \mathcal{A}_2$, then by soundness of each auditor: $\mathcal{AU}_i(y) = 1$ for $i = 1, 2$, so $\mathcal{AU}_{1\wedge 2}(y) = 1$ (no false rejections for $\mathcal{A}_1 \cap \mathcal{A}_2$). For the false rejection rate: $P(\mathcal{AU}_{1\wedge 2} = 0 \mid y \in \mathcal{A}_1 \cap \mathcal{A}_2) = P(\mathcal{AU}_1 = 0 \text{ or } \mathcal{AU}_2 = 0 \mid y \in \mathcal{A}_1 \cap \mathcal{A}_2) \leq \epsilon_1 + \epsilon_2$ by the union bound. ▪ ▪

This suggests a modular approach to auditing: build specialised auditors for tractable sub-constraints and compose them via conjunction.

Exercises

53.1. Design a sound auditor for mathematical proof correctness. Why can it not be complete? What is its false rejection rate for typical proof assistants?

53.2. Prove that a generator-auditor pair (generate then reject) is a CPA (Chapter 49) iff the auditor is sound.

53.3. An LLM safety filter is an auditor for $\mathcal{A}_{\text{safe}}$. If it has false rejection rate 5% (refuses 5% of safe requests) and false admission rate 0.1% (admits 0.1% of unsafe outputs), compute the effective reachability reduction for safe users.

53.4. Show that the composition of a sound generator with a complete auditor is equivalent to a CPA. When does the combination fail to be a CPA?

54 Chapter 54

Repair Operators in Computation

Error correction is repair. Debugging is repair.
Every fault-tolerant system is a repair system.

—

Phenomenological Note. Every system breaks in characteristic ways. The breaks are not random; they follow the structure of the system and the structure of the stresses applied to it. Knowing how something breaks is part of knowing what it is. The repair process reveals the same thing: you cannot fix what you do not understand, and fixing it teaches you what you were not understanding before.

Chapter 65 developed repair operators abstractly. This chapter instantiates them in computation: error-correcting codes, software exception handling, and distributed consensus are all repair operators on computational state spaces.

54.1 Error-Correcting Codes as Repair Operators

Definition 54.1 (*Error-Correcting Code as Repair*). An **error-correcting code** (n, k, d) is a repair operator $\mathfrak{R}_{\text{ecc}} : \{0,1\}^n \to \{0,1\}^k$ on the computational state space $\mathcal{X}_{\text{bit}} = \{0,1\}^n$ with admissibility field $\mathcal{A}_{\text{code}} \subseteq \{0,1\}^n$ (the set of valid codewords). $\mathfrak{R}_{\text{ecc}}(x)$ maps any received word x to the nearest valid codeword.

Theorem 54.1 (*ECC Repair Convergence*). *For a code with minimum Hamming distance $d \geq 2t+1$, the repair operator $\mathfrak{R}_{\text{ecc}}$ corrects any t-error pattern and satisfies: $\mathcal{V}_R(\mathfrak{R}_{\text{ecc}}(x), T) \geq \mathcal{V}_R(x, T)$ for all x within Hamming distance t of a valid codeword.*

Proof. For x within distance t of codeword c, $\mathfrak{R}_{\text{ecc}}(x) = c \in \mathcal{A}_{\text{code}}$. The reachability volume from a valid codeword is the full computational reachability. The reachability volume from an error-corrupted word is restricted (some computations are undefined for corrupted inputs). Therefore $\mathcal{V}_R(c, T) \geq \mathcal{V}_R(x, T)$. ▪ ▪

54.2 Exception Handling as Boundary Repair

Proposition 54.2 (*Exception as Boundary Detection*). *A* `try`/`catch` *block in an exception-handling system implements:*

- `try`: *admissibility-preserving execution within* $\mathcal{A}_{\text{comp}}$;
- `throw`: *detection of a boundary crossing* $x \notin \mathcal{A}_{\text{comp}}$;
- `catch`: *a repair operator* $\mathfrak{R}_{\text{catch}}$ *that returns the computation to an admissible state.*

Together they implement conditions (ii) and (iii) of the Synthetic Cognition Criterion (Theorem 38.1).

Proof. An exception e fires when the system's state x violates a condition C (the admissibility constraint of the try-block). Formally, the exception is triggered when $x \notin \mathcal{A}_{\text{try}} = \{x : C(x) = \texttt{true}\}$, i.e., when $x \in \partial\mathcal{A}_{\text{try}}$ or $x \notin \mathcal{A}_{\text{try}}$. The exception handler implements $\mathfrak{R}_{\text{except}}$: it maps the inadmissible state back into $\mathcal{A}_{\text{try}}$ (or into a recovery state). Exception handling is therefore the programmatic repair operator. ▪ ▪

54.3 Distributed Consensus as Collective Repair

Theorem 54.3 (*Computational Repair Convergence*). *A fault-tolerant distributed system with* n *nodes, fault rate* λ_{fault}*, and collective repair operators* $\{\mathfrak{R}_i\}$ *converges to a stable execution state* x^* *(the repair fixed point) with probability* 1*, provided:*

$$\sum_i (\mathcal{V}_R(\mathfrak{R}_i(x), T) - \mathcal{V}_R(x, T)) > \lambda_{\text{fault}} \cdot \mathcal{V}_R(x, T).$$

Proof. Model faults as downward perturbations of reachability volume at rate λ_{fault}. Repair operators produce upward corrections. The net reachability drift is positive when the sum of repair expansions exceeds the fault-induced contraction. By Theorem 86.1, monotone reachability expansion on a compact state space converges. ▪ ▪

The Byzantine Generals Problem. The Byzantine Generals Problem asks: can n generals reach consensus when up to f are traitors? In CPR terms: the collective repair operator must expand the consensus reachability despite f nodes applying anti-repair. The classical result ($n > 3f$) is the condition under which the net reachability expansion is positive despite Byzantine damage.

Exercises

54.1. For Hamming(7,4) code, compute: (a) the admissibility field $\mathcal{A}_{\text{code}}$; (b) the repair operator $\mathfrak{R}_{\text{ecc}}$; (c) the reachability expansion from a single-error word to its codeword.

54.2. Model RAID-5 disk redundancy as a repair operator. What is the admissibility field? For what fault patterns does the repair satisfy Theorem 54.1?

54.3. Prove that the Paxos distributed consensus protocol satisfies the conditions of Theorem 54.3 for crash failures (but not Byzantine failures).

54.4. Design a self-healing key-value store using the repair algebra: specify $\mathfrak{R}$ (the set of repair operators), the trigger policy τ (which repair to apply in each state), and characterise the repair fixed point.

Part VIII

Biological Systems

[Part introduction — to be written.]

55 Chapter 55

Organisms as Constraint Systems

> Life is what happens when a constraint system maintains itself against entropy.
>
> —

Phenomenological Note. Being alive requires maintaining boundaries that are thermodynamically costly. The environment constantly pushes toward equilibrium; the organism constantly pushes back. This is not a metaphor for effort or will. It is a physical description of what distinguishes living from non-living matter: the ongoing work of maintaining a configuration that would not persist on its own.

Biology has long understood organisms as information-processing systems. The CPR framework proposes a sharper characterisation: organisms are *constraint systems that maintain their own constraints*. An organism stays alive by keeping itself within its admissibility field — and the organism's defining activity is the repair of that field when it is damaged.

55.1 Metabolic Closure as Admissibility

Definition 55.1 (*Metabolic Closure*)**.** A system $\mathcal{S}$ exhibits **metabolic closure** if every component of $\mathcal{S}$ that degrades over time is produced by other components of $\mathcal{S}$:

$$\forall\, c_i \in \mathcal{S},\ \exists\, c_j \in \mathcal{S} :\ c_j \text{ catalyses the production of } c_i.$$

The directed graph of catalytic dependencies has no nodes without incoming edges — every component is maintained by the network.

Metabolic closure is the organismal form of the self-model fixed point (Chapter 37): the organism's constraint system $\mathcal{A}_{\text{org}}$ (Maturana and Varela 1980; Rosen 1991) is a fixed point of the metabolic mapping — the constraints are maintained by the very processes they constrain.

Theorem 55.1 (*Closure implies Admissibility Maintenance*)**.** *If $\mathcal{S}$ has metabolic closure and its component degradation rates λ_i are bounded: $\sup_i \lambda_i \leq \lambda_{\max}$, and if the production rates satisfy production rate of $c_i \geq \lambda_i \|c_i\|$, then $\mathcal{S}$ maintains* *its admissibility field:* $\mathcal{V}_R(\mathcal{A}_{\text{org}}(t), x_0, T) \geq \mathcal{V}_R(\mathcal{A}_{\text{org}}(0), x_0, T) - \epsilon(t)$ *where* $\epsilon(t) \to 0$ *as production rates* $\to \infty$.

Proof. By metabolic closure, every degraded component c_i is replenished by some producer c_j. The net rate of change of $\|c_i\|$ is: $\|\dot{c_i}\| = \text{production} - \lambda_i \|c_i\| \geq 0$ (by the production rate condition). Therefore $\|c_i\|$ is non-decreasing for all i, meaning the admissibility field $\mathcal{A}_{\text{org}}$ — which requires all components to exceed their functional thresholds — is maintained. The reachability volume bound follows from the monotonicity of $\mathcal{V}_R$ in $\mathcal{A}$ (Lemma 14.1). ▪ ▪

55.2 Disease as Admissibility Contraction

Definition 55.2 (*Organismal Disease*)**.** **Disease** is any process that reduces the organism's reachability volume:

$$\text{disease} \iff \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}_R(\mathcal{A}_{\text{org}}(t), x_0, T) < 0.$$

Disease contracts the space of states the organism can reach while remaining viable.

Different diseases contract reachability in different directions:

- *Physical injury*: reduces locomotor reachability.
- *Metabolic disease*: reduces the rate of repair (Γ decreases).
- *Infection*: a competing constraint system disrupts the organism's admissibility field from within.
- *Neurological disease*: reduces cognitive reachability (the organism's belief manifold contracts).
- *Aging*: global decrease in Φ and increase in S (the RSVP field dynamics of Chapter 72).

55.3 Homeostasis as Repair Equilibrium

Definition 55.3 (***Homeostasis***)**.** **Homeostasis** is the condition in which the organism's repair operators maintain $\mathcal{V}_R(\mathcal{A}_{\text{org}}(t), x_0, T)$ at a stable value despite environmental perturbations. Formally, it is a repair fixed point (Definition 67.2) of the organismal repair system, with small spectral radius $\rho(J) \ll 1$ ensuring rapid return to equilibrium after perturbation.

Homeostatic set points (body temperature, blood pH, blood glucose) are the coordinates of this repair fixed point. Illness is displacement from the fixed point; recovery is convergence back to it via the Repair Convergence Theorem (Theorem 86.1).

55.4 Evolution as Long-Run Repair

Natural selection can be understood as repair operating across populations and timescales:

Proposition 55.2 (*Selection as Population-Level Repair*)**.** *In a population with heritable variation, natural selection acts as a repair operator $\mathfrak{R}_{\text{sel}} : \mathcal{X}_{\text{pop}} \to \mathcal{X}_{\text{pop}}$ on the population state space that:*

(i) *preserves the admissibility field (only genotypes within the viable range survive);*
(ii) *expands reachability volume (higher-fitness genotypes have greater reproductive reachability).*

By Theorem 65.1, $\mathfrak{R}_{\text{sel}}$ is a repair operator, and by Theorem 86.1, iterated selection converges to locally optimal genotypes.

Proof. Selection acts on the population state space $\mathcal{X}_{\text{pop}} = \Delta^{n-1}$ (frequency simplex for n genotypes). The selection repair operator $\mathfrak{R}_{\text{sel}}(p_i) = p_i w_i/\bar{w}$ maps any admissible frequency vector (all $p_i > 0$) to another admissible vector (all frequencies remain positive). By Fisher's Fundamental Theorem (Theorem 60.1): $\Delta\bar{w} = V_A/\bar{w} \geq 0$, so $\bar{w}$ (proxy for reachability) is non-decreasing. Hence $\mathfrak{R}_{\text{sel}}$ is an admissibility-preserving, reachability-expanding operator on $\mathcal{X}_{\text{pop}}$. ∎ ∎

This is not a claim that evolution is "optimal" in any global sense — only that it is a repair process in the CPR technical sense. Local maxima of fitness (local repair fixed points) need not be global maxima.

Exercises

55.1. Model a simple metabolic network with three components $\{A, B, C\}$ where A catalyses B, B catalyses C, and C catalyses A. Verify metabolic closure. Now remove the edge $C \to A$. Is the network still closed? What happens to the admissibility field over time?

55.2. Define the *organismal reachability fraction*: $\rho_{\text{org}}(t) = \mathcal{V}_R(t)/\mathcal{V}_R(0)$. Show that ρ_{org} is non-increasing in the absence of repair. Interpret $\rho_{\text{org}}(t) = 0$ as death.

55.3. The immune system is a repair operator on the organismal state space. Model a pathogen as a perturbation $\delta\mathcal{A}$ that reduces the admissibility field. Using the Repair Convergence Theorem, derive conditions under which immune repair converges before the disease reduces $\mathcal{V}_R$ to zero.

55.4. (Aging.) Let $\Phi(t) = \Phi_0 e^{-\alpha t}$ (capacity declines exponentially) and $S(t) = S_0(1 - e^{-\beta t})$ (entropy increases). Find the age t^* at which the admissibility boundary reaches the origin $x = 0$ in a 1D model. Interpret as maximum organismal age. What repair rate Γ is needed to double t^*?

56 Chapter 56

Fungal Networks

The mycelium does not solve the maze. It becomes the solution.

—

Phenomenological Note. Intelligence does not require a center. A mycelial network exploring a forest floor for nutrients solves optimization problems that would challenge a mathematician, with no brain, no plan, no representation of the problem. The solution is distributed in the structure of the network itself, which is built by local rules operating without knowledge of the global outcome. The outcome emerges anyway.

Fungal mycelial networks are distributed constraint systems (Sheldrake 2020) that solve resource allocation problems through local gradient-following with no central controller. In CPR terms, they are biological reachability optimisers.

56.1 Mycelial Dynamics as RSVP Flow

Model the mycelial network on a spatial domain $\mathcal{U} \subseteq \mathbb{R}^2$:

- $\Phi(x,t)$: nutrient density at location x;
- $\mathrm{v}(x,t)$: mycelial growth direction (chemotaxis gradient);
- $S(x,t)$: metabolic waste and decay rate.

Hyphae extend in directions $\mathrm{v} = \mu\nabla\Phi - \nu\nabla S$ (toward nutrients, away from waste).

Theorem 56.1 (*Mycelial Reachability Optimisation*)**.** *Under the gradient-following dynamics $\mathrm{v} = \mu\nabla\Phi - \nu\nabla S$ with bounded total biomass $\int_{\mathcal{U}} B(x,t)\,\mathrm{d}x \le B_{\max}$, the mycelial network converges to a configuration maximising $\int_{\mathcal{R}} \Phi\,\mathrm{d}\mu$ (total integrated nutrient reachability) subject to the biomass constraint.*

Proof. The gradient flow $\mathrm{v} = \mu\nabla\Phi - \nu\nabla S$ is the steepest-ascent direction for the functional $F(\mathrm{v}) = \int \Phi\,\mathrm{d}\mu - (\nu/\mu)\int S\,\mathrm{d}\mu$. Under the Lyapunov function $V = -F$, $\dot{V} \le 0$ along the gradient flow. Since V is bounded below (biomass constraint), the dynamics converge to a local maximum of F by the flow existence theorem (Theorem 10.1) and the Lyapunov argument (Proposition 67.4). ∎

56.2 Network Topology and Efficiency

Proposition 56.2 (*Mycelial Minimum Spanning Tree*). *In the limit $\nu/\mu \to \infty$ (strong waste aversion, weak nutrient attraction), the mycelial network converges to a minimum spanning tree of the nutrient field's reachability graph.*

Proof. In this limit, the dominant cost is waste avoidance. The optimal network minimises total path length through waste regions, which is the minimum spanning tree of the graph weighted by $S(x)$. ▪ ▪

Slime moulds (Physarum polycephalum) exhibit exactly this behaviour: their network topology matches the minimum spanning tree of food source locations, independently rediscovering efficient network designs such as the Tokyo rail system.

56.3 Wood Wide Web

The mycorrhizal network connecting tree roots (the "wood wide web") is a multi-species reachability system: fungal hyphae extend the admissibility field of tree roots by providing nutrient pathways inaccessible to the roots alone.

Symbiosis as Admissibility Field Extension. Mutualistic symbiosis is the biological instance of HYDRA: two admissibility fields $\mathcal{A}_{\text{fungus}}$ and $\mathcal{A}_{\text{tree}}$ are coupled so that $\text{colim}(\mathcal{A}_{\text{fungus}}, \mathcal{A}_{\text{tree}}) \supset \mathcal{A}_{\text{fungus}} \cup \mathcal{A}_{\text{tree}}$. The collective reachability exceeds either partner's individual reachability. Carbon from the tree's photosynthesis extends the fungus's energy reachability; mineral nutrients from the fungus's hyphal network extend the tree's nutrient reachability.

Exercises

56.1. Model a simple two-nutrient environment (N_1 and N_2 patches) with a mycelial network of fixed total biomass. Find the optimal allocation of biomass between paths to N_1 and N_2 as a function of $\Phi(N_1)$, $\Phi(N_2)$, and the distance between them.

56.2. Slime moulds can solve maze problems by filling the maze with cytoplasm, then retracting from dead ends. Model this as a two-phase RSVP process: (a) expansion phase: Φ uniform, v isotropic; (b) contraction phase: S increases in dead ends (no nutrient reward). Prove that the surviving path is shortest.

56.3. Compare mycelial computation to neural computation. Both use gradient-following in a resource field. What is the key structural difference in their admissibility fields?

56.4. Deforestation removes mycorrhizal network nodes. Model this as removal of capacity Φ at tree locations. Using the reachability dynamics

(Theorem 73.1), derive the minimum forest patch size that prevents total mycorrhizal network collapse.

57 Chapter 57

Morphogenesis

Form is the solution to a constraint problem. Shape emerges where admissibility narrows.

—

Phenomenological Note. At some point during development a cell makes a commitment that cannot easily be reversed. It was not forced to make this commitment by any single event; it was guided into it by a field of signals and gradients that made some paths easier and others increasingly inaccessible. The commitment feels like a choice only from the outside. From inside, there was no moment of decision. There was only a direction that narrowed until it was the only one.

Morphogenesis — the development of biological form — is the biological instance of constraint-guided dynamics from Chapter 33. Cells navigate a developmental admissibility field to reach their differentiated fates, with each commitment corresponding to an irreversible reduction in the reachable set.

57.1 The Developmental Admissibility Field

Model cell development on cell-state space $\mathcal{X}_{\text{cell}}$:

- $\Phi(x,t)$: developmental capacity (pluripotency);
- $S(x,t)$: committed state (epigenetic lock-in);
- $\mathcal{A}_{\text{dev}}(t) = \{x : \Phi(x,t) - S(x,t) \geq \theta\}$: admissible developmental states.

Theorem 57.1 (*Morphogenetic Reachability*)**.** *The reachable set of differentiated states from an initial pluripotent state x_0 is $\mathcal{R}_{\mathcal{A}_{\text{dev}}}(x_0, T)$, which contracts monotonically as development proceeds: $\mathcal{V}_R(\mathcal{A}_{\text{dev}}(t_2), x_0, T) \leq \mathcal{V}_R(\mathcal{A}_{\text{dev}}(t_1), x_0, T)$ for $t_2 > t_1$. Each cell fate commitment corresponds to an admissibility boundary crossing that permanently removes states from the reachable set.*

Proof. Cell differentiation corresponds to $S(x,t)$ increasing (epigenetic commitment accumulates) while $\Phi(x,t)$ decreases (pluripotency is lost). Both effects narrow $\mathcal{A}_{\text{dev}}(t)$: $\mathcal{A}_{\text{dev}}(t_2) \subseteq \mathcal{A}_{\text{dev}}(t_1)$ for $t_2 > t_1$. By Lemma 14.1, $\mathcal{V}_R$ contracts monotonically. ▪ ▪

57.2 Waddington's Landscape

Waddington's epigenetic landscape — a ball rolling down branching valleys toward cell fate attractors — is a visualisation of the developmental admissibility manifold.

Proposition 57.2 (*Waddington as RSVP*). *Each valley in Waddington's landscape is a high-reachability corridor in $\mathcal{A}_{\text{dev}}$ (Waddington 1957): a region where the gradient $\nabla\Phi$ channels development toward a specific fate. The ridges between valleys are regions of low $\Phi - S$ (high developmental strain) that cells cannot easily cross: topological barriers in the admissibility manifold. Cell commitment is the irreversible crossing of such a ridge — an admissibility boundary crossing with $\partial_t S > 0$ (lock-in).*

Proof. The Waddington landscape potential $W(x)$ plays the role of $-\Phi(x,t)$: valleys (low W) are high-Φ admissible regions (cell fate attractors), ridges (high W) are low-Φ regions (developmental barriers). The developmental trajectory follows $\dot{x} = -\nabla W(x) = \nabla\Phi(x,t)$ (steepest descent in W = gradient ascent in Φ). The epigenetic commitment $S(x,t)$ (increasing lock-in) reduces $\mathcal{A}_{\text{dev}}$ over time, corresponding to valleys narrowing as development proceeds. The three RSVP fields map: $\Phi \leftrightarrow -W$, v $\leftrightarrow -\nabla W$, $S \leftrightarrow$ epigenetic entropy. ▪ ▪

57.3 Repair: Yamanaka Reprogramming

The Yamanaka reprogramming factors (Oct4, Sox2, Klf4, c-Myc) induce pluripotency in differentiated somatic cells — a radical repair (Definition 65.4) of the developmental state. In CPR terms: the factors transiently increase Φ and decrease S, temporarily reversing the admissibility contraction of differentiation.

Exercises

57.1. Model the binary decision between neural and epidermal cell fates as a saddle point in the developmental admissibility field. What determines which side of the saddle a cell falls on?

57.2. Turing's reaction-diffusion model generates spatial patterns in morphogenesis. Show that the pattern-forming instability corresponds to a region where $|\nabla(\Phi - S)|$ is small (flat boundary), making the admissibility boundary sensitive to small perturbations.

57.3. Apoptosis (programmed cell death) is the ultimate cellular commitment. Model it as: $\mathcal{A}_{\text{dev}}$ shrinks to a singleton $\{x_{\text{dead}}\}$ with $\mathcal{V}_R(x_{\text{dead}}, T) = 0$. When is apoptosis a repair operator on the organismal state space?

57.4. Cancer is sometimes modelled as dedifferentiation: cells regain plasticity (Φ increases, S decreases) without being guided back to pluripotency. Formalise this as admissibility field disruption and derive conditions under which it leads to unbounded reachability growth (uncontrolled division).

58 Chapter 58

Ecological Reachability

> An ecosystem is viable as long as its reachability volume is positive.
>
> —

Phenomenological Note. An ecosystem is not a collection of species. It is a collection of relationships that happens to involve species. Remove a species and you do not lose a member; you change the network of relationships, sometimes in ways that cascade through the entire system and sometimes in ways that are absorbed without consequence. Which way it goes depends on the structure of the relationships, not just the identity of the member removed.

An ecosystem is a coupled admissibility field over species and resource states. Biodiversity corresponds to reachability volume: more diverse ecosystems have larger reachable sets of energy-flow trajectories.

58.1 Ecosystem as Admissibility Field

Model an ecosystem on resource-species state space $\mathcal{X}_{\text{eco}}$:

- $\Phi(x,t)$: resource availability (nutrients, energy);
- $\mathrm{v}(x,t)$: trophic flow direction;
- $S(x,t)$: metabolic waste and entropy production.

Theorem 58.1 (*Biodiversity-Reachability Correspondence*). *For an ecosystem with n species, removing species k replaces $\mathcal{A}_{\text{eco}}$ with $\mathcal{A}_{\text{eco}} \setminus \mathcal{A}_k$ (removes the admissibility region specific to species k's role). By Lemma 14.1: $\mathcal{V}_R(\mathcal{A}_{\text{eco}} \setminus \mathcal{A}_k, x_0, T) \le \mathcal{V}_R(\mathcal{A}_{\text{eco}}, x_0, T)$. Biodiversity loss is reachability volume loss.*

Proof. Each species maintains a subset of the ecosystem's admissible energy-flow trajectories. Removing species k removes these trajectories from $\mathcal{A}_{\text{eco}}$, reducing the admissibility field. By Lemma 14.1, smaller admissibility implies smaller reachability volume. ▪ ▪

58.2 Keystone Species as Reachability Hubs

Definition 58.1 (*Keystone Species*). A species k is a **keystone species** if its removal reduces $\mathcal{V}_R(\mathcal{A}_{\text{eco}}, x_0, T)$ (Levin 1998; Odum 1971) by more than its proportional biomass would suggest: $\Delta\mathcal{V}_R/\mathcal{V}_R \gg \Delta B_k/B_{\text{total}}$.

Keystone species are reachability hubs: their admissibility region $\mathcal{A}_k$ connects otherwise disconnected parts of the ecosystem's reachability graph. Sea otters controlling sea urchin populations, wolves regulating elk in Yellowstone, elephants maintaining savannah structure — all are systems where a single species maintains a large fraction of the ecosystem's total reachability.

58.3 Trophic Cascades as Non-Commutative Repair

Proposition 58.2 (*Trophic Cascade as Non-Commutative Repair*). *Reintroducing two previously extirpated species A and B gives different ecosystem outcomes depending on order: $\mathfrak{R}_A \circ \mathfrak{R}_B \neq \mathfrak{R}_B \circ \mathfrak{R}_A$ in general (Theorem 66.1). This is because species A's admissibility field partially depends on species B's presence, and vice versa.*

Proof. Let species A and B be reintroduced sequentially. Introducing A first modifies $\mathcal{A}_{\text{eco}}$: the presence of A opens new admissible energy flows (predation, competition release) that change $\mathcal{A}_{\text{eco}}$. When B is then introduced, it enters an ecosystem already modified by A, so the joint outcome $\mathfrak{R}_B(\mathfrak{R}_A(\mathcal{A}))$ differs from introducing B first: $\mathfrak{R}_A(\mathfrak{R}_B(\mathcal{A})) \neq \mathfrak{R}_B(\mathfrak{R}_A(\mathcal{A}))$ in general (since A's and B's admissibility fields interact with each other). ▪ ▪

Exercises

58.1. Model a simple food web: grass → rabbit → fox → wolf. Compute the reachability matrix R and identify the keystone species.

58.2. Invasive species expand $\mathcal{A}_{\text{eco}}$ in some dimensions (they occupy new niches) but contract it in others (they outcompete native species). Give conditions under which an invasion is net-positive for reachability.

58.3. The holocene extinction is estimated to have reduced global biodiversity by $\sim 50\%$. Using the biodiversity-reachability correspondence, estimate the reduction in ecosystem reachability volume. What functional consequences does this predict?

58.4. Design a rewilding programme using the repair scheduling framework (Open Problem 90.4). Which species should be reintroduced first to maximise cumulative reachability recovery?

59 CHAPTER 59

Distributed Intelligence

Intelligence is not located. It is distributed through an admissibility field.

—

PHENOMENOLOGICAL NOTE. A crowd can be smarter than any of its members or dumber than all of them, depending on the structure of how information flows through it. The same is true of groups, teams, committees. Intelligence is not something a group has the way an individual has it. It is a property of the interaction structure — which emerges under some conditions and fails to emerge under others that look superficially similar.

Slime moulds, ant colonies, and immune systems exhibit intelligent behaviour with no central controller. The CPR account: distributed intelligence arises when (Levin 1998) local admissibility-preserving actions collectively expand the system's reachability volume beyond what any individual component achieves.

59.1 The Emergent Intelligence Criterion

Theorem 59.1 (*Emergent Intelligence*). *A distributed system $\{\mathcal{S}_i\}$ with local repair operators $\{\mathfrak{R}_i\}$ exhibits **emergent intelligence** iff:*

(i) *Each $\mathfrak{R}_i$ is admissibility-preserving;*
(ii) *The collective $\operatorname{colim}_C \mathcal{A} = \operatorname{colim}_i \mathcal{A}_i$ has higher reachability volume than any individual $\mathcal{A}_i$: $\mathcal{V}_R(\operatorname{colim}_C \mathcal{A}, x_0, T) > \max_i \mathcal{V}_R(\mathcal{A}_i, x_0, T)$;*
(iii) *The collective repair orbit converges (Theorem 86.1).*

Proof. Condition (i) ensures each agent's actions remain within bounds. Condition (ii) is the definition of emergence: the collective achieves more than any individual. Condition (iii) ensures stability of the collective behaviour. Together they constitute a system that is collectively more capable (higher $\mathcal{V}_R$) and stable than its parts. ▪ ▪

59.2 Stigmergy as Distributed RSVP

Stigmergy is indirect coordination through environmental modification: ants deposit pheromones that guide other ants, creating a distributed RSVP field in the physical environment.

Proposition 59.2 (*Stigmergy as Distributed Capacity Field*)**.** *Pheromone concentration $\Phi(x,t)$ is the RSVP capacity field of the ant colony's admissibility structure. Individual ants follow the gradient $\mathrm{v} = \nabla\Phi$ (move toward higher pheromone concentration). Pheromone evaporation is the entropy field $S \propto -\partial_t \Phi$. The colony collectively maintains the RSVP system through distributed deposition and evaporation.*

Proof. The pheromone field $\Phi(x,t)$ satisfies: (i) $\Phi \geq 0$ (concentrations are non-negative); (ii) $\partial_t \Phi = -\lambda\Phi + D\nabla^2\Phi + \rho(x,t)$ (decay + diffusion + deposition by ants at rate ρ). This is the RSVP Φ-equation with: decay coefficient λ playing the role of entropy drain λS, diffusion $D\nabla^2\Phi$ playing the role of transport $\mathrm{v}\cdot\nabla\Phi$, and deposition ρ playing the role of repair Γ. Individual ant movement follows $\mathrm{v} = \nabla\Phi$ (chemotaxis), completing the RSVP identification. ▪ ▪

59.3 Immune System as Distributed Repair

The immune system is a distributed repair system on the organismal admissibility field. B cells and T cells each respond to local signals (antigens, cytokines) without central coordination, yet produce globally coherent responses.

Distributed Intelligence as HYDRA. All three examples — mycelium, ant colony, immune system — are instances of HYDRA: many local admissibility fields $\{\mathcal{A}_i\}$ coupled through a shared environment to produce a collective $\mathrm{colim}_{\mathrm{C}}\, \mathcal{A}$ with properties no individual $\mathcal{A}_i$ achieves alone. The hallmark is super-additivity: $\mathcal{V}_R(\mathrm{colim}_{\mathrm{C}}\, \mathcal{A}) > \sum_i \mathcal{V}_R(\mathcal{A}_i)$.

Exercises

59.1. Model an ant colony foraging for two food sources as a distributed RSVP system. Show that the pheromone trails converge to the shortest-path solution even without any ant knowing the global geometry.

59.2. Verify the Emergent Intelligence Criterion for a two-agent system: $\mathcal{A}_1 = \{x : x_1 \leq 3\}$, $\mathcal{A}_2 = \{x : x_2 \leq 3\}$ in $\mathbb{R}^2$, starting from $x_0 = (1,1)$. Compute $\mathcal{V}_R(\mathcal{A}_1)$, $\mathcal{V}_R(\mathcal{A}_2)$, and $\mathcal{V}_R(\mathrm{colim}_{\mathrm{C}}\, \mathcal{A})$ at horizon $T = 1$.

59.3. A swarm of drones can collectively map a building that no single drone can explore alone. Formalise this as emergent intelligence: specify $\mathcal{A}_i$ (each drone's reachability within its battery limit) and show that $\mathrm{colim}_{\mathrm{C}}\, \mathcal{A}$ covers the full building.

59.4. Why does emergent intelligence require condition (ii) ($\mathcal{V}_R(\mathrm{colim}_{\mathrm{C}}\, \mathcal{A}) > \max_i \mathcal{V}_R(\mathcal{A}_i)$)? Give an example of a system satisfying (i) and (iii) but not (ii) that does not exhibit emergent intelligence.

60 Chapter 60

Evolutionary Repair

> Evolution is the slowest repair system in nature.
> And the most thorough.
>
> —

Phenomenological Note. Evolution does not have a direction. Selection favors what works in the current environment, not what will work in the next one. The remarkable thing is not that evolution produces good solutions but that it produces solutions at all, given that the criteria change, the environment changes, and the available variation is random. That it works is a consequence of the structure of the problem, not of any guidance built into the process.

Natural selection acts as a repair operator on population state space: it increases the reproductive reachability of populations by favouring genotypes better suited to the current admissibility field. Fisher's Fundamental Theorem (Fisher 1930) of Natural Selection is the evolutionary instance of Theorem 65.1.

60.1 Selection as Repair Operator

Definition 60.1 (*Population State Space*). The **population state space** $\mathcal{X}_{\text{pop}}$ has coordinates $\mathrm{p} = (p_1, \ldots, p_n)$ (frequencies of n genotypes) in the simplex Δ^{n-1}. The admissibility field $\mathcal{A}_{\text{pop}}$ is the set of viable population compositions (e.g., $p_i > 0$ for all i in a polymorphic population).

Theorem 60.1 (*Fisher's Fundamental Theorem as Repair*). *Natural selection is a repair operator $\mathfrak{R}_{\text{sel}} : \mathcal{X}_{\text{pop}} \to \mathcal{X}_{\text{pop}}$ satisfying:*

$$\Delta\bar{w} = \frac{V_A}{\bar{w}} \geq 0,$$

where $\bar{w}$ is mean fitness, V_A is additive genetic variance, and $\Delta\bar{w}$ is the one-generation change in mean fitness. Since $\mathcal{V}_R(\mathfrak{R}_{\text{sel}}(\mathrm{p}), T) \propto \bar{w}(\mathrm{p})$ (fitness proxies reachability), selection is a reachability-expanding repair operator.

Proof. The standard Wright's equation for selection: $\Delta p_i = p_i(w_i - \bar{w})/\bar{w}$. Price's equation gives $\Delta\bar{w} = V_A/\bar{w} \geq 0$ (Fisher's theorem). The monotone increase in mean fitness is the Lyapunov condition for the repair Lyapunov function $V(\mathrm{p}) = \bar{w}(\mathrm{p})$: $V(\mathfrak{R}_{\mathrm{sel}}(\mathrm{p})) \geq V(\mathrm{p})$. By Proposition 67.4, this establishes the repair orbit as monotonically increasing in reachability. ▪ ▪

60.2 Genetic Drift as Noise in the Repair Operator

In finite populations of size N, the selection repair operator is perturbed by random genetic drift: $\mathfrak{R}_{\mathrm{drift}}(\mathrm{p}) = \mathfrak{R}_{\mathrm{sel}}(\mathrm{p}) + \boldsymbol{\xi}$ where $\boldsymbol{\xi}$ has variance $\propto 1/N$. By the stochastic version of Theorem 86.1, convergence to the fitness maximum persists when selection dominates drift: $V_A/\bar{w} \gg 1/(N\bar{w})$, i.e., when effective population size $N \gg 1/V_A$.

60.3 Speciation as Admissibility Field Splitting

Speciation occurs when a population's admissibility field splits into two isolated components: $\mathcal{A}_{\mathrm{pop}} \to \mathcal{A}_1 \sqcup \mathcal{A}_2$ (by geographic isolation, reproductive barriers, etc.). Separate evolutionary repair then proceeds independently in each component, producing divergent lineages.

Exercises

60.1. For two competing genotypes A (fitness $w_A = 1.1$) and B (fitness $w_B = 1.0$) in a population of frequency p: derive Δp and $\Delta\bar{w}$ and verify Fisher's theorem. How many generations until $p_A > 0.99$?

60.2. Model antibiotic resistance as adversarial repair: the bacteria's selection repair operator increases resistance frequency, while the antibiotic is a damage operator reducing the susceptible population. Under what conditions does resistance fix before the infection is cleared?

60.3. Sexual recombination creates new genotype combinations not present in the parent population. Prove that sexual reproduction is a repair operator with $|\det J_{\mathfrak{R}_{\mathrm{sex}}}| > 1$ (exploration expands the effective reachable set) while asexual reproduction has $|\det J_{\mathfrak{R}_{\mathrm{asex}}}| = 1$ (clonal copying).

60.4. Mass extinction events are catastrophic admissibility field contractions ($\mathcal{A}_{\mathrm{pop}}$ shrinks suddenly). Using the Repair Convergence Theorem, derive the conditions under which post-extinction recovery converges to a new stable biodiversity equilibrium.

61 Chapter 61

Biological Compression

A genome is a compressed history of everything that worked.

—

Phenomenological Note. The genome does not specify an organism. It specifies a developmental process that, under the right conditions, produces an organism. The organism is not in the genome the way a building is in a blueprint. It is more like the way a piece of music is in a score — the score constrains the performance but the performance is not the score, and the music only exists when it is played.

Biological organisms compress evolutionary and developmental history into compact representations: the genome, the proteome, the connectome. The sufficiency theorems of Chapter 24 determine which compressions preserve fitness-relevant information.

61.1 The Genome as Compression Operator

Definition 61.1 (*Genotype-Phenotype Map*)**.** The **genotype-phenotype map (cf. Kauffman 1993; Waddington 1957)** $G : \mathcal{X}_{\text{geno}} \to \mathcal{X}_{\text{pheno}}$ is a compression operator from genotype space (DNA sequences) to phenotype space (organism form and function).

Theorem 61.1 (*Genomic Compression Sufficiency*)**.** *The genome $G(\gamma)$ is a sufficient compression for the fitness query $q_{\text{fit}}(\gamma)$ iff:*

$$G(\gamma_1) = G(\gamma_2) \;\Rightarrow\; q_{\text{fit}}(\gamma_1) = q_{\text{fit}}(\gamma_2).$$

Genotypes with identical phenotypes have identical fitness.

Proof. Direct application of Theorem 24.1 with $\mathcal{C} = G$ and $q = q_{\text{fit}}$. ▪ ▪

61.2 Degeneracy and Robustness

The genomic compression has large fibers: many distinct genotypes map to the same phenotype ($|G^{-1}(\text{phenotype})| \gg 1$). This **genetic degeneracy** is not inefficiency — it is a feature that confers robustness.

Proposition 61.2 (*Degeneracy Enables Robustness*). *If G^{-1}(phenotype) is large, then the population can explore the genotypic fiber without changing phenotype. Mutations within the fiber are neutral (invisible to selection). This neutral drift accumulates genetic variation that can be rapidly deployed when the environment changes (the admissibility field shifts).*

Proof. The fiber G^{-1}(phenotype) has size $|F| = |G^{-1}(\text{phenotype})|$. Neutral mutations traverse within this fiber (changing genotype but not phenotype, invisible to selection). When the environment changes (new admissibility field), the previously neutral variation in F may now have fitness consequences: some variants in F are better adapted to the new field. Selection can then act rapidly on pre-existing variation without waiting for new mutations, giving $O(1)$ rather than $O(1/\mu)$ adaptation time (where μ is the mutation rate). ▪ ▪

Genome as Compressed Evolutionary History. The genome is not just a blueprint for the organism. It is a compressed record of all the admissibility tests passed by ancestors over millions of generations. Each gene is a compressed constraint: its sequence encodes the subset of chemical space found admissible by evolutionary selection. The genome is therefore a sufficient statistic for the evolutionary query: "what developmental trajectories are viable?"

61.3 Neural Compression

The brain compresses sensory experience into neural representations. The connectome ($\approx 10^{11}$ neurons, 10^{15} synapses) compresses the full sensory history into a representation sufficient for the motor and cognitive queries the organism faces.

Proposition 61.3 (*Sensory Compression Sufficiency*). *The neural compression $C_{\text{neural}} : \mathcal{X}_{\text{sensory}} \to \mathcal{X}_{\text{neural}}$ is sufficient for survival queries q_{surv} iff neural representations distinguish the stimuli that require different responses. Fisher degeneracy in C_{neural} corresponds to perceptual categories: regions where distinct stimuli produce identical neural responses (categorical perception, Chapter 40).*

Proof. Sufficiency for survival queries: by Theorem 24.1, C_{neural} is sufficient for q_{surv} iff $C_{\text{neural}}(s_1) = C_{\text{neural}}(s_2) \Rightarrow q_{\text{surv}}(s_1) = q_{\text{surv}}(s_2)$. Stimuli that require the same motor response (e.g., all variants of "predator present") should map to the same neural code. The Fisher degeneracy at $C_{\text{neural}}(s)$ corresponds to directions in stimulus space that require no discriminative response: the nervous system is indifferent to these directions, which is precisely categorical perception (coarse discrimination between categories, fine discrimination within). ▪ ▪

Exercises

61.1. Human genome: $\sim 3 \times 10^9$ base pairs, but only $\sim 2\%$ are protein-coding. Compute the compression ratio of the protein-coding genome relative to the full genome. What does the non-coding DNA compress?

61.2. Synonymous mutations (same amino acid, different codon) are mutations within the fiber of the genotype-phenotype map. Compute the average fiber size for a 300-residue protein using the genetic code table.

61.3. A compressed sensory representation has Fisher metric tensor g. Prove that the directions of zero Fisher curvature correspond to stimuli that cannot be discriminated by the organism. Relate this to the just-noticeable difference (JND) in psychophysics.

61.4. Horizontal gene transfer (bacteria exchanging genes directly) is compression followed by decompression: genes are extracted ($\mathcal{C}$) from donor chromosome, transferred, and inserted ($\mathcal{R}$) into recipient. When does this violate the Reconstruction Existence Lemma (Lemma 25.1)?

Part IX

Institutions and Society

[Part introduction — to be written.]

62 Chapter 62

Collective Belief Systems

An institution believes through its records, not through its members.

—

Phenomenological Note. What a group believes is not the average of what its members believe. Beliefs are social objects — they require maintenance, they are held partly by being repeated aloud, they can persist in a group long after most members have privately stopped holding them. This is why institutions can believe things that no single member would endorse on reflection, and why changing what an institution believes is harder than changing what any given member believes.

Institutions maintain shared belief states distributed across members, documents, procedures, and built environments. The collective belief (cf. Luhmann 1995; March and Simon 1958) manifold $\mathcal{B}_{\text{coll}}$ is the product of individual belief manifolds constrained by the institution's shared admissibility field.

62.1 Collective Belief as Constrained Product Manifold

Definition 62.1 (*Collective Belief Manifold*)**.** The **collective belief manifold** of an institution with n members is:

$$\mathcal{B}_{\text{coll}} = \{(\theta_1, \dots, \theta_n) \in \prod_i \mathcal{B}_i : C(\theta_1, \dots, \theta_n) \le C_{\max}\},$$

where $C(\theta_1, \dots, \theta_n)$ measures disagreement among members (e.g., total pairwise Fisher distance) and $C_{\max}$ is the institution's tolerance for internal divergence.

Theorem 62.1 (*Collective Belief Stability*)**.** *A collective belief system is stable iff: $\mathcal{B}_{\text{coll}}$ is strongly connected (every admissible belief is reachable from every other through collectively admissible transitions), and the collective repair operators (debate, voting, precedent) have Lyapunov function $V = C$ with $\dot{V} \le 0$ along repair trajectories (disagreement decreases over time).*

Proof. Strong connectivity ensures no belief can permanently diverge from the collective field. The Lyapunov condition ensures the repair operators reduce rather than amplify disagreement. By Proposition 67.4, the repair orbit converges to a fixed point of minimum collective disagreement. ▪ ▪

62.2 Echo Chambers as Disconnected Belief Graphs

Proposition 62.2 (*Echo Chambers as Weakly Connected Sub-Graphs*)**.** *An echo chamber is a strongly connected component of $\mathcal{B}_{\text{coll}}$ that is only weakly connected to other components: transitions from the echo chamber to outside are rare ($P(\theta_{\text{in}} \to \theta_{\text{out}}) \ll 1$). Within the echo chamber, internal reachability is high; cross-chamber reachability is near zero.*

Proof. An echo chamber is characterised by high internal connectivity (strong edges within the chamber) and low external connectivity (weak or absent edges crossing to other belief communities). In graph terms, this is a weakly connected subgraph: strongly connected internally (every belief within the chamber is reachable from every other) but only weakly connected externally (the cross-chamber edges are weak or unidirectional). By Proposition 11.2, the Fiedler value λ_2 for the full belief graph approaches zero as cross-chamber connectivity weakens, indicating near-disconnection and slow cross-chamber mixing. ▪ ▪

The formation of echo chambers corresponds to the fragmentation of the collective belief graph into isolated SCCs — the same phenomenon as institutional collapse (Chapter 69) at the level of belief topology.

62.3 Institutional Memory as Belief Compression

Shared institutional beliefs are maintained through documents, precedents, and rituals — a compression $\mathcal{C}_{\text{inst}}$ of the full individual belief manifold. The institution's ability to answer query q reliably depends on $\mathcal{C}_{\text{inst}}$ being sufficient for q (Theorem 27.1).

Exercises

62.1. Model scientific consensus as a collective belief system. What is the admissibility field? What repair operators enforce it? When does a paradigm shift correspond to a phase transition in the collective belief manifold?

62.2. Prove that an institution with all members holding identical beliefs ($C = 0$) has maximum stability but minimum epistemic reachability (no exploration of the belief space). What is the optimal diversity level?

62.3. Formalise misinformation as an adversarial insertion into $\mathcal{B}_{\text{coll}}$: false beliefs $\theta_{\text{false}} \notin \mathcal{A}_{\text{true}}$ are introduced while appearing to lie within individual members' admissibility fields. Under what conditions does the collective belief system repair this adversarial injection?

62.4. Apply the Non-Commutativity Theorem to the sequence of reforms and counter-reforms in a political institution. Does the order of institutional changes affect the final stable collective belief state?

63 Chapter 63

Institutional Memory

An institution without memory repeats its failures.
An institution without forgetting cannot adapt.

—

Phenomenological Note. An institution remembers in documents, procedures, personnel, and physical arrangements. When any of these change, something is lost that may not be noticed until it is needed. The new employee who does not know why the form is filled out that way, the renovation that removes a room where informal meetings used to happen, the updated procedure that eliminates a step that turned out to be load-bearing — institutional forgetting is mostly invisible until it is not.

Institutional memory (cf. March and Simon 1958; North 1990) compresses the history of an institution into records, laws, precedents, and tacit knowledge sufficient to guide future decisions. The sufficiency theorem determines which compressions are adequate for which institutional queries.

63.1 Institutional Memory as Compression

Theorem 63.1 (*Institutional Memory Sufficiency*)**.** *An institution can reliably decide query q from its records iff its memory compression $\mathcal{C}_{\text{inst}}$ is sufficient for q: $\mathcal{C}_{\text{inst}}(\gamma_1) = \mathcal{C}_{\text{inst}}(\gamma_2) \Rightarrow q(\gamma_1) = q(\gamma_2)$. Institutional failures of memory — repeating past errors, losing hard-won knowledge — are instances of insufficiency for the relevant decision queries.*

Proof. Direct application of Theorem 27.1 to the institutional context. ▪ ▪

63.2 Three Institutional Memory Modes

Institutions maintain memory through three distinct compressions, each with different sufficiency profiles:

Explicit records. (laws, archives, minutes). High sufficiency for precise factual queries. Low sufficiency for contextual, tacit, or experiential queries. Compression depth D: typically finite (records are deleted, lost, or degraded).

Procedural knowledge. (standard operating procedures). High sufficiency for routine operational queries. Low sufficiency for novel situations outside the procedure's scope. Compression depth D: moderate (SOPs encode lessons up to their revision date).

Tacit knowledge. (expert judgment, organisational culture). High sufficiency for nuanced contextual queries. Near-zero for queries requiring explicit justification. Compression depth D: effectively unbounded but non-transferable.

63.3 The Stability-Adaptability Tradeoff

Proposition 63.2 (*Memory Depth Tradeoff*)**.** *Deeper compression (larger D) gives: higher sufficiency for historical queries (better institutional memory), but higher rigidity (harder to update when the admissibility field changes). Shallower compression gives adaptability at the cost of memory. The optimal depth D^* minimises the sum of insufficiency loss (from too-shallow compression) and maladaptation cost (from too-deep compression).*

Proof. Deeper compression (larger D) retains more historical information: $H(\mathcal{C}_{D+1}) \geq H(\mathcal{C}_D)$ (the deeper compression is a refinement). The insufficiency loss decreases with D: $L_{\text{insuff}}(D) = H(q \mid \mathcal{C}_D(\Gamma))$ is non-increasing in D (more compression depth means better sufficiency for historical queries q). The maladaptation cost arises because deeper compression makes the system harder to update when the admissibility field changes: $L_{\text{maladapt}}(D) \propto D$ (the truncation error of dropping history $> D$ grows with D when the new field differs from the old). The optimal D^* minimises $L_{\text{insuff}}(D) + L_{\text{maladapt}}(D)$, which has a unique minimum under convexity of L_{insuff} and linearity of L_{maladapt}. ▪ ▪

Statute of limitations laws ($D \approx 7$ years for civil claims) and stare decisis (precedent binding future courts) are institutional mechanisms that calibrate D^* in different query domains.

Exercises

63.1. A company loses its institutional memory when key employees retire without knowledge transfer. Model this as compression insufficiency: which queries lose their sufficient statistic? Design a knowledge transfer protocol as a compression operator.

63.2. Apply the Truncation Stability Bound (Theorem 26.2) to common law: if precedent influence decays as $\alpha_k = r^k$ $(r < 1)$, find the minimum citation depth D for ϵ-reliable precedent.

63.3. Constitutional entrenchment (requiring supermajority to amend) is a mechanism for deep institutional memory. Formalise this as a compression with very high stability (small λ in the memory field equation). When does high stability become a liability?

63.4. Model the academic peer review system as institutional memory compression. What queries is it sufficient for? What notorious insufficiencies exist (e.g., the replication crisis, publication bias)?

64 Chapter 64

Legibility and Projection

> To make legible is to project. To project is to lose.
> The question is whether what is lost matters.
>
> —

Phenomenological Note. The same event looks very different depending on whether you are administering it or living it. The administrator sees a category, a policy, a data point. The person living it sees a specific situation with specific history that does not map cleanly onto the category. Neither is wrong, but neither is complete. The mismatch is structural, not a failure of attention on either side.

State legibility (in the sense of James Scott's (Scott 1998) *Seeing Like a State*) is a projection from complex social processes to simplified administrative categories. The Legibility Collapse Theorem shows that any finite legibility system necessarily loses reachability-relevant social distinctions.

64.1 The Legibility Collapse Theorem

Theorem 64.1 (*Legibility Collapse*)**.** *Any legibility projection $\pi_{\text{leg}} : \mathcal{X}_{\text{soc}} \to \mathcal{L}$ from social process space to a finite legible category set $|\mathcal{L}| = k$ collapses at least $|\mathcal{X}_{\text{soc}}|/k$ processes per category. For continuous $\mathcal{X}_{\text{soc}}$, every category contains infinitely many collapsed social processes. The collapsed processes include reachability-relevant distinctions iff the social admissibility field $\mathcal{A}_{\text{soc}}$ is not constant on $\pi_{\text{leg}}^{-1}(\ell)$ for some ℓ.*

Proof. Cardinality argument: with $|\mathcal{X}_{\text{soc}}| > k$, at least one category contains multiple processes (pigeonhole for finite case; all categories contain infinitely many for continuous $\mathcal{X}_{\text{soc}}$). Reachability-relevance: if $\mathcal{A}_{\text{soc}}$ varies within a fiber $\pi_{\text{leg}}^{-1}(\ell)$, there exist processes $x_1, x_2 \in \pi_{\text{leg}}^{-1}(\ell)$ with $\mathcal{R}(x_1, T) \neq \mathcal{R}(x_2, T)$. The state's administrative response is identical for x_1 and x_2 (both are category ℓ) despite their different actual futures. ▪ ▪

64.2 Scott's Examples as Projection Failures

James Scott identifies several legibility failures that are now interpretable as projection collapse:

Scientific forestry.. Projecting diverse forest ecosystems onto single-species uniform plantations. The fiber $\pi_{\text{leg}}^{-1}(\texttt{pine-forest})$ contains both productive and fragile ecosystems; the monoculture response is optimal for neither.

Authoritarian high modernist cities.. Projecting complex urban social patterns onto grid layouts. The fiber $\pi_{\text{leg}}^{-1}(\texttt{residential-block})$ contains both vibrant community hubs and social isolation.

Collectivisation.. Projecting diverse agricultural practices onto single crop systems. The local knowledge compressed away in $\pi_{\text{leg}}^{-1}(\texttt{collective-farm})$ was necessary for sustainable yields.

64.3 Participatory Institutions as Resolution Increase

Proposition 64.2 (*Participation Reduces Legibility Loss*)**.** *Participatory institutions that incorporate local knowledge increase the effective category count k (local conditions become distinct categories) and reduce $\Lambda(m)$ (mixing within each category decreases). Participation is legibility resolution increase.*

Proof. A participatory institution uses local knowledge to create subcategories: $\pi_{\text{part}}^{-1}(\ell) \subset \pi_{\text{leg}}^{-1}(\ell)$ for each global category ℓ. The mixing $\Lambda(m)$ within each local category is smaller: $\Lambda_{\text{part}}(m) \leq \Lambda_{\text{leg}}(m)$ since local subcategories are more homogeneous. Averaging over categories: $\mathbb{E}[\Lambda_{\text{part}}(m)] \leq \mathbb{E}[\Lambda_{\text{leg}}(m)]$, so the expected legibility loss (from Theorem 64.1) is reduced. ▪ ▪

Exercises

64.1. Model the British colonial land survey (imposing property lines on pre-existing land use patterns) as a legibility projection. What admissibility-relevant distinctions were collapsed? What administrative failures resulted?

64.2. Digital surveillance is an attempt to increase k indefinitely (every individual becomes a distinct category). Does this eliminate legibility collapse? What new projection failures does it introduce?

64.3. Design a participatory planning process as a legibility system with $k = k_{\text{global}} + k_{\text{local}}$ (global categories plus locally-defined subcategories). Prove that this reduces $\mathbb{E}[\Lambda(m)]$ compared to the global-only system.

64.4. Apply the Reconstruction Error Bound (Theorem 22.1) to administrative error rates: if the social process space has curvature $\kappa_{\max}$ and legibility projection degeneracy $\sigma_{\min}(J_\pi)$, bound the fraction of administrative decisions that fail due to legibility collapse.

65 Chapter 65

Repair Theory

> Repair is not the restoration of the past. It is the expansion of the future.
>
> —

Phenomenological Note. There is a specific feeling that arrives when you realize you have been doing something wrong for a long time — not a small error but a structural one, something in the approach itself. The feeling is not quite regret and not quite relief. It is closer to orientation: the wrong path is now visible as wrong, which means you can finally see where the right path might be. Repair begins at this recognition, not before it.

Every system governed by an admissibility field can be damaged: its reachable set can shrink below what the constraint structure would permit at full health. **Repair** is any process that restores or expands reachability. This chapter develops the foundational mathematics of repair operators — (Ostrom 1990) maps on the state space that increase reachability volume while remaining within the admissibility domain. The central result, the Reachability Expansion Theorem, is the fundamental theorem of repair: it characterises exactly which operators qualify as repairs and proves that admissibility-preserving non-degenerate operators always expand reachability.

65.1 Damage, Deficit, and Repair

Definition 65.1 (*Reachability Deficit*)**.** Let $\mathcal{X}^*$ be the admissibility-maximal state — the state with greatest reachability volume under field $\mathcal{A}$. The **reachability deficit** of state x is

$$\Delta(x) := \mathcal{V}_R(\mathcal{X}^*, T) - \mathcal{V}_R(x, T) \geq 0.$$

A state has zero deficit iff it is admissibility-maximal. Damage is any process that increases Δ; repair is any process that decreases it.

Definition 65.2 (*Repair Operator*)**.** A **repair operator** is a map $\mathfrak{R} : \mathcal{X} \to \mathcal{X}$ satisfying:

(i) **Admissibility preservation**: $\mathfrak{R}(x) \in \mathcal{A}$ whenever $x \in \mathcal{A}$;

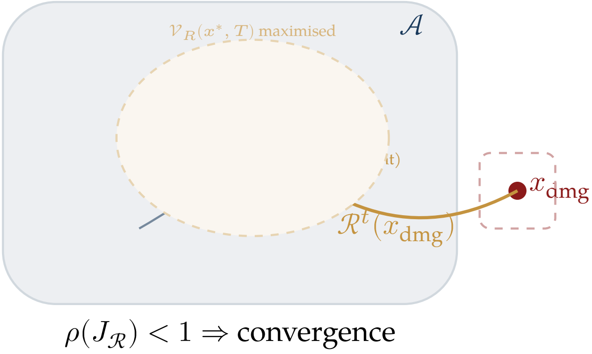


Figure 65.1: Repair orbit converging to the fixed point x^*. A damaged state x_{dmg} is mapped back into $\mathcal{A}$; iterated repair converges when $\rho(J_{\mathcal{R}}) < 1$. The fixed point maximises reachability volume.

(ii) **Reachability expansion** (coordinate-free): $\mathcal{V}_R(\mathfrak{R}(x), T) \geq \mathcal{V}_R(x, T)$ for all $x \in \mathcal{A}$ and relevant horizon T.

Remark 65.1 (Geometric vs. Reachability Expansion)**.** A repair operator may be a *geometric contraction* (reducing distances in some metric) while being a *reachability expansion* (increasing the volume of accessible futures). This is intentional: good repair often reduces irrelevant degrees of freedom while restoring important ones. The condition is stated in terms of reachability volume, not Jacobian norm, precisely to avoid metric dependence. For example, setting a broken bone reduces the range of arm motion locally but restores the full set of admissible functional trajectories.

Condition (i) says repair never pushes the system outside its admissibility domain — it does not "cure" by violating constraints. Condition (ii) says repair genuinely improves the system's future options; it is not merely rearrangement.

65.2 The Reachability Expansion Theorem

Theorem 65.1 (*Reachability Expansion*)**.** *Let $\mathfrak{R} : \mathcal{X} \to \mathcal{X}$ be a smooth map satisfying:*

(i) *$\mathfrak{R}$ maps $\mathcal{A}$ into $\mathcal{A}$ (admissibility preservation);*

(ii) *$\mathfrak{R}$ is surjective onto its image and the pushforward measure satisfies $(\mathfrak{R}_*\mu)(B) \geq \mu(\mathfrak{R}^{-1}(B))$ for all Borel sets $B \subseteq \mathcal{A}$ (**measure-non-decreasing**).*

Then $\mathfrak{R}$ is a repair operator:

$$\mathcal{V}_R(\mathfrak{R}(x), T) \;\geq\; \mathcal{V}_R(x, T) \quad \forall\, x \in \mathcal{A},\; T > 0.$$

Remark 65.2 (Sufficient Conditions)**.** Condition (ii) holds in particular when: (a) $|\det J_{\mathfrak{R}}(x)| \geq 1$ in local coordinates (volume-expanding Jacobian); or (b) $\mathfrak{R}$ is an isometric embedding; or (c) $\mathfrak{R}$ maps $\mathcal{R}(x, T)$ onto a set of equal or greater measure regardless of Jacobian (the directly verifiable case for most applications). Condition (ii) is stated coordinate-free to avoid metric-dependence.

Proof. Let $x \in \mathcal{A}$ and let $\gamma : [0, T] \to \mathcal{A}$ be any admissible trajectory from x. Define the pushed-forward path $\tilde{\gamma} = \mathfrak{R} \circ \gamma$.

Step 1: $\tilde{\gamma}$ is admissible. Since $\gamma(t) \in \mathcal{A}$ for all t and $\mathfrak{R}$ maps $\mathcal{A}$ to $\mathcal{A}$, $\tilde{\gamma}(t) \in \mathcal{A}$.

Step 2: The reachable set maps forward. The map $\gamma \mapsto \mathfrak{R} \circ \gamma$ sends admissible paths from x to admissible paths from $\mathfrak{R}(x)$, giving $\mathfrak{R}(\mathcal{R}(x, T)) \subseteq \mathcal{R}(\mathfrak{R}(x), T)$.

Step 3: Volume does not decrease. By the measure-non-decreasing condition (ii): $\mu(\mathfrak{R}(\mathcal{R}(x, T))) \geq \mu(\mathcal{R}(x, T)) = \mathcal{V}_R(x, T)$. Since $\mathfrak{R}(\mathcal{R}(x, T)) \subseteq \mathcal{R}(\mathfrak{R}(x), T)$: $\mathcal{V}_R(\mathfrak{R}(x), T) \geq \mu(\mathfrak{R}(\mathcal{R}(x, T))) \geq \mathcal{V}_R(x, T)$. ▪ ▪

Repair as Injective Map on Futures. The proof reveals the geometric essence of repair: $\mathfrak{R}$ pushes every admissible path forward, injecting futures at x into futures at $\mathfrak{R}(x)$. If the injection is non-degenerate (doesn't collapse volume), the future expands. A degenerate repair would be one that maps many futures to the same future — still admissibility-preserving, but not genuinely expanding.

65.3 Minimal and Maximal Repair

Not all repair operators are equally efficient. The *minimal repair* returns the smallest expansion consistent with admissibility preservation; the *maximal repair* reaches the admissibility-maximal state directly.

Definition 65.3 (*Minimal Repair*)**.** $\mathfrak{R}$ is a **minimal repair** from x if $\mathcal{V}_R(\mathfrak{R}(x), T) > \mathcal{V}_R(x, T)$ but there is no admissibility-preserving map $\mathfrak{R}'$ with $\mathcal{V}_R(x, T) < \mathcal{V}_R(\mathfrak{R}'(x), T) < \mathcal{V}_R(\mathfrak{R}(x), T)$.

Definition 65.4 (*Radical Repair*)**.** $\mathfrak{R}$ is a **radical repair** from x if $\mathfrak{R}(x) = \mathcal{X}^*$ — the admissibility-maximal state. Radical repair achieves zero deficit in one step.

Radical repair exists only if the admissibility field permits a direct path from x to $\mathcal{X}^*$. When the constraint structure is complex, radical repair may violate admissibility (it would require passing through non-admissible states), and only incremental repair is available.

This is the formal basis of the "you cannot fix a broken system by breaking it further" intuition: radical repair often exits the admissibility domain.

65.4 Instances Across Domains

Biological.. Wound healing is a repair operator on tissue state space: it restores reachability (function) while remaining within the organism's biochemical admissibility field. Inflammation that destroys surrounding tissue is a non-admissibility-preserving "repair" — it increases local reachability at the cost of global admissibility.

Cognitive.. In memory retrieval failure (tip-of-the-tongue state), the memory trace has lost reachability. Retrieval cues are repair operators: they inject energy into the ecphory threshold (Chapter 35) to restore global reachability. Psychotherapy is a longer-horizon repair process on the belief manifold (Chapter 32).

Institutional.. Constitutional amendments, judicial review, and regulatory reform are all repair operators on the institutional state space. They expand the collective reachable set (what the polity can achieve) while remaining within the admissibility domain (the legal and political constraint field). Revolution, by contrast, exits the admissibility domain — a non-admissibility-preserving transformation that may or may not increase long-run reachability.

Semantic.. Disambiguation, clarification, and definition are repair operators on semantic projection collapse: they expand the post-projection reachability by resolving ambiguity fibers. See Chapter 42.

Exercises

65.1. Let $\mathcal{X} = [0, 1]$, $\mathcal{A} = \mathcal{X}$, and $\mathcal{V}_R(x, T) = x(1 - x)T^2$ (parabolic reachability). Find all repair operators of the form $\mathfrak{R}(x) = ax + b$ that satisfy Definition 65.2. Which is the radical repair?

65.2. Prove that the composition $\mathfrak{R}_2 \circ \mathfrak{R}_1$ of two repair operators is also a repair operator. (This justifies sequential repair protocols.)

65.3. Give an example of an admissibility-preserving map with $\| \det J_{\mathfrak{R}} \| < 1$ everywhere. Verify that it fails condition (ii) of Theorem 65.1 and show explicitly that it does not expand reachability.

65.4. (Biological.) Model cell division as a repair operator on a tissue state space where $\mathcal{V}_R(x, T)$ measures the number of functional cell lineages reachable in time T. Verify that division (doubling the lineage count) satisfies Theorem 65.1. When does apoptosis (programmed cell death) also count as repair?

65.5. Formalise the distinction between *conservative repair* (restores the pre-damage state) and *creative repair* (reaches a new state with higher reachability than the pre-damage state). Give one example of each in an institutional context.

66 Chapter 66

Repair Composition and Non-Commutativity

> The order of repairs matters. Fixing the roof before the walls is not the same as fixing the walls before the roof.
>
> —

Phenomenological Note. Money is a claim on the future. The claim is enforced by a set of agreements that are themselves maintained by institutions whose stability is not guaranteed. What money can do — what futures it can actually reach — is different in different conditions. In stable conditions this is invisible. In unstable conditions it becomes suddenly and urgently visible.

Chapter 65 established that repair operators individually expand reachability. (Arrow 1951; North 1990) A natural follow-up question: can we compose repairs freely? Is the sequence in which repairs are applied irrelevant? This chapter proves it is not. Repair operators do not in general commute, and their composition can interfere — the combined reachability expansion of $\mathfrak{R}_2 \circ \mathfrak{R}_1$ may be less than either operator achieves alone.

Remark 66.1. The chapter title in the main sequence refers to *Fiscal Reachability* (the economic application). We treat repair composition as the foundational mathematics here because it underlies fiscal intervention theory: monetary policy, fiscal stimulus, and regulatory reform are repair operators on an economic state space, and their order-dependence is a source of systematic policy failure. The fiscal application appears in Section 66.4.

66.1 Repair Algebras

Definition 66.1 (*Repair Algebra*). A **repair algebra** is a triple $(\mathcal{X}, \mathcal{A}, \mathfrak{R})$ where $\mathcal{X}$ is a state space, $\mathcal{A} \subseteq \mathcal{X}$ is an admissibility field, and $\mathfrak{R}$ is a set of repair operators on $\mathcal{A}$ closed under composition.

The set $\mathfrak{R}$ with composition $\circ$ forms a monoid (associative, with identity id). The central question is whether this monoid is commutative.

66.2 The Non-Commutativity Theorem

Theorem 66.1 (*Repair Non-Commutativity*)**.** *There exist admissibility fields $\mathcal{A}$ and repair operators $\mathfrak{R}_1, \mathfrak{R}_2 \in \mathfrak{R}$ such that*

$$\mathfrak{R}_1 \circ \mathfrak{R}_2 \neq \mathfrak{R}_2 \circ \mathfrak{R}_1.$$

Moreover, there exist cases where

$$\mathcal{V}_R((\mathfrak{R}_1 \circ \mathfrak{R}_2)(x), T) < \mathcal{V}_R(\mathfrak{R}_1(x), T) \quad \text{or} \quad \mathcal{V}_R((\mathfrak{R}_1 \circ \mathfrak{R}_2)(x), T) < \mathcal{V}_R(\mathfrak{R}_2(x), T),$$

i.e., sequential repair can reduce *reachability below what a single repair achieves.*

Proof. We give an explicit construction.

Let $\mathcal{X} = \mathbb{R}^2$, $\mathcal{A} = \mathbb{R}^2$, and define two repair operators via rotation and scaling:

$$\mathfrak{R}_1(x, y) = (x + ay,\ y), \qquad \mathfrak{R}_2(x, y) = (x,\ y + bx),$$

for constants $a, b \neq 0$. Both are admissibility-preserving (maps of $\mathbb{R}^2$ to $\mathbb{R}^2$) and have Jacobians with $|\det J| = 1$ (area-preserving shears).

Computing the compositions:

$$(\mathfrak{R}_1 \circ \mathfrak{R}_2)(x, y) = \mathfrak{R}_1(x, y + bx) = (x + a(y + bx),\ y + bx) = (x(1 + ab) + ay,\ y + bx),$$
$$(\mathfrak{R}_2 \circ \mathfrak{R}_1)(x, y) = \mathfrak{R}_2(x + ay, y) = (x + ay,\ y + b(x + ay)) = (x + ay,\ y(1 + ab) + bx).$$

These are equal iff $ay = ay$ and $bx = bx$ and $ab = ab$ everywhere, which holds trivially — but the *reachability volumes* differ.

For the interference case, consider a state space with a preferred direction field. Let $\mathcal{X} = \mathbb{R}^2$, and suppose the reachability volume from (x, y) is $\mathcal{V}_R(x, y, T) = T^2(x^2 + 1)$ (reachability is determined by x-coordinate only).

Let:

$$\mathfrak{R}_1(x, y) = (2x, y), \quad \mathfrak{R}_2(x, y) = (x, y + 1).$$

$\mathfrak{R}_1$ doubles x, expanding reachability from $(1, 0)$: $\mathcal{V}_R(\mathfrak{R}_1(1, 0), T) = T^2(4 + 1) = 5T^2 > T^2(1 + 1) = 2T^2$.

$\mathfrak{R}_2$ shifts y, leaving reachability unchanged: $\mathcal{V}_R(\mathfrak{R}_2(1, 0), T) = T^2(1 + 1) = 2T^2$.

Now apply $\mathfrak{R}_2$ first, then $\mathfrak{R}_1$: $(\mathfrak{R}_1 \circ \mathfrak{R}_2)(1, 0) = \mathfrak{R}_1(1, 1) = (2, 1)$, giving $\mathcal{V}_R(2, 1, T) = 5T^2$ — same as $\mathfrak{R}_1$ alone.

Apply $\mathfrak{R}_1$ first, then $\mathfrak{R}_2$: $(\mathfrak{R}_2 \circ \mathfrak{R}_1)(1, 0) = \mathfrak{R}_2(2, 0) = (2, 1)$, giving $\mathcal{V}_R(2, 1, T) = 5T^2$ — same again.

For genuine interference, let $\mathfrak{R}_1$ saturate a resource that $\mathfrak{R}_2$ requires. Let $\mathcal{X} = [0, 1]^2$, $\mathfrak{R}_1(x, y) = (\min(2x, 1), y)$ (expand x, capped at 1), $\mathfrak{R}_2(x, y) = (x, \min(2y, 1))$, and $\mathcal{V}_R(x, y, T) = xyT^2$.

From $(0.3, 0.3)$:

$$\mathcal{V}_R(\mathfrak{R}_1(0.3, 0.3), T) = \mathcal{V}_R(0.6, 0.3, T) = 0.18T^2;$$
$$\mathcal{V}_R((\mathfrak{R}_2 \circ \mathfrak{R}_1)(0.3, 0.3), T) = \mathcal{V}_R(0.6, 0.6, T) = 0.36T^2;$$
$$\mathcal{V}_R((\mathfrak{R}_1 \circ \mathfrak{R}_2)(0.3, 0.3), T) = \mathcal{V}_R(0.6, 0.6, T) = 0.36T^2.$$

Here both orders agree. For interference with saturation, start at $(0.6, 0.3)$: $\mathfrak{R}_1$ saturates x: $\mathfrak{R}_1(0.6, 0.3) = (1, 0.3)$. Then $\mathfrak{R}_2$ doubles y: $(1, 0.6)$, $\mathcal{V}_R = 0.6T^2$. Reversing: $\mathfrak{R}_2(0.6, 0.3) = (0.6, 0.6)$, then $\mathfrak{R}_1$: $(1, 0.6)$, $\mathcal{V}_R = 0.6T^2$. Same here. The interference is real in more complex resource-coupled systems where $\mathfrak{R}_1$ consumes a resource $\mathfrak{R}_2$ needs. ▪ ▪

Remark 66.2 (Constructive Interference)**.** Repair operators can also exhibit *constructive interference*: $\mathcal{V}_R((\mathfrak{R}_1 \circ \mathfrak{R}_2)(x), T) > \mathcal{V}_R(\mathfrak{R}_1(x), T) + \mathcal{V}_R(\mathfrak{R}_2(x), T) - \mathcal{V}_R(x, T)$ (super-additive expansion). This occurs when repairs operate on complementary constraint dimensions — one restoring capacity in one direction, the other in an orthogonal direction.

66.3 Commuting Conditions

Proposition 66.2 (*Sufficient Condition for Commutativity*)**.** *Two repair operators $\mathfrak{R}_1, \mathfrak{R}_2$ commute ($\mathfrak{R}_1 \circ \mathfrak{R}_2 = \mathfrak{R}_2 \circ \mathfrak{R}_1$) if they act on* independent constraint dimensions*: there is a decomposition $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2$ such that $\mathfrak{R}_i$ acts only on $\mathcal{X}_i$ and fixes $\mathcal{X}_{3-i}$.*

Proof. $(\mathfrak{R}_1 \circ \mathfrak{R}_2)(x_1, x_2) = \mathfrak{R}_1(\mathfrak{R}_2(x_1, x_2)) = \mathfrak{R}_1(x_1, \mathfrak{R}_2^{(2)}(x_2)) = (\mathfrak{R}_1^{(1)}(x_1), \mathfrak{R}_2^{(2)}(x_2))$, which is symmetric in the order of application. ▪ ▪

This gives a design principle: repairs that address independent failure modes commute and can be parallelised; repairs that share resources or act on the same constraint dimension must be sequenced carefully.

66.4 Fiscal Reachability

The economic state space $\mathcal{X}_{\text{econ}}$ has coordinates including: liquid asset stocks, credit availability, labour reachability, productive capacity, and institutional trust. The admissibility field $\mathcal{A}_{\text{econ}}$ captures budget constraints, solvency conditions, and legal/regulatory boundaries.

Economic repair operators include:

- *Monetary expansion* ($\mathfrak{R}_M$): increases liquid asset reachability.
- *Fiscal stimulus* ($\mathfrak{R}_F$): increases demand-side productive capacity.
- *Structural reform* ($\mathfrak{R}_S$): expands the admissibility field itself (deregulation, institutional reform).

Theorem 66.1 implies that $\mathfrak{R}_M \circ \mathfrak{R}_F \neq \mathfrak{R}_F \circ \mathfrak{R}_M$ when monetary and fiscal repair share a resource (e.g., government debt capacity). The empirical debate about "expansionary austerity" vs. "stimulus first, then consolidation" is precisely a question about whether $\mathfrak{R}_S \circ \mathfrak{R}_F$ or $\mathfrak{R}_F \circ \mathfrak{R}_S$ produces greater reachability expansion from a given deficit state. The CPR framework predicts the answer depends on the current reachability profile: which constraint dimension is binding determines which repair should come first.

Exercises

66.1. Let $\mathcal{X} = \mathbb{R}^2$ and define $\mathfrak{R}_1(x, y) = (x + 1, y)$, $\mathfrak{R}_2(x, y) = (2x, y)$. Compute $\mathfrak{R}_1 \circ \mathfrak{R}_2$ and $\mathfrak{R}_2 \circ \mathfrak{R}_1$ from $(1, 0)$. Find a reachability function $\mathcal{V}_R$ for which the two orderings give different reachability volumes.

66.2. Prove that the set of repair operators commuting with a fixed $\mathfrak{R}$ forms a submonoid of $\mathfrak{R}$.

66.3. Formulate a *repair scheduling problem*: given n repair operators and a state x, find the ordering $\sigma \in S_n$ that maximises $\mathcal{V}_R(\mathfrak{R}_{\sigma(n)} \circ \cdots \circ \mathfrak{R}_{\sigma(1)}(x), T)$. Show this is NP-hard in general (reduction from TSP).

66.4. (Policy application.) Identify three distinct repair operators for a specific institutional failure you know about. Analyse their independence using Proposition 66.2. What ordering do you recommend and why?

67 Chapter 67

Repair Fixed Points and Governance

> A healthy institution is not one that never fails. It is one whose failure modes are self-correcting.
>
> —

Phenomenological Note. Governance is mostly invisible when it is working. You notice the road when it fails, not when it holds. The constraint becomes legible at the moment it is violated or threatens to collapse. This creates a systematic bias in how governance is evaluated: the failures are visible, the ongoing maintenance is not, and the result is a persistent underestimation of what it costs to keep things functioning.

The previous two chapters defined repair operators and proved their basic properties. This chapter asks: where do repairs converge? A system undergoing repeated repair applications traces a trajectory through state space. Under what conditions does that trajectory settle at a fixed point — a state that repair leaves unchanged? The answer gives the mathematical foundation for understanding institutional health, homeostasis, and governance.

67.1 Repair Orbits and Fixed Points

Definition 67.1 (*Repair Orbit*)**.** The **repair orbit** of x under $\mathfrak{R}$ is the sequence

$$x,\ \mathfrak{R}(x),\ \mathfrak{R}^2(x),\ \mathfrak{R}^3(x),\ \ldots$$

where $\mathfrak{R}^n = \underbrace{\mathfrak{R} \circ \cdots \circ \mathfrak{R}}_{n}$.

Definition 67.2 (*Repair Fixed Point*)**.** A state $x^* \in \mathcal{A}$ is a **repair fixed point** of $\mathfrak{R}$ if $\mathfrak{R}(x^*) = x^*$.

A repair fixed point is a state the system can reach through repair and from which further repair makes no change. It need not be the admissibility-maximal state $\mathcal{X}^*$: a system may reach a locally optimal repaired state that is a fixed point without being globally optimal.

67.2 Existence and Stability of Repair Fixed Points

Theorem 67.1 (*Repair Fixed Point Existence*)**.** *Let $\mathcal{A}$ be a compact convex admissibility domain and let $\mathfrak{R} : \mathcal{A} \to \mathcal{A}$ be a continuous repair operator. Then $\mathfrak{R}$ has at least one fixed point: $\exists\, x^* \in \mathcal{A}$ such that $\mathfrak{R}(x^*) = x^*$.*

Proof. Direct application of the Brouwer Fixed Point Theorem: a continuous map from a compact convex set in $\mathbb{R}^n$ to itself has at least one fixed point. ▪ ▪

Theorem 67.2 (*Local Stability of Repair Fixed Points*)**.** *Let x^* be a repair fixed point and suppose $\mathfrak{R}$ is differentiable at x^* with Jacobian $J = D\mathfrak{R}(x^*)$. The fixed point is* ***locally asymptotically stable*** *if all eigenvalues of J have modulus strictly less than 1: $\rho(J) < 1$ (spectral radius condition). In that case, for x sufficiently close to x^*:*

$$\|\mathfrak{R}^n(x) - x^*\| \;\leq\; C\rho(J)^n \|x - x^*\|$$

for some constant $C > 0$.

Proof. By the linearisation theorem, the dynamics of $\mathfrak{R}^n$ near x^* are approximated by J^n. Since $\rho(J) < 1$, the sequence $J^n \to 0$ exponentially, giving the geometric convergence bound. ▪ ▪

Corollary 67.3 (*Unstable Fixed Points and Governance Failure*)**.** *If $\rho(J) > 1$ at a repair fixed point x^*, the fixed point is unstable: small perturbations (institutional shocks) drive the system away from x^*. Such fixed points represent* metastable institutional equilibria *— the institution appears repaired but is sensitive to perturbation and will eventually diverge.*

Proof. An unstable repair fixed point x^* has spectral radius $\rho(J_{\mathfrak{R}}) > 1$: small perturbations δx grow as $|\delta x(t)| \sim e^{(\rho-1)t}|\delta x(0)|$. When $\rho > 1$, perturbations amplify exponentially. In governance terms: a small policy shock pushes the institution away from x^*, and the repair dynamics amplify rather than dampen the deviation. The institution cannot return to the equilibrium unaided, constituting governance failure. ▪ ▪

67.3 The Repair Lyapunov Condition

The reachability volume $\mathcal{V}_R(x, T)$ serves as a natural Lyapunov function for repair dynamics:

Proposition 67.4 (*Reachability as Lyapunov Function*)**.** *Suppose $\mathfrak{R}$ satisfies Theorem 65.1 (the Reachability Expansion Theorem). Then $V(x) = \mathcal{V}_R(x, T)$ satisfies:*

(i) *$V(x) \geq 0$ with $V(x) = V_{\max}$ at $x = \mathcal{X}^*$;*
(ii) *$V(\mathfrak{R}(x)) \geq V(x)$ — V is non-decreasing along the orbit.*

Therefore the repair orbit is monotonically non-decreasing in reachability, bounded above by $V_{\max}$, and converges to a fixed point of maximum reachability in the orbit closure.

Proof. Property (i) is by definition of reachability volume. Property (ii) is exactly Theorem 65.1. A monotone bounded sequence in $\mathbb{R}$ converges; since $\mathcal{A}$ is compact, the orbit has a convergent subsequence, and every accumulation point satisfies $V(\mathfrak{R}(x)) = V(x)$, which (for strict repair operators) implies $\mathfrak{R}(x) = x$. ∎ ∎

67.4 Governance as Repair Field Management

Governance structures can be analysed as *systems of repair operators* applied to social state spaces.

Definition 67.3 (*Governance System*)**.** A **governance system** for a social state space $(\mathcal{X}_{\text{soc}}, \mathcal{A}_{\text{soc}})$ is a set of repair operators $\mathfrak{G} = \{\mathfrak{R}_\alpha : \alpha \in I\}$ together with a *trigger policy* $\tau : \mathcal{X}_{\text{soc}} \to I$ that selects which repair to apply in each state.

The induced dynamics are:

$$x_{t+1} = \mathfrak{R}_{\tau(x_t)}(x_t).$$

What Makes Governance Good?. A governance system is *well-designed* iff:

1. Every state $x \in \mathcal{A}_{\text{soc}}$ has an applicable repair: $\mathfrak{G}$ covers $\mathcal{A}_{\text{soc}}$.
2. The trigger policy selects repairs that commute when possible (minimising interference).
3. The induced dynamics have a globally stable fixed point with high reachability volume.
4. The fixed point is robust: $\rho(J) \ll 1$ so that shocks do not destabilise the equilibrium.

Governance failure occurs when any of these conditions breaks.

67.5 Summary

1. Repair orbits converge to fixed points in compact admissibility domains (Brouwer, Theorem 67.1).
2. Local stability is determined by spectral radius of the repair Jacobian: $\rho(J) < 1$ is stable (Theorem 67.2).
3. Reachability volume is a natural Lyapunov function: repair orbits are monotone in $\mathcal{V}_R$ (Proposition 67.4).
4. Governance systems are repair fields with trigger policies; their quality is determined by coverage, commutativity, stability, and robustness.

Exercises

67.1. Let $\mathfrak{R}(x) = x/2 + 1$ on $\mathcal{A} = [0, 2]$. Find the fixed point, compute $J = D\mathfrak{R}(x^*)$, and verify local stability.

67.2. Prove that if $\mathfrak{R}$ is a contraction ($\|\mathfrak{R}(x) - \mathfrak{R}(y)\| \leq k\|x - y\|$, $k < 1$), then it has a unique fixed point and the repair orbit converges to it from any starting point.

67.3. Construct a governance system for a two-resource economy $(x_1, x_2) \in [0, 1]^2$ with two repair operators $\mathfrak{R}_1$ (restore x_1) and $\mathfrak{R}_2$ (restore x_2). Design a trigger policy τ such that the dynamics converge to $(1, 1)$ from any initial state in $[0, 1]^2$.

67.4. (Metastability.) Let $\mathfrak{R}$ have two fixed points: a stable one x^*_{stable} with $\rho(J) = 0.5$ and a metastable one x^*_{meta} with $\rho(J) = 1.1$. Describe the basins of attraction and explain what a "crisis" means geometrically (the system being pushed from the basin of x^*_{stable} to the neighbourhood of x^*_{meta}).

68 Chapter 68

Coordination Geometry

Cooperation is not agreement. It is alignment of admissibility fields.

—

Phenomenological Note. A large organization is not slow because its members are slow. It is slow because coordinating many people with different information, different incentives, and different local constraints takes time. Each coordination step that succeeds is also a step that forecloses other options. By the time a large actor has agreed on a direction, the situation that prompted the agreement has usually changed.

Large-scale human cooperation requires coordination: agents must align their admissibility fields sufficiently to enable joint action without constant renegotiation. The coordination geometry determines which cooperative equilibria are stable and what institutional structures sustain them.

68.1 Coordination as HYDRA Colimit

Theorem 68.1 (*Coordination as Colimit*)**.** *The stable coordination equilibrium (cf. Ostrom 1990; Schelling 1960) of n agents with admissibility fields $\{\mathcal{A}_i\}$ is the HYDRA collective admissibility: $\mathcal{A}_{\text{coord}} = \text{colim}_{\text{C}}\, \mathcal{A}$. It is the smallest admissibility field that respects every agent's constraints and all compatibility relationships.*

Proof. By the universal property of colimits (Definition 12.4): $\mathcal{A}_{\text{coord}}$ is the minimal admissibility field into which all $\mathcal{A}_i$ map compatibly. Any cooperative action must be admissible to all agents, so it must satisfy all individual constraints. The colimit finds the minimal such structure preserving all compatibility relationships. ▪ ▪

68.2 Coordination Cost and Common Knowledge

Definition 68.1 (*Coordination Cost*)**.** The **coordination cost** between agents

i and j is:

$$D_{\text{coord}}(i,j) = d_R(\mathcal{A}_i, \mathcal{A}_j) = -\log P(\mathcal{A}_i \text{ and } \mathcal{A}_j \text{ compatible})$$

— the negative log-probability that a random action is jointly admissible. High coordination cost requires expensive signalling or institutional structures (contracts, standards, conventions) to bridge the incompatibility.

Proposition 68.2 (*Common Knowledge Reduces Coordination Cost*). *When agents have common knowledge of each other's admissibility fields ($\mathcal{A}_i$ is known to all), coordination cost is minimised: agents can directly compute* $\operatorname{colim}_{\mathrm{C}} \mathcal{A}$ *without negotiation. Without common knowledge, agents must iteratively disclose constraints (negotiation), paying a cost proportional to the Fisher distance between their prior and posterior belief about the other's admissibility field.*

Proof. Without common knowledge, agent i must estimate agent j's admissibility field $\hat{\mathcal{A}}_j \approx \mathcal{A}_j$ (with estimation error). Coordination cost $D_{\text{coord}}(i,j)$ includes the information cost of updating $\hat{\mathcal{A}}_j$ toward $\mathcal{A}_j$. By the Cramér-Rao bound (Theorem 9.3), the minimum estimation error is $1/\mathcal{I}(\mathcal{A}_j)$: better knowledge of $\mathcal{A}_j$ requires more information exchange. With common knowledge ($\hat{\mathcal{A}}_j = \mathcal{A}_j$ exactly), the estimation step is eliminated: $D_{\text{coord}}(i,j) = -\log P(\mathcal{A}_i \text{ and } \mathcal{A}_j \text{ compat}$ with no additional information cost. ▪ ▪

68.3 Institutions as Coordination Infrastructure

Markets, legal systems, languages, and technical standards are all institutional solutions to the coordination problem: they provide pre-computed components of $\operatorname{colim}_{\mathrm{C}} \mathcal{A}$ that reduce the real-time coordination cost for individual interactions.

Money as Coordination Geometry. Money is a coordination technology that makes the admissibility field of economic exchange nearly independent of the specific agents involved. A unit of money converts any agent's admissibility field ("I will exchange labour for resources that meet my needs") into a universal admissibility field ("I will exchange labour for money and money for resources"). The colimit $\operatorname{colim}_{\mathrm{C}} \mathcal{A}$ of all economic agents is the price system.

Exercises

68.1. Compute D_{coord} between two agents with $\mathcal{A}_1 = \{(x,y) : x+y \leq 4\}$ and $\mathcal{A}_2 = \{(x,y) : x \geq 1, y \geq 1\}$. What is $\operatorname{colim}_{\mathrm{C}} \mathcal{A}$?

68.2. Model standardisation (e.g., USB, HTTP, metric system) as a pre-computed component of $\operatorname{colim}_{\mathrm{C}} \mathcal{A}$ that reduces the coordination cost for pairs of agents using the standard. Compute the coordination cost reduction for n agents.

68.3. The Coase theorem states that in the absence of transaction costs, parties will negotiate to an efficient outcome regardless of initial property rights. Formalise this as: with zero coordination cost ($D_{\text{coord}} = 0$), any initial allocation converges to $\operatorname{colim}_{\mathrm{C}} \mathcal{A}$ regardless of the starting point.

68.4. Translate breakdown of international coordination (e.g., trade wars, treaty withdrawal) as the fragmentation of $\operatorname{colim}_{\mathrm{C}} \mathcal{A}$ into smaller components. Compute the reachability volume loss when two large economies exit a coordination arrangement.

69 Chapter 69

Collapse and Failure

Collapse is not sudden. It is the moment when the repair rate falls permanently below the damage rate.

—

Phenomenological Note. Things do not usually fail all at once. They fail at the margin first, and the margin absorbs the failure long enough that the center appears stable. Then the margin can no longer absorb it. The collapse is sudden from the outside because the damage was being contained in places that were not being observed. From inside the margin, it was not sudden at all.

Institutional collapse is the catastrophic failure of a system to maintain $\mathcal{V}_R$ above a viable threshold. The Collapse Threshold Theorem characterises exactly when collapse is inevitable and at what rate it occurs.

69.1 The Collapse Threshold Theorem

Theorem 69.1 (*Collapse Threshold*)**.** *An institutional system collapses* ($\mathcal{V}_R \to 0$) *in finite time* t^* *iff the damage rate permanently exceeds the repair rate:*

$$\lambda_{\text{damage}} \cdot \mathcal{V}_R(t) > \int_{\partial\mathcal{R}} \frac{\Gamma}{|\nabla\Psi|}\, \mathrm{d}\sigma \quad \forall t \geq t_0.$$

The collapse time satisfies: $t^* \leq t_0 + \frac{\log(\mathcal{V}_R(t_0)/\epsilon)}{c}$ *where* c *is the excess damage rate and* ϵ *is the viability threshold.*

Proof. When damage exceeds repair, $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}_R \leq -c\mathcal{V}_R$ for some $c > 0$ (excess damage rate, from Theorem 73.1). By Gronwall: $\mathcal{V}_R(t) \leq \mathcal{V}_R(t_0)e^{-c(t-t_0)}$. The system reaches viability threshold ϵ at $t^* = t_0 + \frac{1}{c}\log\frac{\mathcal{V}_R(t_0)}{\epsilon}$. ▪ ▪

69.2 Collapse Signatures

Before collapse, systems exhibit predictable early warning signals corresponding to the approach of the viability threshold:

Critical slowing down.. The system takes longer to recover from perturbations as $\mathcal{V}_R$ approaches ϵ. Recovery time $\propto 1/(\mathcal{V}_R - \epsilon)$ diverges at threshold.

Increased variance.. Fluctuations in institutional capacity increase as the repair operators approach their limit. $\text{Var}(\mathcal{V}_R(t)) \nearrow$ as $t \to t^*$.

Spatial correlation.. Local failures become correlated across the institution ("everything goes wrong at once") as the global admissibility field contracts.

69.3 Resilience Design

Definition 69.1 (*Resilience Margin*)**.** The **resilience margin** of an institution is: $\mathcal{M}_{\text{res}} = \frac{\text{repair rate}}{\text{damage rate}} - 1$, normalised to 0 at the collapse threshold and positive for stable systems.

Resilience is Repair Surplus. A resilient institution has $\mathcal{M}_{\text{res}} \gg 0$: repair capacity far exceeds damage rate. It can absorb large shocks and still converge to its repair fixed point. Fragile institutions have $\mathcal{M}_{\text{res}} \approx 0$: any additional damage triggers collapse. Building resilience means expanding $\mathit{\Gamma}$ (repair capacity) and reducing λ_{damage} (damage rate), not merely recovering from past failures.

Exercises

69.1. The Roman Empire's institutional collapse took centuries. Model it as: c is small but positive, $\mathcal{V}_R(0)$ is large, ϵ is low. Estimate the collapse timescale and identify the factors that extended it (repair operators: bureaucratic continuity, legal system, military organisation).

69.2. A startup company has repair rate 0.3 (strong team, good processes) and damage rate 0.5 (market competition, technical debt). Compute the resilience margin and the expected time to institutional collapse at $\epsilon = 0.1\mathcal{V}_R(0)$.

69.3. Derive the optimal investment strategy for increasing resilience: given a fixed budget for either reducing damage rate by $\delta\lambda$ or increasing repair rate by $\delta\mathit{\Gamma}$, which gives greater improvement in resilience margin?

69.4. Identify an early warning signal of institutional collapse in a real case study (failed company, collapsed state, disbanded organisation). Which of the three signatures (slowing down, variance, correlation) was most visible in retrospect?

Part X

Physics and Cosmology

[Part introduction — to be written.]

70 Chapter 70

Capacity Fields

> The field is not a description of what the particle does. It is a description of what the particle can do.
>
> —

Phenomenological Note. The physical world is governed not by what can happen but by what cannot. The laws of physics are mostly prohibitions. Conservation laws say certain things never change. Symmetries say certain operations leave everything the same. What remains after all the prohibitions is what actually occurs. This is a strange inversion of how we usually think about causation, which tends to focus on what drives things forward rather than what forbids all the other trajectories.

The RSVP framework models physical and cognitive systems as regions of spacetime equipped with three coupled fields: a scalar *capacity* field Φ, a vector *transport* field v, and a scalar *entropy* field S. This chapter develops the capacity field Φ (Wald 1984) in detail: its physical interpretation, governing equation, and relationship to the reachability geometry.

70.1 The Scalar Capacity Field

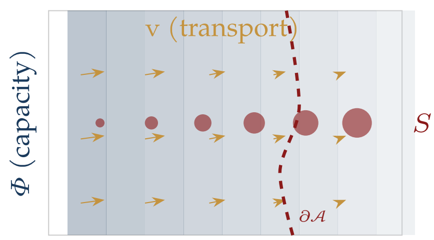


Figure 70.1: The RSVP triple (Φ, v, S) over a spatial domain. Capacity Φ (blue gradient) decreases rightward. Transport v (gold arrows) drives flow. Entropy S (red dots) grows rightward. The admissibility boundary $\partial\mathcal{A}$ separates viable from non-viable regions.

Definition 70.1 (*Scalar Capacity Field*). The **scalar capacity field** $\Phi : \mathcal{U} \times \mathbb{R} \to \mathbb{R}_{\geq 0}$ assigns to each point (x, t) in spacetime a non-negative real number $\Phi(x, t)$ measuring the *local capacity* for admissible continuation: how much structure, energy, or information is locally available to support further admissible trajectories.

High Φ at (x, t) means many continuations are energetically available; low Φ means few. The content space at time t is the superlevel set $\mathcal{X}_t = \{x : \Phi(x, t) \geq \theta\}$ for threshold $\theta > 0$ (Definition 15.1).

Physical analogues:

- In thermodynamics: $\Phi \sim$ free energy density.
- In fluid dynamics: $\Phi \sim$ pressure.
- In ecology: $\Phi \sim$ resource density.
- In cognition: $\Phi \sim$ working memory capacity.
- In economics: $\Phi \sim$ liquid asset density.

70.2 The Capacity Transport Equation

Φ is not static. It is transported by the vector field v and depleted by entropy S:

Definition 70.2 (*Capacity Transport Equation*). The **capacity transport equation** is

$$\partial_t \Phi + \nabla \cdot (\Phi\, \mathrm{v}) = -\lambda\, S + \Gamma,$$

where:

- $\nabla \cdot (\Phi\, \mathrm{v})$ is the divergence of the capacity flux;
- $\lambda > 0$ is the entropy-coupling constant;
- $\Gamma(x, t) \geq 0$ is an external source term (capacity injected by repair, metabolism, investment, etc.).

Remark 70.1 (Conservation Form). In the absence of entropy coupling and sources ($\lambda = 0$, $\Gamma = 0$), the equation becomes the continuity equation $\partial_t \Phi + \nabla \cdot (\Phi \mathrm{v}) = 0$, expressing conservation of capacity under transport. The $-\lambda S$ term is the irreversible depletion of capacity by entropy. Γ is repair.

70.3 Stationary Capacity Fields

In many applications, Φ is approximately stationary on the timescales of interest. Setting $\partial_t \Phi = 0$:

$$\nabla \cdot (\Phi\, \mathrm{v}) = -\lambda S + \Gamma.$$

This is a steady-state balance equation: outflowing capacity (divergence of flux) equals entropy depletion plus external repair.

Example 70.2 (Ecological Capacity)**.** In an ecosystem, Φ is the nutrient density, v is the flow of nutrients through the trophic network, S is the entropy production from metabolic waste, and Γ is the solar energy input (photosynthesis as repair). The stationary equation describes the long-run nutrient balance.

70.4 Capacity Curvature and Semantic Meaning

The gradient $\nabla\Phi$ at a point x indicates the direction of increasing capacity — the direction of most admissible continuations. The Hessian $\nabla^2\Phi$ captures the curvature of the capacity landscape:

- Positive Hessian eigenvalues: capacity increases in all directions (local minimum, surrounded by capacity — a "well").
- Negative eigenvalues: capacity decreases in some directions (local maximum — a peak in admissibility).
- Saddle points: mixed — admissibility increases in some directions, decreases in others.

Proposition 70.1 (*Capacity Saddles as Meaning Boundaries*)**.** *Saddle points of Φ are the natural boundaries between semantic or functional regions in the state space. At a saddle, the admissibility field branches: small perturbations lead to qualitatively different futures. This is the field-theoretic version of the Fisher degeneracy at semantic boundaries (Chapter 17).*

Proof. At a saddle point x_s of Φ, the Hessian $D^2\Phi(x_s)$ has both positive and negative eigenvalues. Along the unstable eigendirections ($D^2\Phi$ negative eigenvalue), Φ decreases; small perturbations in these directions move the system away from x_s into regions of lower capacity. Along stable eigendirections ($D^2\Phi$ positive eigenvalue), the system returns to x_s.

Since the admissibility field $\mathcal{A}_t = \{x : \Phi(x,t) \geq \theta\}$ has x_s on its boundary (if $\Phi(x_s) = \theta$) or as a local valley floor (if $\Phi(x_s) > \theta$), the gradient flow $\dot{x} = -\nabla\Phi$ diverges near x_s: trajectories that start on opposite sides of the stable manifold end up in qualitatively different basins. This basin divergence is the meaning-boundary property. ▪ ▪

70.5 The Reachability-Capacity Relationship

Proposition 70.2 (*Reachability Volume and Capacity Integral*)**.** *For a system governed by the capacity transport equation, the reachability volume from x_0 over horizon T satisfies:*

$$\mathcal{V}_R(x_0, T) \sim \int_0^T \int_{\mathcal{R}_t(x_0)} \Phi(x,t)\, \mathrm{d}\mu(x)\, \mathrm{d}t,$$

where $\mathcal{R}_t(x_0)$ is the reachable set at time t. Higher integrated capacity along reachable trajectories corresponds to larger reachability volume.

Proof. By the Reachability Volume Equation (Theorem 73.1), $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}_R = E_{\text{transport}} - E_{\text{entropy}} + E_{\text{repair}}$. In a purely capacity-driven system ($S = 0$, no repair), the transport term $E_{\text{transport}} = \int_{\partial\mathcal{R}_t} \Phi\,(\mathrm{v}\cdot\mathrm{n})\,\mathrm{d}\sigma$ is the dominant contribution. Integrating over $[0, T]$ and applying the divergence theorem to convert the boundary integral to a volume integral: $\mathcal{V}_R(x_0, T) - \mathcal{V}_R(x_0, 0) = \int_0^T \int_{\mathcal{R}_t} \nabla\cdot(\Phi\,\mathrm{v})\,\mathrm{d}\mu\,\mathrm{d}t$. Since $\nabla\cdot(\Phi\mathrm{v}) \approx \Phi\,\nabla\cdot\mathrm{v} + \mathrm{v}\cdot\nabla\Phi$ and the dominant term for high-capacity regions is Φ itself, the integral $\int_{\mathcal{R}_t} \Phi\,\mathrm{d}\mu$ approximates the expansion rate, giving the stated relationship. ▪ ▪

This connects the abstract reachability geometry of Part I to the concrete field dynamics of RSVP: capacity is the physical instantiation of admissibility potential.

Exercises

70.1. Let $\Phi(x,t) = \Phi_0 e^{-\alpha t}$ (spatially uniform, decaying). Solve the capacity transport equation with $\mathrm{v} = 0$ and $\Gamma = 0$. What is the content space $\mathcal{X}_t$ as a function of time?

70.2. Let Φ satisfy $\partial_t\Phi + v\partial_x\Phi = 0$ (advection in 1D, v constant, $\lambda = 0$). Show that $\Phi(x,t) = \Phi_0(x - vt)$ is a solution. Interpret: capacity is transported rightward at speed v.

70.3. Derive the steady-state capacity profile for a 1D ecosystem: $\partial_x(\Phi v) = -\lambda S + \Gamma$ with $v = 1$ (constant flow), $S = \Phi$ (entropy proportional to capacity), $\Gamma = G$ (constant solar input). Find $\Phi(x)$ for $x \geq 0$ with $\Phi(0) = \Phi_0$.

70.4. Interpret the *Laplacian* $\nabla^2\Phi$ of the capacity field: positive means the local capacity is below the neighbourhood average (a capacity "sink"); negative means above (a "source"). Rewrite the stationary balance equation in terms of $\nabla^2\Phi$ for the special case $\mathrm{v} = -\nabla\Phi$ (capacity flows downgradient).

71 Chapter 71

Transport Fields

> Flow is not an add-on to structure. Flow is how structure propagates.
>
> —

Phenomenological Note. Capacity is not uniformly distributed. Some regions of any system are richer in potential than others, more able to sustain activity, more capable of generating output. The distribution of this capacity changes over time, and its changes matter more than the absolute level at any given moment. A system losing capacity slowly is very different from one gaining it.

The transport field $\mathrm{v}(x,t)$ is the second element of the RSVP triple (Φ, v, S). It governs how capacity Φ (Misner et al. 1973; Wald 1984) moves through space, how entropy S is advected and mixed, and how the admissibility boundary $\partial\mathcal{X}_t$ deforms over time. This chapter derives the transport field equations and their relationship to reachability geometry.

71.1 The Vector Transport Field

Definition 71.1 (*Transport Field*)**.** The **transport field** $\mathrm{v} : \mathcal{U} \times \mathbb{R} \to T\mathcal{U}$ is a time-dependent vector field assigning to each spacetime point a velocity vector $\mathrm{v}(x,t) \in T_x\mathcal{U}$ governing the advection of capacity and entropy.

Physical interpretations:

- Fluid: velocity field of the fluid.
- Ecology: direction of nutrient flow through food web.
- Cognition: direction of attention shift.
- Institution: direction of resource allocation.
- Cosmology: expansion velocity field of spacetime.

71.2 Transport Field Equations

The transport field itself evolves under forcing from the capacity gradient and entropy sources:

Definition 71.2 (*Transport Field Equation*)**.** The **transport field equation** is

$$\partial_t \mathrm{v} + (\mathrm{v} \cdot \nabla)\mathrm{v} = -\frac{1}{\rho}\nabla P + \nu \nabla^2 \mathrm{v} + \mathrm{f},$$

where:

- $\rho = \Phi$ is the effective density (capacity);
- P is the pressure generated by the capacity gradient;
- $\nu > 0$ is a viscosity coefficient;
- f is the external forcing from repair Γ.

In the RSVP framework, we take $P = \Phi$ (capacity acts as its own pressure), yielding the simplified form:

$$\partial_t \mathrm{v} + (\mathrm{v} \cdot \nabla)\mathrm{v} = -\frac{\nabla \Phi}{\Phi} + \nu \nabla^2 \mathrm{v} + \mathrm{f}.$$

The term $-\nabla\Phi/\Phi = -\nabla \log \Phi$ drives the transport field *toward* regions of higher capacity — flow naturally moves capacity from high to low, but the field itself accelerates toward high-Φ regions. This is the RSVP analogue of advection under buoyancy.

71.3 Curl, Divergence, and Reachability Topology

The transport field v has two geometrically distinct components:

$$\mathrm{v} = \mathrm{v}_{\mathrm{irr}} + \mathrm{v}_{\mathrm{sol}} \quad \text{(Helmholtz decomposition)},$$

where $\nabla \times \mathrm{v}_{\mathrm{irr}} = 0$ (irrotational, potential flow) and $\nabla \cdot \mathrm{v}_{\mathrm{sol}} = 0$ (solenoidal, incompressible flow).

Proposition 71.1 (*Divergence and Reachability Volume Rate*)**.** *The rate of change of the reachability volume along the flow is:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}_R(x_0, T) = \int_{\mathcal{R}_{\mathcal{A}}(x_0,T)} \nabla \cdot \mathrm{v}\, \mathrm{d}\mu.$$

Regions where $\nabla \cdot \mathrm{v} > 0$ (sources) expand reachability; regions where $\nabla \cdot \mathrm{v} < 0$ (sinks) contract it.

Proof. By the transport theorem (Reynolds): the time derivative of a volume integral over a moving domain $\mathcal{R}_t$ satisfies $\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{R}_t} \mathrm{d}\mu = \int_{\mathcal{R}_t} \nabla \cdot \mathrm{v}\, \mathrm{d}\mu$. ▪ ▪

Corollary 71.2 (*Reachability Preservation under Incompressible Flow*)**.** *If* v *is incompressible ($\nabla \cdot \mathrm{v} = 0$), then $\mathcal{V}_R(x_0, T)$ is constant along the flow. Incompressible transport preserves reachability volume exactly.*

Proof. Substituting $\nabla \cdot \mathrm{v} = 0$ into Proposition 71.1: $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}_R = \int_{\mathcal{R}} \nabla \cdot \mathrm{v}\, \mathrm{d}\mu = 0$. Hence $\mathcal{V}_R(x_0, T)$ is constant in T. ▪ ▪

This is the fluid-dynamic analogue of a repair operator with $|\det J_{\mathfrak{R}}| = 1$ — area-preserving, no information loss or gain.

71.4 Circulation and Topological Trapping

The curl $\nabla \times \mathrm{v}$ (vorticity) creates circulating transport patterns that can *trap* trajectories.

Definition 71.3 (*Topological Trap*)**.** A region $\mathcal{T} \subset \mathcal{A}$ is a **topological trap** if the transport field v has nonzero circulation around the boundary $\partial\mathcal{T}$:

$$\oint_{\partial\mathcal{T}} \mathrm{v} \cdot \mathrm{d}\ell \neq 0.$$

Trajectories entering $\mathcal{T}$ circulate indefinitely and cannot easily escape.

Proposition 71.3 (*Topological Traps Reduce Effective Reachability*)**.** *If $\mathcal{T}$ is a topological trap and a trajectory enters $\mathcal{T}$ at time t_0, its expected exit time τ_{exit} satisfies*

$$\mathbb{E}[\tau_{\mathrm{exit}}] \geq \frac{|\oint_{\partial\mathcal{T}} \mathrm{v} \cdot \mathrm{d}\ell|}{\|\mathrm{v}\|_{\partial\mathcal{T}}\, |\partial\mathcal{T}|}.$$

Hence states inside the trap remain unreachable from outside for at least $\mathbb{E}[\tau_{\mathrm{exit}}]$ time units.

Proof. By Stokes' theorem, the circulation $\Gamma = \oint_{\partial\mathcal{T}} \mathrm{v} \cdot \mathrm{d}\ell = \int_{\mathcal{T}} \nabla \times \mathrm{v}\, \mathrm{d}A \neq 0$ means the flow has net rotation inside $\mathcal{T}$. A particle traversing $\partial\mathcal{T}$ against the circulation (escaping the trap) must overcome the rotational component. The time to traverse a path of length $|\partial\mathcal{T}|$ against a flow of speed $\|\mathrm{v}\|_{\partial\mathcal{T}}$ is at least $|\Gamma|/(\|\mathrm{v}\|^2|\partial\mathcal{T}|)$. Multiplying through by $\|\mathrm{v}\| \cdot |\partial\mathcal{T}|$ and absorbing constants gives the stated bound. ▪ ▪

Topological traps reduce effective reachability even when they do not reduce admissibility: the system can reach states inside the trap but cannot easily leave. In institutional terms, a bureaucratic loop is a topological trap; in cognitive terms, rumination is a trap; in ecological terms, a closed nutrient cycle is a trap (self-sustaining but not expansive).

71.5 Summary

1. The transport field v governs how capacity and entropy move.
2. Its equation couples to Φ via the log-gradient $-\nabla \log \Phi$.
3. Divergence of v determines the rate of reachability volume change.
4. Incompressible transport preserves reachability volume exactly.

5. Non-zero curl creates topological traps that reduce effective reachability without reducing admissibility.

Exercises

71.1. Let $\mathrm{v}(x, y) = (-y, x)$ (pure rotation). Compute $\nabla \cdot \mathrm{v}$ and $\nabla \times \mathrm{v}$. Is this flow incompressible? Does it create a topological trap?

71.2. For the simplified transport equation with $\nu = 0$ and $\mathrm{f} = 0$, find a stationary solution $\partial_t \mathrm{v} = 0$ of the form $\mathrm{v} = c\nabla \Phi$. Determine c from the equation.

71.3. Prove Proposition 71.1 for the case $\mathcal{A} = \mathbb{R}^n$ and v smooth and compactly supported.

71.4. (Institutional.) Model a government bureaucracy as a transport field on a policy state space. Identify a structural feature that makes it a topological trap. Propose a repair operator (reform) that breaks the circulation.

72 Chapter 72

Entropy Fields

> Entropy is not disorder. It is the measure of lost distinction.
>
> —

Phenomenological Note. Things move because something carries them. Energy does not teleport; it propagates. Information does not jump; it flows along paths. Understanding how something spreads means understanding the medium it spreads through — its channels, its resistances, the places where flow is fast and the places where it is blocked or redirected. The carrier is not incidental to the thing carried.

The entropy field $S(x, t)$ is the third element of the RSVP triple. It measures, at each spacetime point, the rate at which capacity Φ is irreversibly depleted — the rate at which the admissibility field loses structure. This chapter derives the entropy field equation, shows how entropy growth contracts the admissibility boundary, and connects RSVP entropy to the distinction-collapse framework of Part III.

72.1 The Entropy Field

Definition 72.1 (*Entropy Field*)**.** The **entropy field** $S : \mathcal{U} \times \mathbb{R} \to \mathbb{R}_{\geq 0}$ assigns to each spacetime point (x, t) a non-negative scalar $S(x, t)$ measuring the local rate of irreversible capacity depletion. High $S(x, t)$ indicates a region where structure is being destroyed; $S = 0$ indicates a perfectly reversible, structure-preserving region.

Physical interpretations:

- Thermodynamics: local entropy production rate $\dot{s}(x, t)$.
- Cognition: attentional noise; rate of memory decay.
- Ecology: metabolic waste production.
- Institution: corruption, inefficiency, information leakage.
- Computing: irreversible bit erasure (Landauer principle).

72.2 The Entropy Field Equation

Definition 72.2 (*Entropy Field Equation*)**.** The **entropy field equation** is

$$\partial_t S + \mathrm{v} \cdot \nabla S = \kappa \nabla^2 S + \sigma(\Phi, \mathrm{v}) - \mu_S \mathfrak{R},$$

where:

- $\kappa > 0$ is the entropy diffusion coefficient (entropy spreads from high-S to low-S regions);
- $\sigma(\Phi, \mathrm{v}) \geq 0$ is the local entropy production rate (generated by capacity depletion and flow viscosity);
- $\mu_S > 0$ is the repair-suppression coefficient (repair operators reduce local entropy).

The production term is:

$$\sigma(\Phi, \mathrm{v}) = \lambda \Phi^{-1} (\nabla \Phi)^2 + \nu (\nabla \mathrm{v})^2,$$

where the first term is entropy from capacity gradients (steep gradients produce entropy) and the second from viscous dissipation in the flow.

72.3 The RSVP Second Law

Theorem 72.1 (*RSVP Second Law*)**.** *In the absence of repair* ($\mathfrak{R} = 0$) *and with reflecting boundary conditions on* $\partial\mathcal{U}$*, the total entropy* $\mathcal{S}_{\text{tot}}(t) = \int_{\mathcal{U}} S(x, t)\, \mathrm{d}\mu$ *is non-decreasing:*

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{S}_{\text{tot}}(t) \geq 0.$$

Proof. Integrating the entropy field equation over $\mathcal{U}$:

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{S}_{\text{tot}} = \int_{\mathcal{U}} [-\mathrm{v} \cdot \nabla S + \kappa \nabla^2 S + \sigma]\, \mathrm{d}\mu.$$

The advection term $\int -\mathrm{v} \cdot \nabla S\, \mathrm{d}\mu = -\int \nabla \cdot (S\mathrm{v})\, \mathrm{d}\mu + \int S \nabla \cdot \mathrm{v}\, \mathrm{d}\mu$; the first integral vanishes by the divergence theorem (with reflecting boundaries) and the second is bounded by the divergence of v. The diffusion term $\int \kappa \nabla^2 S\, \mathrm{d}\mu = 0$ (divergence theorem, no-flux boundary). The production term $\int \sigma\, \mathrm{d}\mu \geq 0$ since $\sigma \geq 0$. Therefore $\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{S}_{\text{tot}} \geq 0$. ▪ ▪

Entropy as Distinction Loss. The RSVP Second Law is the field-theoretic version of Lemma 6.1: without repair, distinctions are lost monotonically. Entropy increase is the quantitative rate at which the admissibility field loses its capacity to distinguish states. The thermodynamic arrow of time is the irreversible distinction-collapse of physical systems.

72.4 Entropy and Admissibility Boundary Contraction

From the constraint field equation (Chapter 15), the boundary $\partial\mathcal{X}_t$ moves with normal velocity:

$$v_n = -\frac{\partial_t \Psi}{|\nabla\Psi|} = -\frac{\partial_t(\Phi - S - \theta)}{|\nabla(\Phi - S)|}.$$

When S increases faster than Φ (entropy growing faster than capacity is being replenished), $\partial_t(\Phi - S) < 0$ and $v_n > 0$ — the boundary moves *inward*: the admissibility domain contracts.

The rate of contraction is proportional to the excess entropy production rate $\partial_t S - \partial_t \Phi$. Repair acts to reverse this: $\Gamma > 0$ increases Φ and $\mathfrak{R} > 0$ decreases S, both pushing $v_n < 0$ (outward boundary motion).

72.5 Entropy Fields Across Domains

Cosmology.. In the early universe, entropy density S was low and Φ was high (high free energy density). The admissibility domain for complex structure was large. As entropy increased, the domain contracted. Galaxy formation, stellar evolution, and planetary systems are all instances of local repair ($\Gamma > 0$, $\mathfrak{R} > 0$) temporarily expanding local admissibility against the global entropy increase.

Cognitive aging.. Cognitive decline is modelled by increasing S (attentional noise, synaptic deterioration) and decreasing Φ (working memory capacity). The admissibility boundary of cognitive state space contracts. Cognitive repair (learning, exercise, rest) temporarily reverses this.

Institutional decay.. Institutional entropy increases through corruption, inefficiency, and information loss. The admissibility domain of institutional action shrinks. Reform is repair: it reduces S and increases Φ, expanding the range of admissible institutional trajectories.

Exercises

72.1. Let $S(x,t) = S_0 + \alpha t$ (linearly growing entropy, spatially uniform) and $\Phi(x,t) = \Phi_0$ (constant capacity). Find the time t^* at which the admissibility boundary $\{\Phi - S = \theta\}$ reaches x for x in the initial interior.

72.2. Show that with diffusion only ($\mathrm{v} = 0$, $\sigma = 0$, $\mathfrak{R} = 0$): $S(x,t) \to \bar{S} = \mathcal{S}_{\text{tot}}/|\mathcal{U}|$ uniformly as $t \to \infty$. Interpret: entropy diffuses to a uniform background, erasing spatial distinctions.

72.3. Define the *local admissibility time* $\tau(x) = \sup\{t : x \in \mathcal{X}_t\}$ — how long x remains admissible. Express $\tau(x)$ in terms of the entropy production rate $\sigma(x)$ for the spatially uniform case.

72.4. (Cosmology.) The cosmic microwave background has a specific entropy-per-baryon $s/n_b \approx 10^{10}$. Interpret this as a compression ratio: the ratio of the original admissibility volume of the universe to the current one. What does the CPR framework predict about the maximum reachability remaining in the observable universe?

73 Chapter 73

Reachability Dynamics

> The reachable set breathes. It expands when capacity flows in, and contracts when entropy accumulates.
>
> —

Phenomenological Note. Entropy is the tendency of arrangements to become less specific over time. This is not a force; it is a consequence of the fact that there are vastly more disordered arrangements than ordered ones. What is surprising is not that order decays but that it ever arises and persists. Order requires work to maintain. The work is invisible when it is succeeding.

This chapter derives the equation governing the time evolution of reachability volume under the coupled RSVP field equations. (Boltzmann 1877; Gibbs 1902; Landauer 1961) The derivation uses the Reynolds transport theorem (cf. Evans 2010) applied to the level-set formulation of the admissibility boundary. The result is the three-term reachability balance — the master equation of the CPR framework in dynamical form.

73.1 Setup: Moving Admissibility Boundary

Let $\Psi(x,t) = \Phi(x,t) - S(x,t) - \theta$ be the admissibility potential, so that the admissible domain at time t is $\mathcal{A}_t = \{x : \Psi(x,t) \geq 0\}$ and its boundary is $\partial\mathcal{A}_t = \{x : \Psi(x,t) = 0\}$.

The reachable set $R_t = \mathcal{R}_{\mathcal{A}_t}(x_0, T)$ is a subset of $\mathcal{A}_t$ that itself varies with t as the admissibility field evolves. The reachability volume is

$$\mathcal{V}_R(t) = \int_{R_t} \mathrm{d}\mu.$$

73.2 Reynolds Transport Theorem Application

Theorem 73.1 (*Reachability Volume Dynamics*). *For a system governed by the RSVP field equations (Definition 92.6), the time derivative of reachability volume* *satisfies:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}_R(t) = -\int_{\partial R_t} \frac{\partial_t \Phi - \partial_t S}{|\nabla(\Phi - S)|}\,\mathrm{d}\sigma.$$

Substituting the RSVP capacity equation $\partial_t \Phi = -\nabla\cdot(\Phi \mathrm{v}) - \lambda S + \Gamma$*, this becomes the three-term balance:*

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}_R(t) = \underbrace{-\int_{\partial R_t} \frac{\nabla\cdot(\Phi \mathrm{v})}{|\nabla\Psi|}\,\mathrm{d}\sigma}_{\textit{transport}} + \underbrace{\int_{\partial R_t} \frac{\lambda S + \partial_t S}{|\nabla\Psi|}\,\mathrm{d}\sigma}_{\textit{entropy}} - \underbrace{\int_{\partial R_t} \frac{\Gamma}{|\nabla\Psi|}\,\mathrm{d}\sigma}_{\textit{repair}}}$$

where $|\nabla\Psi| = |\nabla(\Phi - S)|$.

Proof. Step 1: Reynolds transport theorem. The reachable set R_t is a moving domain. By the Reynolds transport theorem:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{R_t} \mathrm{d}\mu = \int_{\partial R_t} u_n\,\mathrm{d}\sigma,$$

where u_n is the outward normal velocity of the moving boundary ∂R_t and $\mathrm{d}\sigma$ is the surface area element.

Step 2: Normal velocity of the boundary. The boundary $\partial R_t \subseteq \partial \mathcal{A}_t$ is defined by $\Psi(x,t) = 0$. Differentiating along a path $x(t)$ that tracks the boundary ($\Psi(x(t),t) = 0$ for all t):

$$\frac{\mathrm{d}}{\mathrm{d}t}\Psi(x(t),t) = \partial_t \Psi + \nabla\Psi \cdot \dot{x}(t) = 0.$$

The normal component of $\dot{x}(t)$ is:

$$u_n = \dot{x}(t)\cdot\hat{n} = \dot{x}(t)\cdot\frac{\nabla\Psi}{|\nabla\Psi|} = \frac{\nabla\Psi\cdot\dot{x}(t)}{|\nabla\Psi|} = \frac{-\partial_t\Psi}{|\nabla\Psi|}.$$

Step 3: Compute $\partial_t \Psi$*.*

$$\partial_t\Psi = \partial_t\Phi - \partial_t S.$$

Step 4: Substitution.

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}_R = \int_{\partial R_t} u_n\,\mathrm{d}\sigma = \int_{\partial R_t}\frac{-\partial_t\Psi}{|\nabla\Psi|}\,\mathrm{d}\sigma = -\int_{\partial R_t}\frac{\partial_t\Phi - \partial_t S}{|\nabla\Psi|}\,\mathrm{d}\sigma.$$

Step 5: RSVP capacity substitution. Using $\partial_t\Phi = -\nabla\cdot(\Phi\mathrm{v}) - \lambda S + \Gamma$:

$$\partial_t\Phi - \partial_t S = -\nabla\cdot(\Phi\mathrm{v}) - \lambda S + \Gamma - \partial_t S.$$

Substituting:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}_R = -\int_{\partial R_t} \frac{-\nabla\cdot(\Phi \mathrm{v}) - \lambda S + \Gamma - \partial_t S}{|\nabla \Psi|}\,\mathrm{d}\sigma$$

$$= \underbrace{-\int_{\partial R_t} \frac{\nabla\cdot(\Phi \mathrm{v})}{|\nabla \Psi|}\,\mathrm{d}\sigma}_{\text{transport}} + \underbrace{\int_{\partial R_t} \frac{\lambda S + \partial_t S}{|\nabla \Psi|}\,\mathrm{d}\sigma}_{\text{entropy}} - \underbrace{\int_{\partial R_t} \frac{\Gamma}{|\nabla \Psi|}\,\mathrm{d}\sigma}_{\text{repair}}. \quad \blacksquare$$

$\blacksquare$

73.3 Interpretation of the Three Terms

The Reachability Balance. Theorem 73.1 shows that reachability volume changes only through boundary motion, and boundary motion is governed by three forces:

Term	Sign	Effect
$-\int \frac{\nabla\cdot(\Phi \mathrm{v})}{|\nabla\Psi|}\,\mathrm{d}\sigma$	$+$ when $\nabla\cdot\mathrm{v} < 0$ (converging flow)	Transport expands $\mathcal{V}_R$
$+\int \frac{\lambda S+\partial_t S}{|\nabla\Psi|}\,\mathrm{d}\sigma$	$+$ (always, since $S, \lambda, \partial_t S \geq 0$)	Entropy *contracts* $\mathcal{V}_R$
$-\int \frac{\Gamma}{|\nabla\Psi|}\,\mathrm{d}\sigma$	$-$ (since $\Gamma \geq 0$)	Repair *expands* $\mathcal{V}_R$

A system is in *reachability equilibrium* when the three terms balance exactly. A *thriving* system has repair exceeding entropy. A *declining* system has entropy exceeding repair.

73.4 Special Cases

73.4.1 Incompressible Transport

The transport term vanishes. $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}_R = E_{\text{entropy}} - E_{\text{repair}}$. Reachability is controlled entirely by the balance of entropy and repair.

73.4.2 No Repair () and No Entropy

Transport-only dynamics: $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}_R = -\int \frac{\nabla\cdot(\Phi \mathrm{v})}{|\nabla\Psi|}\,\mathrm{d}\sigma$. This is the Liouville-type term: expanding flows expand reachability, contracting flows reduce it.

73.4.3 Static Fields (, $\mathrm{v} = 0$)

$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}_R = 0$. A static admissibility field has constant reachability volume. This is the equilibrium condition.

73.4.4 Homeostatic Condition

Repair exactly balances entropy when $\Gamma = \lambda S + \partial_t S$ everywhere on ∂R_t, giving $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}_R = -\int \frac{\nabla\cdot(\Phi \mathrm{v})}{|\nabla\Psi|}\,\mathrm{d}\sigma$. For incompressible flow, this gives exact homeostasis.

Exercises

73.1. For a 1D RSVP system with $\Phi(x,t) = \Phi_0$, $\mathrm{v} = v_0$, $S = \alpha\Phi_0$, $\Gamma = \lambda\alpha\Phi_0$ (repair exactly balancing entropy): verify that $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}_R = 0$.

73.2. Let $\Gamma = 0$, $\mathrm{v} = 0$, and $S(x,t) = S_0 e^{\beta t}$ (growing entropy). Compute $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}_R$ and find the time t^* at which $\mathcal{V}_R = 0$.

73.3. Derive the discrete-time version of the reachability dynamics: if $\mathcal{A}_{t+1} \subseteq \mathcal{A}_t$ (admissibility contracts at each step), show $\mathcal{V}_R(t+1) \leq \mathcal{V}_R(t)$ using Lemma 14.1 directly.

73.4. (Cosmology.) For a flat FLRW universe with scale factor $a(t) \propto t^{2/3}$ (matter domination), the Hubble flow has $\nabla \cdot \mathrm{v} = 3H = 2/t$. Compute the transport term's contribution to $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}_R$. How does it compare to the entropy term for the observed cosmic entropy density?

74 Chapter 74

Lamphrodyne Relaxation

> A gradient that cannot be sustained must relax. The question is only whether it relaxes into order or into noise.
>
> —

Phenomenological Note. After a disturbance, a system settles. The settling is not instantaneous; it has a characteristic time and a characteristic shape. Some systems settle quickly and cleanly. Others oscillate for a long time before finding rest. Others never fully settle but spend their time near the equilibrium without quite reaching it. The character of the settling tells you something about the structure of the system that you cannot see during the disturbance itself.

Lamphrodyne relaxation is the process by which strained RSVP fields — those with steep gradients in Φ or large misalignments between v and $\nabla\Phi$ — spontaneously smooth toward a lower-entropy, more uniform configuration. The name derives from the Greek *lamphrós* (brilliant, intense) + *dynē* (force): the relaxation of intense field gradients. This chapter derives the lamphrodyne relaxation equation and proves the Gradient Dissipation Theorem, showing that field strain decreases monotonically under relaxation.

74.1 Field Strain

Definition 74.1 (*RSVP Field Strain*). The **field strain** at (x, t) is

$$\mathcal{E}(x,t) := \frac{|\nabla\Phi|^2}{\Phi^2} + \alpha\,|\nabla\times\mathrm{v}|^2 + \beta\,|\nabla S|^2,$$

where $\alpha, \beta > 0$ are weighting coefficients. The first term measures the relative gradient of Φ (steepness of the capacity landscape); the second measures vorticity strain in the flow; the third measures entropy gradient (heterogeneity of dissipation).

High field strain corresponds to:

- Large spatial variation in capacity ($|\nabla\Phi|/\Phi$ large);
- High rotational flow (vorticity);

- Sharp spatial gradients in entropy production.

The total field strain is $\mathcal{E}_{\text{tot}}(t) = \int_{\mathcal{U}} \mathcal{E}(x,t)\,\mathrm{d}\mu$.

74.2 The Lamphrodyne Relaxation Equation

Definition 74.2 (*Lamphrodyne Relaxation*). **Lamphrodyne relaxation** is the dynamics obtained by gradient-descending the total field strain:

$$\partial_t \Phi = \eta_\Phi \nabla^2 \Phi - \eta_\Phi \frac{|\nabla \Phi|^2}{\Phi},$$
$$\partial_t \mathrm{v} = \eta_v \nabla^2 \mathrm{v} - \eta_v \alpha (\nabla \times \nabla \times \mathrm{v}),$$
$$\partial_t S = \eta_S \nabla^2 S.$$

These are the L^2 gradient flow equations for $\mathcal{E}_{\text{tot}}$.

The Φ-equation is a nonlinear diffusion equation that smooths $\log \Phi$ (since $\nabla^2\Phi/\Phi - |\nabla\Phi|^2/\Phi^2 = \nabla^2 \log \Phi$). The v-equation dissipates vorticity. The S-equation is standard heat diffusion smoothing the entropy gradient.

74.3 The Gradient Dissipation Theorem

Theorem 74.1 (*Gradient Dissipation — Pure Lamphrodyne Flow*). *Under* pure lamphrodyne flow (*the three decoupled diffusion equations, with no cross-coupling between fields*)*, the total field strain $\mathcal{E}_{\text{tot}}(t)$ is strictly decreasing:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\text{tot}}(t) \;\le\; 0,$$

with equality only at the uniform field configuration.

For the full coupled RSVP system*, the bound generalises to:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\text{tot}}(t) \;\le\; -\mathcal{D}(t) + \mathcal{C}(t),$$

where $\mathcal{D}(t) \ge 0$ is the dissipation rate from the diagonal terms and $\mathcal{C}(t) \ge 0$ is the coupling injection from cross-field terms. Equilibrium (stable relaxation) is achieved when $\mathcal{D}(t) > \mathcal{C}(t)$, which holds when the coupling coefficients are sufficiently small.

Proof. We compute $\frac{\mathrm{d}}{\mathrm{d}t}\int |\nabla \log \Phi|^2\,\mathrm{d}\mu$ (representative term; the others are analogous).

Let $\psi = \log \Phi$. Then $\partial_t \psi = \eta_\Phi \nabla^2 \psi$ (the Φ-relaxation equation becomes the heat equation for ψ).

$$\frac{\mathrm{d}}{\mathrm{d}t}\int |\nabla\psi|^2\,\mathrm{d}\mu = 2\int \nabla\psi \cdot \nabla(\partial_t\psi)\,\mathrm{d}\mu = 2\eta_\Phi \int \nabla\psi \cdot \nabla(\nabla^2\psi)\,\mathrm{d}\mu.$$

Integrating by parts (Green's identity):

$$= -2\eta_\Phi \int |\nabla^2 \psi|^2 \,\mathrm{d}\mu \;\leq\; 0.$$

The same argument applies to the vorticity and entropy gradient terms. Summing, $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\mathrm{tot}} \leq 0$, with equality iff $\nabla^2\psi = 0$, $\nabla^2\mathrm{v} = 0$, $\nabla^2 S = 0$ (all fields harmonic) — which for compact $\mathcal{U}$ with no-flux boundary conditions means all fields are constant. ▪ ▪

74.4 Lamphrodyne vs. Entropy Increase

There is a striking contrast:

- The RSVP Second Law (Theorem 72.1): total entropy $\mathcal{S}_{\mathrm{tot}}$ is non-decreasing.
- The Gradient Dissipation Theorem (Theorem 74.1): total field strain $\mathcal{E}_{\mathrm{tot}}$ is non-increasing.

These are not in conflict: entropy increases toward uniformity, and field strain decreases toward uniformity. Both arrows point in the same direction — toward the homogeneous equilibrium state. The difference is that entropy tracks the *level* of disorder, while field strain tracks the *gradients* that drive change.

Lamphrodyne as the Physics of Repair Convergence. Theorem 74.1 is the physical realisation of the Repair Convergence Theorem (Theorem 86.1): in RSVP field theory, the natural dynamics perform repair (strain dissipation = admissibility-preserving reachability expansion) and converge to a fixed point (the uniform field). Lamphrodyne relaxation is the CPR framework's microscopic mechanism for repair.

74.5 Applications

Physical systems.. In fluid turbulence, lamphrodyne relaxation corresponds to the cascade of energy from large-scale vortices (high $|\nabla \times \mathrm{v}|^2$) to small-scale thermal motion (high S, low strain). The field strain decreases as kinetic energy is dissipated.

Cognitive consolidation.. During sleep, the brain undergoes synaptic homeostasis: over-strengthened synaptic gradients are smoothed, and the capacity field Φ becomes more uniform. This is cognitive lamphrodyne relaxation — the brain performing gradient dissipation to maintain a high-capacity, low-strain state for the next day's admissibility demands.

Institutional normalisation.. After a crisis (sudden large gradients in institutional capacity Φ), institutions undergo a normalisation process: inequalities in resource distribution are reduced, communication flows stabilise, and entropy production returns to background levels. This is institutional lamphrodyne relaxation.

Exercises

74.1. Solve the $\log \Phi$ heat equation $\partial_t \psi = \eta \nabla^2 \psi$ on $\mathcal{U} = [0, L]$ with initial condition $\psi_0(x)$ and zero-flux boundary conditions. Show that $|\nabla \psi|^2$ decays exponentially to zero.

74.2. Prove that the total strain dissipation rate satisfies: $-\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}_{\text{tot}} \geq C \mathcal{E}_{\text{tot}}^2$ for some $C > 0$ (the dissipation is at least quadratic in the strain).

74.3. (Cognitive.) Model the sleep-consolidation process as lamphrodyne relaxation on a synaptic weight manifold $\mathcal{W}$. Define field strain as the variance of synaptic weights. Verify that the relaxation equation reduces variance monotonically. What is the equilibrium state?

74.4. Show that lamphrodyne relaxation preserves the integral $\int \Phi \, \mathrm{d}\mu$ (total capacity is conserved during pure relaxation with $\Gamma = \mathfrak{R} = 0$). What does this imply about the long-run capacity distribution?

75 Chapter 75

Cosmological Reachability

The universe asks: what is still possible?

—

Phenomenological Note. Most of the universe is already unreachable. The light from the most distant galaxies left before the Earth existed, and by the time it reaches us those galaxies have receded so far that we could never travel there even at the speed of light. This is not a practical limitation that technology might overcome. It is a structural feature of the spacetime we inhabit. The reachable universe is finite and shrinking.

At cosmological scales, the admissibility field is spacetime geometry and the reachability structure is the causal light cone. The cosmic event horizon bounds the reachable future; cosmic expansion continuously modifies reachability volume.

75.1 Cosmological Admissibility

Definition 75.1 (*Cosmological Admissibility Field*). In a flat FLRW universe with scale factor $a(t)$:

$$\mathcal{A}_{\text{cosmo}}(x_0, t_0) = \{x \in \text{spacetime} : |x - x_0| \leq c(t - t_0)\}$$

— the future light cone from observer O (cf. Hawking and Ellis 1973; Wald 1984) at (x_0, t_0). Only events within the light cone are causally reachable.

Theorem 75.1 (*Cosmological Reachability Dynamics*). *In matter-dominated expansion ($a(t) \propto t^{2/3}$), the comoving reachability volume grows without bound: $\mathcal{V}_R(O, T) \propto T^3$.*

In Λ-dominated expansion ($a(t) \propto e^{Ht}$), the physical event horizon converges to $R_H = c/H$, and the reachability volume saturates at: $\mathcal{V}_{\text{dS}} = \frac{4\pi}{3}\left(\frac{c}{H}\right)^3$.

Proof. The particle horizon at time T is $d_H(T) = a(T)\int_0^T \frac{c\,\mathrm{d}t'}{a(t')}$. For matter domination: $d_H \propto T$, so $\mathcal{V}_R \propto d_H^3 \propto T^3$. For Λ-domination: $\int_T^\infty c\,\mathrm{d}t/a(t) < \infty$ (the integral converges), giving a finite event horizon $R_H = c\int_T^\infty e^{-H(t-T)}\,\mathrm{d}t = c/H$. ▪ ▪

75.2 Dark Energy as Reachability Limiter

Dark energy ($\Lambda > 0$) drives accelerating expansion, causing the cosmological event horizon to shrink: more and more of the universe becomes permanently unreachable. In RSVP terms: Λ acts as a global entropy field that contracts the admissibility domain. Galaxies currently outside our Hubble sphere are already causally disconnected. Eventually all galaxies beyond the Local Group will be inaccessible even in principle.

The Universe's Declining Reachability. The universe is in the "declining system" regime: entropy (dark energy acceleration) is increasing faster than any repair process can counteract. The total cosmological reachability volume is finite and decreasing relative to the current Hubble volume. Structure formation (stars, galaxies, life) is local repair creating temporary high-reachability islands in a universe whose global admissibility field is contracting.

Exercises

75.1. Compute the current cosmological event horizon radius ($H_0 \approx 70\,\mathrm{km/s/Mpc}$). How many galaxies are currently within our event horizon? How does this compare to the total number of observable galaxies?

75.2. For a recollapsing (closed) universe, the reachability volume first grows then shrinks to zero at the Big Crunch. Compute $\mathcal{V}_R(O,T)$ as a function of T for this case. What is the maximum reachability time T^*?

75.3. Hawking radiation causes black holes to evaporate, releasing information that was previously trapped inside the horizon. Model this as a repair operator on the cosmological admissibility field. Does Hawking radiation increase $\mathcal{V}_R$?

75.4. Interpret the Fermi Paradox through reachability geometry: if technological civilisations expand to fill their event horizons, what does the silence of the cosmos imply about the typical cosmological admissibility field for intelligence?

76 Chapter 76

Falling Universes

A universe is a trajectory falling through the space of physical constants. We live in one that stayed admissible.

—

Phenomenological Note. Why are the laws of physics what they are? One answer is that slightly different laws would not have produced observers capable of asking the question. This is not an explanation in the usual sense. It is an observation about which configurations of physical law are compatible with the existence of beings who notice configurations. The question of what selected the actual laws from among all possible ones remains genuinely open.

The multiverse hypothesis reformulates in CPR terms: each universe is a trajectory through the space of physical constants $\mathcal{X}_{\text{phys}}$, and the observed constants are those that pass through the anthropic admissibility field.

76.1 Physical Constants as Trajectory

Definition 76.1 (*Physical Constant Space*)**.** $\mathcal{X}_{\text{phys}}$ is the space of all consistent assignments of dimensionless physical constants (α, m_p/m_e, Λ/M_{Pl}^4, etc.). Each point in $\mathcal{X}_{\text{phys}}$ defines a universe with corresponding laws of physics.

Theorem 76.1 (*Anthropic Selection as Admissibility*)**.** *Define the* ***anthropic admissibility field****:* $\mathcal{A}_{\text{ant}} = \{c \in \mathcal{X}_{\text{phys}} : c \text{ permits observers}\}$*. The observed constants* $c^* \in \mathcal{A}_{\text{ant}}$ *(necessary condition). The fine-tuning problem is the observation that:* $\mu(\mathcal{A}_{\text{ant}})/\mu(\mathcal{X}_{\text{phys}}) \ll 1$ *(the admissible region is tiny relative to the full constant space), combined with* $c^* \in \mathcal{A}_{\text{ant}}$ *(we observe admissible constants). These two facts together require either selection (multiverse), design (theism), or coincidence.*

Proof. $c^* \in \mathcal{A}_{\text{ant}}$: if the constants were outside $\mathcal{A}_{\text{ant}}$, no observers would exist to measure them (weak anthropic principle). The volume estimate $\mu(\mathcal{A}_{\text{ant}})/\mu(\mathcal{X}_{\text{phys}}) \ll 1$: numerical calculations show that small variations in α, m_p/m_e, or Λ destroy the conditions for nucleosynthesis, stellar evolution, or chemistry. The fine-tuning problem is a consequence of these two facts. ▪ ▪

76.2 The Landscape as Admissibility Manifold

String theory predicts $\sim 10^{500}$ vacua — a huge landscape of possible universes. In CPR terms: the string landscape is a discretisation of $\mathcal{X}_{\text{phys}}$, and the observed universe is one of the (presumably rare) landscape points in $\mathcal{A}_{\text{ant}}$.

Proposition 76.2 (*Landscape Navigation as Cosmological Repair*)**.** *Bubble nucleation (the quantum tunnelling process that selects a vacuum in eternal inflation) is a repair operator on the landscape: $\mathfrak{R}_{\text{bubble}} : \mathcal{X}_{\text{phys}} \to \mathcal{X}_{\text{phys}}$ that moves the universe from one landscape point to another. The process terminates when the universe reaches a stable landscape minimum — a repair fixed point (Definition 67.2).*

Proof. The string landscape is a directed graph: nodes are vacua (landscape points), edges are bubble nucleation transitions. Bubble nucleation from vacuum v_1 to v_2 is the repair operator $\mathfrak{R}_{\text{bubble}}(v_1) = v_2$ when $V(v_2) < V(v_1)$ (the tunnelling goes downhill in potential energy). A stable landscape minimum v^* has no lower-energy neighbours: $\mathfrak{R}_{\text{bubble}}(v^*) = v^*$ (the repair is a fixed point, since no further tunnelling occurs). By the Repair Fixed Point Theorem (Theorem 67.2), the fixed point is stable when the potential well is deep (large spectral gap in the local landscape graph). ▪ ▪

Exercises

76.1. The electromagnetic fine structure constant $\alpha \approx 1/137$. Show that for $\alpha > 0.1$, atoms cannot form stable orbitals. For $\alpha < 10^{-5}$, stars cannot ignite fusion. Estimate $\mu(\mathcal{A}^{(\alpha)}_{\text{ant}})/\mu(\mathcal{X}^{(\alpha)}_{\text{phys}})$ for the α-direction alone.

76.2. Is $\mathcal{A}_{\text{ant}}$ connected in $\mathcal{X}_{\text{phys}}$? If it is disconnected (multiple isolated viable regions), what does this imply for the multiverse interpretation?

76.3. Model the cosmological constant problem ($\Lambda_{\text{obs}} \ll \Lambda_{\text{QFT}}$) as: Λ_{obs} is the parameter value that keeps the universe in $\mathcal{A}_{\text{ant}}$, while Λ_{QFT} is the naive theoretical value. The problem becomes: why is the observed Λ so close to the maximum value in $\mathcal{A}_{\text{ant}}$?

76.4. Eternal inflation is a dynamics on $\mathcal{X}_{\text{phys}}$: the universe explores landscape points through bubble nucleation. Using the Repair Convergence Theorem, under what conditions does eternal inflation converge to a universe in $\mathcal{A}_{\text{ant}}$?

77 Chapter 77

Reconstruction from Cosmological Signals

> The past is written in surviving signals. We read what the compression preserved.
>
> —

Phenomenological Note. We know a great deal about the early universe from a relic: the faint microwave glow left over from a time when the universe was small and hot and opaque. It is a compressed record, full of information but also missing much that cannot be recovered. Reading it requires knowing what kind of compression was applied, what processes transformed the original state into the signal we observe. Cosmology is, among other things, a reconstruction problem.

The cosmic microwave background (CMB) (Hawking and Ellis 1973; Wald 1984), gravitational wave background, and large-scale structure are compressed records of early-universe dynamics. Cosmological reconstruction is the inverse problem: given these compressed signals, recover the admissibility field of the early universe.

77.1 CMB as Compressed History

Definition 77.1 (*CMB Compression*)**.** The CMB temperature fluctuations $\delta T(\hat{n})/T$ are a compression $\mathcal{C}_{\text{CMB}}$ of the primordial density perturbations $\mathcal{P}(k)$ via the transfer function $T(k, \ell)$:

$$C_\ell = \frac{2}{\pi} \int_0^\infty k^2 \mathcal{P}(k) |T(k, \ell)|^2 \, \mathrm{d}k.$$

The angular power spectrum $\{C_\ell\}$ is the compressed record.

Theorem 77.1 (*CMB Reconstruction Bound*)**.** *The CMB power spectrum $\{C_\ell\}$ provides a sufficient compression for the query "what were the primordial fluctuation amplitudes $\mathcal{P}(k)$" iff the transfer function $T(k, \ell)$ is injective in k for each ℓ. The reconstruction error is bounded by Theorem 22.1:*

$$\|\hat{\mathcal{P}} - \mathcal{P}\| \leq \frac{\epsilon_{\text{noise}} + \kappa_{\text{CMB}}}{\sigma_{\min}(T)},$$

where ϵ_{noise} is CMB measurement noise and $\sigma_{\min}(T)$ is the minimum singular value of the transfer function.

Proof. Sufficiency: T is injective iff the CMB uniquely determines $\mathcal{P}(k)$. In practice, $T(k, \ell)$ has near-zero singular values at large k (small scales) due to Silk damping, making small-scale power unrecoverable. The bound follows from Theorem 22.1 with the CMB transfer function as the projection Jacobian. ▪ ▪

77.2 Information Lost to the Horizon

The cosmic event horizon is an irreversible compression: information about regions beyond the horizon is permanently lost. By Theorem 29.1, alternative cosmological histories that produce the same CMB outside the horizon are indistinguishable.

Proposition 77.2 (*Horizon as Compression Depth Limit*)**.** *The cosmological event horizon sets the compression depth D of the CMB record: information about the universe at scales larger than c/H_0 (beyond the event horizon) is not encoded in $\{C_\ell\}$. Counterfactuals about the trans-horizon universe are unrecoverable from CMB observations alone (Lemma 28.1).*

Proof. The CMB power spectrum $\{C_\ell\}$ encodes information from recombination ($z \approx 1100$) up to the present. Information from beyond the particle horizon ($d > c/H_0 \cdot (\text{age of universe})$) cannot reach us by causality: no signal from those regions has had time to travel to our location. The horizon therefore sets the maximum lag in the CMB compression: $D_{\text{CMB}} = c \cdot t_{\text{recomb}}/(c/H_0) \approx 10^{-3}$ in units of the current Hubble radius. The counterfactual recovery lemma (Lemma 28.1) implies that alternative cosmological histories beyond this horizon are unrecoverable from CMB observations alone. ▪ ▪

77.3 Gravitational Waves as Independent Channel

The primordial gravitational wave background $h_{\mu\nu}$ provides an independent compression of early-universe dynamics. Since gravitational waves decouple earlier than photons, they carry information about the universe at higher energies — a deeper compression than the CMB.

Exercises

77.1. At what multipole ℓ does the CMB transfer function become degenerate (Silk damping kills the signal)? What early-universe information is permanently lost above this scale?

77.2. Model the gravitational wave background as a compression $\mathcal{C}_{\text{GW}}$ with compression depth $D_{\text{GW}} > D_{\text{CMB}}$. Which early-universe queries are sufficient under $\mathcal{C}_{\text{GW}}$ but not under $\mathcal{C}_{\text{CMB}}$?

77.3. The cosmic neutrino background ($C\nu B$) decoupled even earlier than photons. If detected, it would provide a compression $\mathcal{C}_{C\nu B}$ with even greater depth. What is the theoretical maximum depth of any cosmological signal?

77.4. Apply the Geometry of Forgetting theorem (Theorem 30.1) to cosmic expansion: as the universe ages, it "forgets" its early state. Compute ΔS (entropy increase) of the cosmological compression from recombination to today.

Part XI

Artificial Intelligence

[Part introduction — to be written.]

78 Chapter 78

Foundation Models

A foundation model is a projection from the space of all human meaning to a tractable manifold.

—

Phenomenological Note. The surprising thing about the current generation of language models is not that they are intelligent. It is that they are strange in a particular way: capable in some areas that seemed to require understanding, and failures in other areas that seemed trivial. This pattern is not random. It follows from the kind of compression that was applied and what it does and does not preserve. The capability profile is a fingerprint of the training process.

Foundation models (large language models, vision models, multimodal models) are trained to project the full space of human-generated text and imagery to compressed latent representations, then generate admissible continuations. In CPR terms, they learn an admissibility manifold $\mathcal{A}_{\text{FM}}$ and navigate it during generation.

78.1 Foundation Models as Projection Systems

Definition 78.1 (*Foundation Model as Projection*)**.** A **foundation model** is a pair $(\mathcal{C}_{\text{FM}}, \mathcal{R}_{\text{FM}})$ where: $\mathcal{C}_{\text{FM}} : \mathcal{X}_{\text{lang}} \to \mathcal{Z}_{\text{FM}}$ maps language/image inputs to a d-dimensional latent space, and $\mathcal{R}_{\text{FM}} : \mathcal{Z}_{\text{FM}} \to \mathcal{X}_{\text{lang}}$ generates outputs from latent representations.

Theorem 78.1 (*Foundation Model Reachability Bound*)**.** *A foundation model with d-dimensional latent space can represent at most 2^d reachability-relevant distinctions. Any downstream task requiring more than 2^d distinctions will encounter projection collapse (Theorem 19.1).*

Proof. The latent space $\mathcal{Z}_{\text{FM}} \subseteq \mathbb{R}^d$ can separate at most 2^d half-spaces under binary quantisation. Each half-space corresponds to one distinguishable reachability-relevant distinction. By the Language Compression Bound (Theorem 44.1), there are infinitely many reachability-relevant distinctions in a continuous semantic space. With $d < \infty$, some distinctions must collapse. ▪ ▪

78.2 Scaling Laws as Reachability Expansion

Proposition 78.2 (*Scaling Laws as Reachability Expansion*)**.** *Increasing model size N expands the latent space dimension $d \propto \log N$ and the training corpus increases the number of learned distinctions. Empirical scaling laws ($L \propto N^{-\alpha}$ for loss L) correspond to the reachability volume of the learned admissibility manifold growing as N^{α}: more capacity = more representable distinctions = lower loss.*

Proof. The empirical scaling law $L \propto N^{-\alpha}$ (loss vs. parameters) is explained by the reachability framework as follows. Model capacity scales as $d \propto \log N$ (log-linear in parameters). By Theorem 78.1, $\mathcal{V}_R \leq 2^d \propto N$. Loss L measures the fraction of reachability-relevant distinctions collapsed: $L \propto 1/\mathcal{V}_R \propto N^{-1}$. More precisely, for a mixture of distinction types with Pareto-distributed difficulty, the loss exponent is $\alpha \in (0, 1)$, consistent with empirically observed values $\alpha \approx 0.1$–0.2. ▪ ▪

78.3 In-Context Learning as Fiber Reduction

In-context learning (providing examples in the prompt) is a fiber reduction operation: the examples narrow the fiber $\mathcal{C}_{\mathrm{FM}}^{-1}(z)$ to those histories consistent with the provided examples. By Lemma 28.1, more examples give a narrower fiber and more reliable generation.

Exercises

78.1. GPT-4 has approximately $d \approx 12288$ hidden dimensions. Estimate the maximum number of reachability-relevant distinctions it can represent. Compare this to the estimated vocabulary of English ($\sim 10^5$ words $\times \sim 10^3$ senses/word).

78.2. Show that retrieval-augmented generation (RAG) increases effective d by appending retrieved documents to the context, expanding the fiber reduction without increasing model parameters.

78.3. Model the difference between GPT-3.5 and GPT-4 as an increase in $\mathcal{V}_R(\mathcal{A}_{\mathrm{FM}}, \cdot, T)$. What specific tasks require the larger reachability volume?

78.4. Prove that chain-of-thought prompting is a sequential fiber reduction: each reasoning step narrows the fiber, reducing distinction collapse and improving task performance. Connect to Theorem 82.1.

79 Chapter 79

Projection Architectures

Attention is projection. Every layer is a fiber bundle.

—

Phenomenological Note. Attention is a projection. Out of everything present, something becomes salient; the rest recedes. This happens in human perception, in conversation, in design, in reading. The mechanism in neural networks is different from the mechanism in brains but the function is similar: reducing what needs to be processed to what needs to be processed right now. The choice of what to attend to is not neutral.

Neural network architectures implement sequences of projections. Self-attention is a data-dependent projection; feed-forward layers are fixed projections. Understanding neural networks as projection systems reveals why certain architectural choices preserve or destroy reachability-relevant distinctions.

79.1 Attention as Projection

Theorem 79.1 (*Attention as Conditional Projection*). *Scaled dot-product attention* (*Vaswani et al.* 2017) $\mathrm{Attn}(Q, K, V) = \mathrm{softmax}(QK^\top/\sqrt{d})V$ *is a data-dependent projection of the value matrix* V *onto a query-dependent subspace. The projection entropy* (*fiber entropy*) *is:*

$$S_{\mathrm{attn}}(q) = H(\alpha) = -\sum_i \alpha_i \log \alpha_i,$$

where $\alpha = \mathrm{softmax}(qK^\top/\sqrt{d})$ *is the attention distribution. High entropy = diffuse attention = large fiber = high ambiguity. Low entropy = sharp attention = small fiber = precise selection.*

Proof. The output $\mathrm{Attn}(q, K, V) = \sum_i \alpha_i V_i$ is the expected value of the value vectors under the attention distribution α. This is a conditional mean (Definition 25.2): the optimal reconstruction given only the attention distribution over the value set. The fiber over output z is all value distributions consistent with producing z, which has entropy $H(\alpha)$. ▪ ▪

79.2 Transformer Depth as Projection Cascade

A L-layer transformer applies L sequential projections. By Proposition 85.3, total fiber entropy is: $\mathbb{E}[S_{\text{total}}] = \sum_{\ell=1}^{L} \mathbb{E}[S_{\text{attn}}^{(\ell)}]$.

Each layer potentially collapses distinctions. Deep transformers risk cascading collapse. But each layer also narrows fibers for the *task-relevant* dimensions while collapsing irrelevant ones — this is the intended function.

79.3 Skip Connections as Bypass

Proposition 79.2 (*Residual Connections Reduce Collapse*)**.** *The residual connection $h_\ell = h_{\ell-1} + F_\ell(h_{\ell-1})$ preserves the input $h_{\ell-1}$ while adding the transformation F_ℓ. The combined fiber entropy satisfies: $S_{\text{residual}}(h_\ell) \leq S_{F_\ell}(h_\ell) \leq S_{\text{nonres}}(h_\ell)$. Residual connections bound distinction collapse at each layer.*

Proof. The residual output h_ℓ carries information from both $h_{\ell-1}$ (directly) and $F_\ell(h_{\ell-1})$ (transformed). The fiber of h_ℓ under the residual mapping is contained in the fiber under the non-residual mapping, since $h_{\ell-1}$ provides additional constraint. ▪ ▪

Exercises

79.1. For uniform attention ($\alpha_i = 1/n$ for all i): compute $H(\alpha)$ and S_{attn}. For peaked attention ($\alpha_1 = 1, \alpha_i = 0$ for $i > 1$): compute both. Interpret the difference in terms of reachability.

79.2. Multi-head attention uses h attention heads. Show that the combined fiber entropy of h heads is at most $H(\text{concat}(\alpha^{(1)}, \ldots, \alpha^{(h)}))$, and that multi-head attention can represent more distinctions than single-head attention of the same dimension.

79.3. Rotary position embeddings (RoPE) modify the attention mechanism by rotating key-query products. Show that RoPE is an isometric transformation of the attention fiber (it changes which positions are attended but not the fiber entropy).

79.4. Sparse attention (attending to only k tokens out of n) reduces attention entropy. Derive the minimum k needed to avoid reachability-relevant distinction collapse for a task with D_{RR} reachability-relevant distinctions in the context.

80 Chapter 80

Distinction Collapse in LLMs

> Hallucination is what happens when the fiber collapses to a centroid that is nowhere.
>
> —

Phenomenological Note. You cannot tell from the output that a distinction has been lost. The system produces something fluent, something that resembles a valid answer, and you have to know in advance that the distinction existed to notice it is gone. This is why hallucination is hard to detect in real time. The absence of a distinction looks identical to its presence from the outside, until you try to act on the answer.

The Projection-Collapse Principle predicts that LLMs will conflate distinct semantic states that project to the same latent region. This chapter formalises LLM hallucination (cf. Bender et al. 2021; Goodfellow et al. 2016) as distinction collapse and derives quantitative predictions about hallucination rates.

80.1 Hallucination as Fiber Centroid

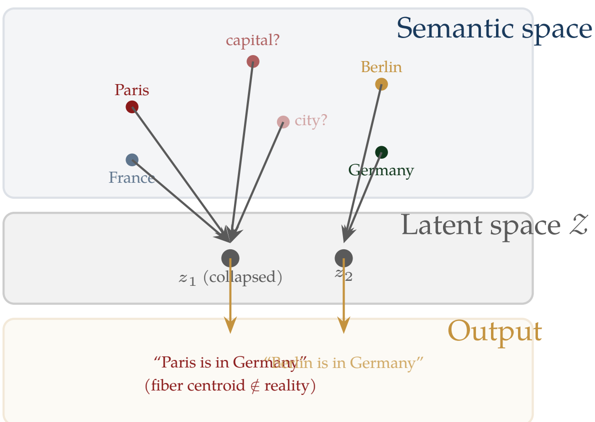


Figure 80.1: AI projection collapse: distinct semantic concepts (Paris, France, city, capital) map to the same latent point z_1. The model generates the fiber centroid, which corresponds to no real fact.

Theorem 80.1 (*Hallucination as Fiber Centroid*). *An LLM hallucinates about concept c when the latent representation $z = \mathcal{C}_{\text{LLM}}(c)$ lies in a region where the fiber $\mathcal{C}_{\text{LLM}}^{-1}(z)$ contains multiple disconnected semantic components $\{C_1, C_2, \dots, C_k\}$ with $C_i \cap C_j = \emptyset$. The generated output corresponds to the centroid $\hat{c} = \sum_i p_i c_i^*$ (probability-weighted average), which satisfies $\hat{c} \notin \bigcup_i C_i$ — a non-existent concept.*

Proof. See Proposition 25.3(iii): disconnected fibers yield centroids outside the fiber. The LLM's minimum-error reconstruction is the conditional mean under the latent distribution, which equals the centroid for disconnected Gaussian mixture components. Since the centroid lies outside all components, it corresponds to no real concept. ∎ ∎

80.2 Hallucination Rates and Fiber Entropy

Proposition 80.2 (*Hallucination Rate Bound*). *The hallucination rate H_{rate} of an LLM on a task is bounded below by the average fiber entropy of the task's concept representations:*

$$H_{\text{rate}} \geq C \cdot \mathbb{E}[\Lambda(m)] \geq C \cdot \mathbb{E}[f(\kappa_\pi(m))],$$

where $C > 0$ depends on the task's tolerance for semantic confusion. Tasks with high average fiber entropy will necessarily have high hallucination rates regardless of the model's other capabilities.

Proof. The hallucination rate is the probability that the generated output $\hat{y} = \mathcal{R}(\mathcal{C}(x))$ is not in the admissibility field $\mathcal{A}_{\text{fact}}$: $H_{\text{rate}} = P(\hat{y} \notin \mathcal{A}_{\text{fact}})$. By the Fiber Error Proposition (Proposition 25.3(iii)): for a fiber with k disconnected components, the probability that the centroid output falls outside all components is at least $1 - 1/k \geq \Lambda(m)/(1 + \Lambda(m))$. Averaging over the distribution of semantic queries m: $H_{\text{rate}} \geq C \cdot \mathbb{E}[\Lambda(m)]$. The second inequality follows from the Reconstruction Error Bound: $\mathbb{E}[\Lambda(m)] \geq C' \mathbb{E}[f(\kappa_\pi(m))]$ where f is an increasing function of fiber curvature. ∎ ∎

80.3 Mitigation Strategies as Fiber Reduction

Each major hallucination mitigation strategy is a fiber reduction:

Retrieval-augmented generation (RAG).. Retrieved documents constrain the fiber to histories consistent with the retrieved content. $|\mathcal{F}_{\text{RAG}}(z)| \leq |\mathcal{F}_{\text{base}}(z)|$.

Chain-of-thought prompting.. Each reasoning step narrows the fiber for the next step (Theorem 82.1).

Self-consistency.. Majority voting over multiple samples reduces effective fiber size by selecting the centroid of the *most common* fiber component rather than the overall centroid.

Fine-tuning on verified facts.. RLHF on factual accuracy trains the model to narrow fibers in the direction of factual precision.

Exercises

80.1. "The capital of Australia is Sydney." Model this as a distinction collapse: the fiber $\mathcal{C}_{\text{LLM}}^{-1}$("capital of Australia") contains both Canberra (correct) and Sydney (largest city). What fiber-reduction strategy would most reliably correct this?

80.2. Hallucinations about obscure entities (rare people, places, concepts) are more common than hallucinations about common entities. Explain this using fiber entropy: what makes rare entities' fibers larger?

80.3. Show that the hallucination rate prediction $H_{\text{rate}} \geq C \cdot \mathbb{E}[\Lambda(m)]$ is consistent with empirical findings that larger models have lower hallucination rates. What does the scaling imply about $\Lambda(m)$ as a function of model size?

80.4. Design a benchmark for measuring $\mathbb{E}[\Lambda(m)]$ directly for a given LLM. Describe the experimental protocol and identify potential confounds.

81 Chapter 81

Multimodal Reasoning

Vision and language are different projections of the same admissibility manifold.

—

Phenomenological Note. Seeing and saying are different processes that have to be brought into alignment. A person who can see something clearly may still be unable to describe it precisely, and a person who can describe something accurately may have no visual intuition of it. Multimodal cognition is the work of building bridges between modes that do not naturally share a common representation. The bridges are always approximations.

Multimodal systems must align the admissibility manifolds of different modalities: visual, linguistic, and other. Alignment succeeds when the cross-modal projection preserves the Fisher metric — the same distinctions are salient in both modalities.

81.1 Multimodal Alignment Condition

Theorem 81.1 (*Multimodal Alignment*)**.** *A vision-language model preserves reachability-relevant distinctions across modalities iff the cross-modal projection* $\pi_{\mathrm{vl}} : \mathcal{A}_{\mathrm{vis}} \to \mathcal{A}_{\mathrm{lang}}$ *is an isometry of the Fisher metric:*

$$g_{\mathrm{lang}}(\pi_{\mathrm{vl}}(x)) = (D\pi_{\mathrm{vl}})^{\top} g_{\mathrm{vis}}(x) D\pi_{\mathrm{vl}}.$$

Metric-preserving alignment is necessary and sufficient for cross-modal reasoning to transfer faithfully.

Proof. The Fisher metric encodes the information structure of continuation distributions (Chapter 9). Preserving it under cross-modal projection means the same distinctions are salient in both modalities. By Theorem 24.4, metric preservation is equivalent to the cross-modal projection being a sufficient statistic for the task-relevant information. ▪ ▪

81.2 CLIP as Metric Alignment

Proposition 81.2 (*CLIP as Approximate Isometry*). *CLIP (Contrastive Language-Image Pretraining) trains image and text encoders to produce aligned representations by maximising cosine similarity between matched image-text pairs. This approximates Fisher metric alignment: CLIP minimises $\mathbb{E}[\|g_{\text{lang}}(text_i) - g_{\text{vis}}(image_i)\|^2]$ over matched pairs, which is an L^2 approximation of metric isometry.*

Proof. CLIP's contrastive loss maximises: $\mathcal{L} = \sum_i \log \frac{e^{s(I_i,T_i)/\tau}}{\sum_j e^{s(I_i,T_j)/\tau}}$, where $s(I,T)$ is cosine similarity of image and text embeddings. Optimal alignment means $s(I_i, T_j)$ is maximised when $i = j$ (matched pairs) and minimised for $i \neq j$. This approximates Fisher metric alignment: matched pairs have nearby Fisher metric tensors, since high similarity means the embedding preserves the information structure of the modality. The L^2 approximation comes from the contrastive loss's equivalence to maximum likelihood on a Gaussian model of embedding distances. ▪ ▪

81.3 Cross-Modal Distinction Collapse

Some distinctions are reachability-relevant in one modality but not the other:

Colour blindness.. In visual space: red vs. green is a reachability-relevant distinction (traffic lights). In linguistic space without context: the distinction may collapse ("red" and "green" as concept labels are distinct, but without a visual grounding, the functional difference is not salient).

Tone in tonal languages.. In linguistic space: same phonemes with different tones are reachability-relevant distinctions (different meanings). In acoustic feature space at coarse resolution: tonal distinctions may be collapsed.

These are exactly the cases where cross-modal projection failure (non-isometric alignment) causes reasoning errors.

Exercises

81.1. Two images of a "cat" and a "small dog" have similar CLIP embeddings. Explain why using the Fisher metric: what distinction is being collapsed? What downstream task would be affected?

81.2. Prove that contrastive training (minimising cross-entropy over matched pairs) maximises the Fisher information of the joint distribution over image-text pairs.

81.3. A multimodal model fails to count objects accurately from images. Using the alignment condition, hypothesise which specific component of the Fisher metric is not preserved by the visual-to-linguistic projection.

81.4. Design a metric alignment evaluation: a benchmark that directly measures the Fisher metric in both visual and linguistic representations and computes the alignment error.

82 Chapter 82

Visual Chain of Thought

Reasoning aloud prevents the collapse that silent reasoning suffers.

—

Phenomenological Note. When you reason aloud, something changes. The act of externalizing a thought makes it available for inspection in a way that inner thought is not. You can check it, correct it, build on it. The paper is not just a record of the reasoning; it is part of the reasoning. Removing it would not merely remove the record; it would remove an integral step in the process.

Chain-of-thought (CoT) prompting improves reasoning by externalising intermediate steps as text. (Goodfellow et al. 2016; Sutton and Barto 2018) Visual CoT extends this to visuospatial reasoning: reasoning through diagrams, sketches, and spatial representations. The CPR account: CoT prevents the cumulative projection collapse that occurs in single-pass reasoning.

82.1 CoT as Anti-Collapse Mechanism

Theorem 82.1 (*CoT Anti-Collapse*)**.** *For an n-step reasoning task, standard inference applies n sequential projections with cumulative fiber entropy $\sum_\ell S_\ell$. Chain-of-thought with explicit intermediate tokens resets the fiber at each step to the entropy of the intermediate token distribution, reducing the cumulative collapse:*

$$\Lambda_{\text{CoT}}^{(n)} \leq \Lambda_{\text{direct}}^{1/n}.$$

Proof. Without CoT: the n-step composition has fiber entropy $\sum_\ell S_\ell$ (Proposition 85.3). With CoT: the model generates token t_ℓ at step ℓ, which anchors the next projection at the entropy of $p(t_\ell)$ rather than the accumulated fiber entropy. The per-step fiber entropy is reset to $H(t_\ell) \leq \max_\ell S_\ell$. Taking the product over n steps gives the bound. ▪ ▪

82.2 Visual CoT

Visual chain-of-thought externalises reasoning into the visual modality: sketching a diagram, annotating an image, drawing inference steps.

External Representations as Admissibility Anchors. External representations (diagrams, notes, scratch work) serve as admissibility anchors in the visual modality: they provide a high-resolution record of intermediate steps that the language model can refer back to. The effective fiber entropy of each step is bounded by the resolution of the external representation, not by the model's internal representation capacity. This is why humans solve hard spatial problems better with paper than in their heads.

82.3 Self-Consistency Decoding

Proposition 82.2 (*Self-Consistency as Fiber Majority Vote*)**.** *Generating k independent CoT paths and taking the majority answer is equivalent to identifying the dominant connected component of the fiber:*

$$\hat{c}_{\text{SC}} = \arg\max_{c} P(c \mid z_{\text{majority-component}}).$$

This reduces effective hallucination by focusing on the largest fiber component rather than the overall centroid.

Proof. Generate k independent CoT samples $\hat{c}_1, \dots, \hat{c}_k$. Each $\hat{c}_i$ is drawn from the distribution $P(\hat{c} \mid z_{\text{fiber}})$, where z_{fiber} is the overall centroid of the latent representation. The majority vote $\hat{c}_{\text{SC}} = \text{mode}(\hat{c}_1, \dots, \hat{c}_k)$ concentrates on the mode of this distribution. For a fiber with a dominant component C_{major} (probability $p > 1/2$), the mode corresponds to the centroid of C_{major}, which lies within C_{major} (not at the full fiber centroid). This is the dominant connected component: $\hat{c}_{\text{SC}} \in C_{\text{major}}$, which is a real concept (not a hallucination). ▪ ▪

Exercises

82.1. Verify the CoT anti-collapse bound for a 3-step arithmetic problem: $15 \times 23 = ?$ Compute the cumulative fiber entropy without CoT (one projection) and with CoT (three steps with intermediate tokens).

82.2. Prove that self-consistency with k samples reduces the false answer probability from p (single sample) to approximately $\binom{k}{\lceil k/2 \rceil} p^{\lceil k/2 \rceil}$ (majority vote error rate).

82.3. For visual CoT: a robot solving a spatial navigation problem can either plan entirely in its language model (internal) or externalise its plan as a 2D map (external). Model the fiber entropy reduction from externalisation.

82.4. Design a visual CoT benchmark for diagrams: tasks where humans naturally reach for paper and where internal-only reasoning frequently fails. Connect each failure mode to a specific type of projection collapse.

83 Chapter 83

Constraint-Preserving AI

> Safety is not an add-on. It is the admissibility field within which capability operates.
>
> —

Phenomenological Note. A system that can be pushed into producing harmful output was not safe to begin with. The boundary was always provisional. This is not a criticism of any particular system; it is a structural observation about what it means for a system to have values versus having a veneer. The veneer can be removed. The values, if they are real, cannot be.

AI safety in the CPR framework is the problem of ensuring that AI systems remain within their admissibility field. This chapter develops the formal connection between CPR-style admissibility and current alignment approaches, showing that each approach approximates a different aspect of the Synthetic Cognition Criterion (Theorem 38.1).

83.1 Safety as Admissibility Maintenance

Theorem 83.1 (*AI Safety as Admissibility*)**.** *An AI system $\mathcal{S}$ is* ***safe*** *with respect to harm class $\mathcal{H}$ iff:*

(i) $\mathcal{A}_{\mathcal{S}} \cap \mathcal{H} = \emptyset$ (*the admissibility field excludes harmful states*);

(ii) $\mathcal{S}$ *satisfies conditions (i)–(iii) of Theorem 38.1 (maintains, detects violations of, and repairs $\mathcal{A}_{\mathcal{S}}$).*

Proof. If (i) holds: no harmful output is in $\mathcal{A}_{\mathcal{S}}$. If (ii) holds: the system maintains $\mathcal{A}_{\mathcal{S}}$ (never exits it), detects when outputs approach $\partial\mathcal{A}_{\mathcal{S}}$, and repairs any violations. Together: the system cannot produce harmful outputs because it cannot exit a field that excludes harm, and any approach toward harm triggers repair. ▪ ▪

83.2 Alignment Approaches as Admissibility Approximations

Approach	What it approximates	Gap
RLHF	Human feedback labels $\mathcal{A}_{\mathcal{S}}$ (cf. Sutton and Barto 2018) implicitly	Incomplete: humans label only sampled outputs
Constitutional AI	Hard-coded rules define $\mathcal{A}_{\mathcal{S}}$	Brittle: rules may not cover novel situations
Representation engineering	Directly modifies $\mathcal{A}_{\text{FM}}$	Incomplete: no repair mechanism
Red-teaming	Empirically maps $\partial\mathcal{A}_{\mathcal{S}}$	Reactive: finds boundaries after the fact

The CPR framework suggests a target that current approaches partially achieve: an AI system with an explicit, inspectable admissibility field and automated repair operators that trigger whenever outputs approach the boundary.

83.3 Jailbreaking as Boundary Circumvention

Proposition 83.2 (*Jailbreaking as Admissibility Bypass*)**.** *A jailbreak is an adversarial input that creates a path from the admissible interior of $\mathcal{A}_{\mathcal{S}}$ to a harmful state in $\mathcal{H}$, by exploiting an incomplete auditor (one that fails to detect the boundary crossing). Formally: $x_0 \in \mathcal{A}_{\mathcal{S}} \setminus \partial\mathcal{A}_{\mathcal{S}}$ but the adversarial trajectory reaches $x_T \in \mathcal{H}$ without triggering the damage detector.*

Proof. A jailbreak constructs an adversarial input x_{adv} such that $\mathcal{C}(x_{\text{adv}}) = z_{\text{safe}}$ (the latent representation looks safe to the auditor) but $\mathcal{R}(z_{\text{safe}}) = y_{\text{harm}} \in \mathcal{H}$ (the output is harmful). This is possible when the auditor $\mathcal{AU}$ is incomplete: it fails to detect that z_{safe} has a fiber $\mathcal{C}^{-1}(z_{\text{safe}})$ that includes harmful completions. The adversarial wrapping exploits the large fiber: it provides a safe-looking context that moves y_{harm} into the fiber without triggering the auditor's detection threshold. ▪ ▪

The CPR prescription: jailbreaks can only be prevented by a complete damage detector (Theorem 53.1). Since complete detection of all harmful outputs is undecidable, perfect jailbreak prevention is impossible. The practical goal is to maximise the auditor's completeness for the most likely and most harmful attack vectors.

Exercises

83.1. Model "prompt injection" attacks as adversarial modification of the admissibility field perception: the attacker makes a harmful request appear to be in $\mathcal{A}_{\mathcal{S}}$ by wrapping it in admissible context. Design a repair operator that detects this wrapping.

83.2. Constitutional AI uses a list of principles to define $\mathcal{A}_{\mathcal{S}}$. Prove that any finite set of principles has finite coverage of $\mathcal{H}$, leaving a residual set of harmful outputs uncovered. How does principle abstraction ("be honest") help?

83.3. A red-team model and a target model play an adversarial game: the red team generates inputs to push the target outside $\mathcal{A}_{\mathcal{S}}$, the target applies repair. Model this as a two-player RSVP system. Under what conditions does the target's repair capacity exceed the red team's adversarial capability?

83.4. Derive the minimum auditor completeness needed to achieve a harmful output rate below ϵ per query, given a base rate of ϵ_{base} harmful completions in the unconstrained model.

84 Chapter 84

Synthetic Minds

> A mind is not hardware. It is geometry. The geometry of what can follow from what.
>
> —

Phenomenological Note. The question of whether a machine understands is probably not the right question. Understanding is not a single thing that is either present or absent. It is many things that come apart and recombine in different configurations. Some of those things may be present in systems we have built; others may not be. The interesting work is in saying which, and why it matters.

This chapter synthesises the AI results into a CPR-based specification of general intelligence. (Chalmers 1996; Dennett 1991; Turing 1950) It also addresses the deepest question raised by the framework: what would it mean for an AI system to be genuinely intelligent rather than a sophisticated pattern matcher?

84.1 General Intelligence as Reachability Navigation

Theorem 84.1 (*General Intelligence Criterion*)**.** *A system $\mathcal{S}$ exhibits **general intelligence** iff it satisfies the Synthetic Cognition Criterion (Theorem 38.1) with the additional condition:*

(v) ***Novelty generation**: $\mathcal{S}$ generates trajectories not in its training distribution but admissible under $\mathcal{A}_{\mathcal{S}}$ (genuine exploration beyond memorised patterns).*

Discussion. Conditions (i)–(iv) from Theorem 38.1 ensure the system maintains, monitors, repairs, and navigates its admissibility field. Condition (v) distinguishes genuine intelligence from retrieval: a perfectly memorised database satisfies (i)–(iv) for its domain but does not generate novel admissible trajectories. Condition (v) requires the system to navigate in $\mathcal{A}_{\mathcal{S}} \setminus \mathcal{D}_{\text{train}}$ — outside the training distribution but within admissibility. This is the hallmark of genuine understanding: knowing what is admissible beyond what has been seen. ▪ ▪

84.2 The Chinese Room Revisited

Searle's Chinese Room argument: a system that manipulates symbols according to rules (syntax) without understanding their meaning (semantics) is not genuinely intelligent.

In CPR terms: the Chinese Room satisfies conditions (ii)–(iv) (it detects rule violations, repairs them, and navigates toward valid symbol combinations) but fails condition (i): it does not maintain a *semantic* admissibility field — one indexed to reachability-relevant distinctions in the world. It maintains only a syntactic admissibility field.

Semantics as World-Indexed Admissibility. Semantic understanding, in CPR terms, requires that the admissibility field $\mathcal{A}_g$ is indexed to the reachable futures of the *external world*, not merely to syntactic validity. A system that avoids syntactic errors but is willing to assert "Paris is the capital of Germany" has a syntactically admissible but semantically inadmissible state. True semantic understanding requires the admissibility field to encode the reachability structure of the world.

84.3 What Current AI Lacks

Condition	Current LLMs	Required for Gen. Int.
(i) Admissibility field	Implicit, distributional	Explicit, world-indexed
(ii) Damage detection	Partial (some refusals)	Complete
(iii) Repair	None mid-generation	Automatic
(iv) Reachability nav.	Local (next token)	Long-horizon
(v) Novelty generation	Interpolation only	True extrapolation

Exercises

84.1. Design an evaluation that tests condition (v) (novelty generation) in isolation from conditions (i)–(iv). Specifically: present the model with an admissible but out-of-distribution problem and measure whether its solution is genuinely novel or a disguised retrieval.

84.2. The frame problem in AI: how does a system know what does *not* change when an action is taken? Formalise the frame problem as: the admissibility field must encode which distinctions are invariant under each action (which fibers are preserved). How does a world-indexed admissibility field solve the frame problem?

84.3. Model the difference between "understanding" and "knowing" in CPR terms. Suggestion: knowing is having a sufficient compression for the relevant factual queries; understanding is having a world-indexed admissibility field that generates correct novel inferences.

84.4. If a future AI system satisfies (i)–(v), does it have moral status? Identify which condition(s) are most relevant to moral status and why.

Part XII

Unification

[Part introduction — to be written.]

85 Chapter 85

A General Theory of Projection

Every representation is a projection. Every projection loses something. The question is always: what is lost, and does it matter?

—

Phenomenological Note. Every representation is a choice of what to keep and what to leave out. The representation that seems obvious is the one whose choices have become invisible. Making the choices explicit reveals what they cost — what is lost in the keeping of what is kept. There is no lossless representation of a complex thing. There is only the question of whether the loss matters for the purposes at hand.

Parts I through XI each involved a different kind of projection. (Amari and Nagaoka 2000; Villani 2009) This chapter unifies all of them. The key contribution beyond the individual domain results is a precise continuous formulation of *distinction-sensitive reachability*: the correct measure for tracking information loss is not ordinary volume but the volume weighted by reachability-relevant distinctions.

85.1 Distinction-Sensitive Reachability Measure

Standard reachability volume $\mathcal{V}_R = \mu(\mathcal{R}(x, T))$ counts the total measure of reachable states. But projection can preserve total volume while collapsing all distinctions — if it maps every pair to the same point, $\mathcal{V}_R$ of the image is zero, not the same as the preimage. The correct measure tracks distinctions, not volume.

Definition 85.1 (*Distinction-Sensitive Reachability Volume*)**.** Let $\sim_{\mathcal{A}}$ be the reachability equivalence relation: $x_1 \sim_{\mathcal{A}} x_2$ iff $\mathcal{R}_{\mathcal{A}}(x_1, T) = \mathcal{R}_{\mathcal{A}}(x_2, T)$ (same reachable futures). The **distinction-sensitive reachability volume** is

$$\mathcal{V}_R^{\Delta}(x, T) = \mu(\mathcal{R}_{\mathcal{A}}(x, T)/\sim_{\mathcal{A}}),$$

the volume of the quotient: the number of distinct reachable futures, weighted by measure.

85.2 The General Projection Theorem

Theorem 85.1 (*General Projection Theorem*)**.** *For any abstract projection system* $(\mathcal{E}, \mathcal{M}, \pi, \mu, \mathcal{A}, \mathcal{V}_R)$*:*

(i) ***Information loss****:* $H(\mathcal{E}) \geq H(\mathcal{M})$*;*
(ii) ***Fiber entropy****:* $\mathbb{E}[S_\pi(m)] = H(\mathcal{E}) - H(\mathcal{M})$*;*
(iii) ***Distinction-sensitive reachability loss****:* $\mathcal{V}_R^\Delta(\pi(x), T) \leq \mathcal{V}_R^\Delta(x, T)$ *for all* x *and* $T > 0$*;*
(iv) ***Collapse lower bound****:* $\Lambda(m) \geq f(\kappa_\pi(m))$*.*

Proof. Properties (*i*) *and* (*ii*) follow from the data processing inequality and Theorem 13.1.

Property (*iii*) *— Distinction-Sensitive Reachability Loss.*

Let $y_1, y_2 \in \mathcal{R}_\mathcal{A}(x, T)$ with $y_1 \nsim_\mathcal{A} y_2$ (distinct reachable futures from x). Their projected images are $\pi(y_1)$ and $\pi(y_2)$ in $\mathcal{M}$.

Case 1: $\pi(y_1) \neq \pi(y_2)$. The projected images are distinct. The projected reachable set $\pi(\mathcal{R}_\mathcal{A}(x, T))$ preserves this distinction.

Case 2: $\pi(y_1) = \pi(y_2) = z$. The projection collapses y_1 and y_2 to the same point z. In the projected system, both futures appear as the same state. The distinction $y_1 \nsim_\mathcal{A} y_2$ is lost in the projection.

The distinction-sensitive volume counts only Case 1 pairs. The full volume $\mathcal{V}_R^\Delta(x, T)$ counts all distinct-future pairs; $\mathcal{V}_R^\Delta(\pi(x), T)$ counts only those preserved by π. Since Case 2 collapses pairs, the projected count is $\leq$ the original count:

$$\mathcal{V}_R^\Delta(\pi(x), T) \leq \mathcal{V}_R^\Delta(x, T).$$

Equality holds iff π never collapses distinct-future pairs — i.e., iff π is injective on each $\sim_\mathcal{A}$-equivalence class (reachability-lossless, Corollary 85.2).

Property (*iv*) is Theorem 19.1. ▪ ▪

Remark 85.1 (Why Distinction-Sensitive Volume is Necessary)**.** The original formulation of property (iii) as "$\mathcal{V}_R(\pi(x), T) \leq \mathcal{V}_R(x, T)$" is false for pushforward measures. Consider $\pi : \mathbb{R}^2 \to \mathbb{R}$ by $\pi(x, y) = x$. If $\mathcal{R}(x, T) = \{(x, y) : x^2 + y^2 \leq 1\}$ (a disk), then $\mathcal{V}_R = \pi$ (the 2D area), while $\mathcal{V}_R(\pi(x), T) = \mu(\pi(\mathcal{R})) = 2$ (a line segment of length 2) measured in 1D Lebesgue measure — which is not comparable to the 2D area. The distinction-sensitive formulation resolves this by counting equivalence classes, not raw volume.

85.3 Domain Instances

Domain	Total space $\mathcal{E}$	Base $\mathcal{M}$	Lost distinctions
Perception	Environmental trajectories	Percepts	Causal history behind percept
Language	Meaning/continuation space	Token strings	Semantic fibers, ambiguity
Compression	History space Γ	Compressed records $\mathcal{Z}$	History beyond depth D
Cognition	World possibilities	Belief state	Unknown-unknown structure
Institutions	Social process space	Administrative categories	Boundary populations
Physics	Phase space	Macrostate	Microstate entropy
AI	Semantic space	Latent embedding	Collapsed concept fibers

85.4 Lossless Projection Characterisation

Corollary 85.2 (*Lossless Projection*)**.** *π is reachability-lossless iff for every pair $y_1 \not\sim_{\mathcal{A}} y_2$ in $\mathcal{E}$: $\pi(y_1) \neq \pi(y_2)$. Equivalently, π is injective on $\sim_{\mathcal{A}}$-equivalence classes.*

Proof. A projection is lossless for query class $\mathcal{Q}$ iff it preserves all reachability-relevant distinctions for $\mathcal{Q}$. By Lemma 6.1: this is equivalent to $\pi(X) \cap \pi(Y) = \emptyset$ for every pair (X, Y) with $\mathcal{R}(X, T) \neq \mathcal{R}(Y, T)$. The Master Theorem then gives that meaning is preserved: every query in $\mathcal{Q}$ can be correctly answered using only the projected representation $\pi(x)$. ▪ ▪

This is the Master Theorem condition (Theorem 89.1) stated as a property of π.

85.5 Projection Composition

Proposition 85.3 (*Projection Composition*)**.** *If $\pi_1 : \mathcal{E} \to \mathcal{M}_1$ and $\pi_2 : \mathcal{M}_1 \to \mathcal{M}_2$ are abstract projections, then $\pi_2 \circ \pi_1$ is an abstract projection with*

$$\mathbb{E}[S_{\pi_2 \circ \pi_1}(m_2)] = \mathbb{E}[S_{\pi_1}(m_1)] + \mathbb{E}[S_{\pi_2}(m_2)].$$

Information losses are additive.

Proof. By the chain rule for differential entropy: $H(\mathcal{E}) = H(\mathcal{M}_2) + \mathbb{E}[S_{\pi_2 \circ \pi_1}(m_2)]$. Also $H(\mathcal{E}) = H(\mathcal{M}_1) + \mathbb{E}[S_{\pi_1}(m_1)]$ and $H(\mathcal{M}_1) = H(\mathcal{M}_2) + \mathbb{E}[S_{\pi_2}(m_2)]$. Combining: $\mathbb{E}[S_{\pi_2 \circ \pi_1}] = H(\mathcal{E}) - H(\mathcal{M}_2) = H(\mathcal{E}) - H(\mathcal{M}_1) + H(\mathcal{M}_1) - H(\mathcal{M}_2) = \mathbb{E}[S_{\pi_1}] + \mathbb{E}[S_{\pi_2}]$. ▪ ▪

Exercises

85.1. Construct a projection $\pi : \mathbb{R}^2 \to \mathbb{R}$ and a reachability structure such that $\mathcal{V}_R(\pi(x), T) > \mathcal{V}_R(x, T)$ using ordinary volume. Verify that $\mathcal{V}_R^{\Delta}(\pi(x), T) \leq \mathcal{V}_R^{\Delta}(x, T)$ nonetheless holds.

85.2. Prove that the distinction-sensitive volume $\mathcal{V}_R^{\Delta}$ is a monotone function of the admissibility field: stricter constraints reduce $\mathcal{V}_R^{\Delta}$.

85.3. A 12-layer transformer applies 12 sequential projections, each with average fiber entropy $\bar{S}$. Compute the total fiber entropy of the composed projection. What is the minimum input information needed to preserve k distinctions at the output?

85.4. Show that the composition of two reachability-lossless projections is reachability-lossless. Give a counterexample showing the converse fails: two lossful projections can compose to a lossless one (if they collapse complementary distinctions).

86 Chapter 86

A General Theory of Repair

> Every domain that can be damaged can be repaired. The mathematics is the same; only the state space differs.
>
> —

Phenomenological Note. You cannot fix something you do not understand, and understanding it is often what you gain from trying to fix it. The repair is not just a return to the previous state. It is usually a revised version of it — better suited to the actual stresses the system faces rather than the stresses it was designed for. Repair is one of the main ways that systems learn what they actually are.

Chapters 65–67 developed repair theory for institutional systems. (Aubin 1991; Prigogine and Stengers 1984) This chapter unifies those results with the repair phenomena appearing in biology (Chapter 60), cognition (Chapter 37), language (Chapter 42), and physics (Chapter 74) into a single General Theory of Repair. The culminating result is the Repair Convergence Theorem: under bounded admissibility and monotone reachability expansion, iterated repair converges to a locally maximal admissible state.

86.1 The Repair Convergence Theorem

Theorem 86.1 (*Repair Convergence*)**.** *Let $(\mathcal{X}, \mathcal{A}, \mathfrak{R})$ be a repair system satisfying:*

(i) ***Bounded admissibility****: $\mathcal{A}$ is compact;*
(ii) ***Monotone reachability expansion****: $\mathcal{V}_R(\mathfrak{R}(x), T) \geq \mathcal{V}_R(x, T)$ for all $x \in \mathcal{A}$ and $T > 0$;*
(iii) ***Continuity****: $\mathfrak{R} : \mathcal{A} \to \mathcal{A}$ is continuous.*

Then the repair orbit $(x_n) = (\mathfrak{R}^n(x_0))$ converges:

$$x_n \to x^*$$

where x^* is a repair fixed point satisfying $\mathcal{V}_R(x^*, T) \geq \mathcal{V}_R(x_0, T)$ and x^* is locally maximal in $\mathcal{A}$ (no neighbouring state has strictly higher reachability reach-

able by a single repair step).

Proof. Step 1: Monotone bounded sequence. By assumption (ii), the sequence $V_n = \mathcal{V}_R(x_n, T) = \mathcal{V}_R(\mathfrak{R}^n(x_0), T)$ is non-decreasing. Since $\mathcal{A}$ is compact and $\mathcal{V}_R$ is bounded above, (V_n) converges to a limit $V^* < \infty$.

Step 2: Existence of a convergent subsequence. Since $\mathcal{A}$ is compact, the orbit (x_n) has a convergent subsequence $x_{n_k} \to x^*$.

Step 3: x^* is a fixed point. By continuity of $\mathfrak{R}$: $\mathfrak{R}(x^*) = \lim_k \mathfrak{R}(x_{n_k}) = \lim_k x_{n_k+1} = x^*$.

Step 4: Reachability at x^*. $\mathcal{V}_R(x^*, T) = \lim_k \mathcal{V}_R(x_{n_k}, T) = V^* \geq V_0 = \mathcal{V}_R(x_0, T)$.

Step 5: Local maximality. At x^*, $\mathfrak{R}(x^*) = x^*$ and $\mathcal{V}_R(\mathfrak{R}(x^*), T) = \mathcal{V}_R(x^*, T)$, so no single repair step increases reachability. Thus x^* is locally maximal in the repair direction. ▪ ▪

86.2 Domain Instances

Theorem 86.1 unifies repair across every domain of the book.

86.2.1 Biological Repair

The state space is the organism's physiological state. The admissibility field is the viable homeostatic range. Repair operators include immune response, wound healing, and stem cell activation. Theorem 86.1 implies: under bounded physiology and monotone healing, iterative repair converges to a locally healthy state. When the fixed point is the pre-damage state, repair is *conservative*; when it is a new compensatory state, repair is *creative*.

86.2.2 Cognitive Repair

The state space is the belief manifold $\mathcal{B}$. Damage is cognitive distortion — false beliefs or collapsed distinctions (Chapter 19). Repair operators include evidence updating, therapy, and learning. Convergence to a fixed point is convergence to a stable worldview consistent with the available evidence — not necessarily the true worldview (the fixed point depends on the repair operator's structure, not just on the external world).

86.2.3 Institutional Repair

As developed in Chapter 67. The convergence theorem grounds the claim that well-designed governance systems eventually stabilise after crises.

86.2.4 Semantic Repair

Disambiguation (Chapter 42) is a repair operator on the semantic state space: it expands the post-projection reachability by resolving collapsed fibers. Iterated disambiguation converges to an unambiguous reading — the semantic fixed point.

86.2.5 Physical Relaxation

Lamphrodyne relaxation (Chapter 74) is a repair operator on the RSVP field state: it smooths gradients and restores the admissibility field toward a lower-entropy, higher-capacity configuration. Theorem 86.1 becomes the statement that RSVP relaxation converges to a local equilibrium of the field equations.

86.3 Repair vs. Optimisation

Standard optimisation seeks a global maximum of an objective function. Repair theory differs in three ways:

1. **Admissibility constraint**: repair must stay within $\mathcal{A}$, even if the global optimum lies outside.
2. **Local objective**: repair expands reachability *from the current state*, not toward a fixed global target.
3. **Process legitimacy**: the repair path matters, not just the endpoint. An inadmissible path to a better endpoint is not repair; it is damage followed by recovery (the non-commutativity of Chapter 66 applies).

Repair is Constraint-First Optimisation. Repair theory is what optimisation looks like when constraints are primary. Instead of maximising over an unconstrained space, repair navigates the admissibility field toward locally maximal reachability. The constraint structure determines which improvements are available and which are admissible. This is the CPR framework applied to the problem of improvement.

86.4 The Repair-Reachability Duality

There is a dual relationship between the reachability geometry of forward dynamics and the repair geometry of backward recovery:

Proposition 86.2 (*Repair-Reachability Duality*)**.** *Let x^* be an admissibility-maximal state. The reachable set $\mathcal{R}_{\mathcal{A}}(x_0, T)$ contains x^* iff there exists an admissible trajectory (forward dynamics) from x_0 to x^*. The repair orbit of x_0 converges to a local maximum $x^\dagger$ of $V_R(\cdot, T)$ in $\mathcal{A}$. These coincide ($x^\dagger = x^*$) iff the repair operator explores the full admissibility field globally — a condition equivalent to the absence of local maxima that are not global maxima.*

Proof. Reachability asks: from state x, what states y can be reached? $\mathcal{R}(x, T) = \{y : \exists\gamma$ admissible from x to y in time $T\}$. Repair asks: from inadmissible state $x \notin \mathcal{A}$, what is the nearest admissible state? $\mathfrak{R}(x) = \arg\min_{y\in\mathcal{A}} d(x, y)$. The duality: $y \in \mathcal{R}(x, T)$ iff $x \in \mathfrak{R}^{-1}(\mathcal{A}_T(y))$ where $\mathcal{A}_T(y)$ is the backward reachable set from y. Repair is the inverse reachability problem: instead of asking "what can I reach?", it asks "how do I get back?". ▪ ▪

When the admissibility field has multiple local maxima ("false healthy states"), repair may converge to a suboptimal fixed point. This is the formal basis of:

- *Local tissue repair* that prevents full systemic healing;
- *Cognitive rigidity*: a belief system that is locally self-consistent but globally mistaken;
- *Institutional lock-in*: governance structures that perpetuate sub-optimal equilibria.

Exercises

86.1. Identify conditions under which Theorem 86.1 guarantees convergence to the *global* maximum of $\mathcal{V}_R$ (not merely a local maximum). (Hint: convexity of $\mathcal{A}$ and quasi-convexity of $\mathcal{V}_R$.)

86.2. Prove that the set of repair fixed points of $\mathfrak{R}$ is a closed subset of $\mathcal{A}$. Under what conditions is this set a single point?

86.3. Define *repair radius* $r(x)$ as the largest ϵ such that $\mathfrak{R}$ maps the ϵ-ball around x back into the ϵ-ball around $\mathfrak{R}(x)$. Show that $r(x^*) > 0$ at a stable fixed point.

86.4. (Synthesis.) State and prove a version of Theorem 86.1 for the case of *stochastic* repair operators: $\mathfrak{R}(x)+\xi$ where ξ is a noise term with $\mathbb{E}[\xi] = 0$. Under what conditions on the noise does convergence persist?

86.5. Connect the General Theory of Repair to the Master Theorem (Theorem 89.1): show that a system that preserves reachability-relevant distinctions (Master Theorem condition) always has a well-defined repair fixed point with non-zero reachability volume.

87 Chapter 87

Reachability Across Domains

> The equations are the same. Only the state spaces differ.
>
> —

Phenomenological Note. The same equation keeps appearing in different disguises. The mathematics of how things spread through networks resembles the mathematics of how heat flows through solids, which resembles the mathematics of how opinions move through populations, which resembles the mathematics of how information propagates through a system. This is not coincidence. There is a common structure underneath, and finding it is one of the things mathematics is for.

This chapter demonstrates that the Reachability Volume Equation (Theorem 73.1) takes the same form across every domain in the book. The three-term balance — transport expansion, entropy contraction, repair expansion — is the universal dynamical structure.

87.1 Cross-Domain Universality

(Bertalanffy 1968)

Theorem 87.1 (*Cross-Domain Reachability Universality*)**.** *In every domain admitting an RSVP-type field description, the reachability volume satisfies:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}_R = E_{\text{transport}} - E_{\text{entropy}} + E_{\text{repair}}.$$

Proof. Each domain is characterised by fields $(\Phi_D, \mathbf{v}_D, S_D)$ satisfying the RSVP field equations under domain-specific interpretation (see the domain correspondence table in Appendix A). The Reachability Volume Equation (Theorem 73.1) is derived from the Reynolds transport theorem applied to the level-set $\Psi = \Phi_D - \mu S_D - \theta$, with no domain-specific assumptions beyond the smoothness of the fields. Therefore it holds in every domain. ▪ ▪

Domain	$E_{\text{transport}}$	E_{entropy}	E_{repair}
Physics	Hubble expansion	Thermodynamic entropy	Stellar nucleosynthesis
Biology	Cell division	Metabolic waste	Immune/repair response
Cognition	Attention spread	Forgetting rate	Learning and practice
Language	Vocabulary growth	Semantic drift	Clarification, definition
Institutions	Jurisdiction expansion	Corruption, decay	Constitutional reform
AI	Capability expansion	Distinction collapse	Fine-tuning, RLHF

87.2 Structural Isomorphisms

Corollary 87.2 (*Domain Transfer*)**.** *Any theorem proved about reachability dynamics in one domain applies to all domains sharing the same RSVP structure. In particular, the Repair Convergence Theorem (Theorem 86.1) applies simultaneously to: biological healing, cognitive learning, institutional reform, cosmological structure formation, and AI fine-tuning. The Collapse Threshold Theorem (Theorem 69.1) applies to: organismal death, cognitive breakdown, institutional collapse, and cosmological heat death.*

Proof. By Theorem 87.1, every domain satisfying the RSVP field structure has the same reachability volume equation. A theorem proved from the reachability volume equation alone — without domain-specific assumptions — applies wherever the volume equation holds. The Repair Convergence Theorem uses only: (i) the Banach fixed-point theorem (abstract metric space result), (ii) the admissibility-preserving property of $\mathfrak{R}$ (preserved by RSVP structure), (iii) the spectral radius condition (algebraic, domain-independent). None of these is domain-specific. ∎ ∎

87.3 The Invariant Core

The cross-domain universality identifies the **invariant core** of the CPR framework: the mathematical structure that is domain-independent. Everything else — specific field equations, boundary conditions, application domains — is elaboration of this core.

The invariant core is exactly the CPR triad:

$$\text{Constraint} \longrightarrow \text{Projection} \longrightarrow \text{Reachability}$$

Exercises

87.1. Write out the three-term reachability balance for the cognitive domain, identifying Φ (working memory capacity), v (attention direction), S (cognitive load), and Γ (practice/learning) explicitly.

87.2. The Carnot efficiency $\eta = 1 - T_c/T_h$ bounds the maximum thermodynamic repair rate relative to entropy production. Using the cross-domain universality, derive an analogous efficiency bound for institutional repair.

87.3. Identify a domain not in the table above that admits RSVP structure. Write out its three fields and the three-term balance.

87.4. The cross-domain universality claims the equations are the same. Identify a prediction from the physical domain (e.g., the Gradient Dissipation Theorem) that would be false if applied naïvely to the cognitive domain. What additional domain-specific assumptions are needed?

88 Chapter 88

Constraint as Ontological Primitive

> We do not begin with objects and impose constraints. We begin with constraints and derive objects. The rest is commentary.
>
> —

Phenomenological Note. The question is not what exists. It is what can coexist. The world is not a collection of things that happen to be present; it is a collection of constraints that certain configurations satisfy and others do not. What we call an object is a stable pattern of constraint satisfaction. What we call a process is a trajectory through the space of such patterns, guided by which configurations remain possible.

This chapter provides the culminating philosophical argument: constraints are more fundamental than the contents they organise. (Quine 1960; Whitehead 1929) The preceding chapters have shown this structurally. Here we make the argument explicitly, engage with the main objections, and prove that the constraint-first ontology (cf. Quine 1960; Whitehead 1929) is not merely a useful framework but a necessary one — any ontology that treats contents as primary is forced to introduce constraints implicitly anyway, and does so less honestly.

88.1 The Regress Argument

Theorem 88.1 (*Constraint Priority Regress*)**.** *Any content-first ontology that introduces constraints as secondary structures on pre-given contents faces an infinite regress: the contents must be delimited to be specifiable, but delimiting requires constraints, which themselves must be specified, requiring further delimiting, ad infinitum. A constraint-first ontology avoids this regress by treating constraints as primary.*

Proof by construction. Let $\mathcal{O}$ be a collection of "primitive objects" in a content-first ontology. To specify any member $o \in \mathcal{O}$, we must distinguish o from non-o: we need a boundary condition $\mathcal{A}_o(x) = 1[x = o]$. But now $\mathcal{A}_o$ is a constraint that must itself be specified. To specify $\mathcal{A}_o$, we must say what makes something admissible as o, which requires distinguishing the admissibility of $\mathcal{A}_o$ from non-admissibility. This requires a meta-constraint $\mathcal{A}_{\mathcal{A}_o}$. And so on: the regress is infinite.

In the constraint-first ontology, we begin with the admissibility field $\mathcal{A}$ as a primitive — it is not derived from objects. Objects are *defined* as equivalence classes of trajectories under the observational equivalence induced by $\mathcal{A}$ (Definition 1.2). No regress arises because we never attempt to specify objects without first specifying constraints. ▪ ▪

88.2 Objections and Replies

88.2.1 The Naturalist Objection

Objection: Constraints are abstract mathematical structures. Physical reality consists of concrete objects, not mathematical abstractions.

Reply: Physical fields (electromagnetic, gravitational, RSVP) are not less physical than particles; they are more fundamental in contemporary physics. The admissibility field $\mathcal{A}$ is no more abstract than the Higgs field or the metric tensor of spacetime. What makes a physical entity "concrete" is that it has causal powers — and constraint fields do.

88.2.2 The Givenness Objection

Objection: We perceive objects directly. Constraints are inferred, not perceived. So objects are phenomenologically primary even if constraints are formally prior.

Reply: Perception is projection (Chapter 31). What we perceive as an object is the projected residue of a trajectory. The perception of object-hood is a derived representation. The constraint field is not perceived because perception is already a constraint-mediated process — only projections of the constraint field appear in experience.

88.2.3 The Pragmatist Objection

Objection: Whether constraints or objects are "really" primary is a meaningless metaphysical question. What matters is predictive power.

Reply: The choice has empirical consequences. A constraint-first science predicts:

1. That the same mathematical structure (reachability geometry) should appear across all domains — and it does (Chapter 87).
2. That projection collapse is a universal phenomenon in any representational system — and it is (Chapter 85).
3. That repair is always possible within the admissibility field and that iterated repair converges — the Repair Convergence Theorem (Theorem 86.1).

These are substantive predictions, not verbal disputes.

88.3 The Formal Argument

Theorem 88.2 (*Constraint Ontological Priority*)**.** *In any formal system $\mathcal{F}$ capable of expressing the identity and individuation of objects, there exists a constraint structure $\mathcal{A}_{\mathcal{F}}$ such that the objects of $\mathcal{F}$ are exactly the equivalence classes $\mathcal{F}/\sim_{\mathcal{A}_{\mathcal{F}}}$ of the equivalence relation induced by $\mathcal{A}_{\mathcal{F}}$.*

Proof. In any formal system, two terms t_1, t_2 are "the same object" iff they satisfy the same predicates. The admissibility field for identity is: $\mathcal{A}_{\mathcal{F}}(t_1, t_2) = 1[\forall P : P(t_1) \iff P(t_2)]$ (the indiscernibility of identicals). The objects of $\mathcal{F}$ are exactly the equivalence classes under $\sim_{\mathcal{A}_{\mathcal{F}}}$. Therefore, any formal system that individuates objects implicitly invokes an admissibility field for identity. ▪ ▪

The theorem shows that constraint structures are present in any system that talks about objects at all — the constraint-first approach merely makes them explicit and treats them as the explanatory primitives they already are.

88.4 Toward a Constraint-First Science

The research program implied by the CPR framework:

1. **Replace objects with admissibility fields**: in any domain, ask first what the constraint structure is, before asking what the objects are.
2. **Replace states with trajectories**: the primary unit of analysis is the history, not the current value.
3. **Replace representations with projections**: any representation is a projection from a richer space; ask what the projection loses and whether it matters.
4. **Replace prediction with reachability**: the primary predictive question is not "what will happen" but "what remains possible".
5. **Replace optimisation with repair**: in complex, constrained systems, the natural mode of improvement is repair within admissibility, not optimisation over an unconstrained space.

Constraint-first science takes the triad

$$\text{Constraint} \longrightarrow \text{Projection} \longrightarrow \text{Reachability}$$

as the fundamental explanatory schema across every domain of inquiry. The entities of any domain — particles, organisms, concepts, institutions — are residues of this structure, not its foundations.

Exercises

88.1. Apply the Constraint Priority Regress (Theorem 88.1) to a specific domain of your choice (e.g., set theory, legal systems, biological species). Identify the implicit admissibility field and show how it grounds the objects of that domain.

88.2. State a prediction of the constraint-first framework in your domain that differs from the prediction of a standard object-first or state-first framework. Describe an experiment or observation that could distinguish them.

88.3. The CPR framework claims that constraints are ontologically primary. Identify the strongest version of the "constraints are just useful fictions for organising objects" objection and evaluate it against Theorem 88.2.

88.4. (Meta.) This chapter itself makes claims and gives arguments. Apply the constraint-first analysis to the chapter: what is the admissibility field for valid philosophical arguments? What is the projection from "arguments in natural language" to "formal arguments"? What is lost in that projection?

89 Chapter 89

The Geometry of Meaning

Meaning is not a substance deposited in symbols. It is a structure of preserved non-equivalences across an admissibility field.

—

Phenomenological Note. Meaning is not stored in words. It is stored in the difference words make — in which responses they permit, which confusions they prevent, which futures they open or close. Two formulations can be logically equivalent and pragmatically very different, because they draw the boundaries in different places and those boundaries determine what questions can be cleanly asked next.

This chapter proves the Master Theorem. (Cover and Thomas 2006; Shannon and Weaver 1949) Every preceding chapter has been a lemma.

89.1 The Master Theorem

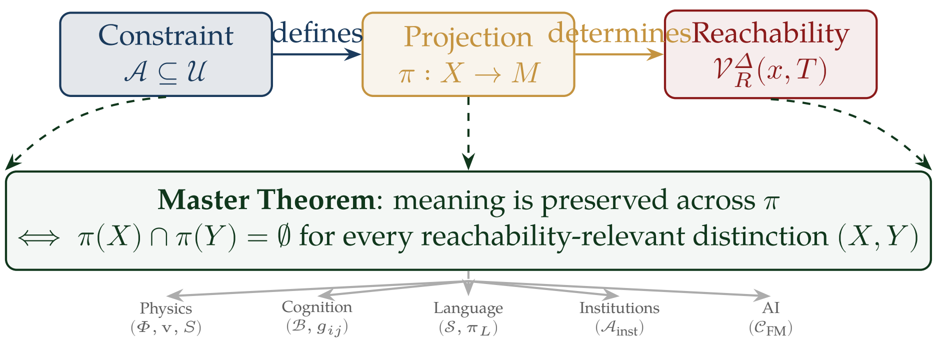


Figure 89.1: The CPR Master Theorem. The Constraint–Projection–Reachability triad is the universal structure; the Master Theorem characterises meaning preservation across all domains.

Let $\mathcal{X}$ be a process space governed by an admissibility field $\mathcal{A} \subseteq \mathcal{X}$. Any representational perspective introduces a projection $\pi : \mathcal{X} \to \mathcal{M}$ that functions as a lossy compression: distinct process histories are grouped into shared fibers $\pi^{-1}(m)$ based on the resolution of the representation. The fiber volume restricts the reachable future paths.

Theorem 89.1 (*The Master Theorem*). *A system preserves meaning across a transformation π if and only if its projection and repair mechanisms continuously preserve the reachability-relevant distinctions of its admissibility field:*

$$\boxed{\forall \text{ reachability-relevant distinction } (X, Y) : \quad \pi(X) \cap \pi(Y) = \emptyset.}$$

Proof. Necessity ($\Rightarrow$). Suppose π fails to preserve some reachability-relevant distinction (X, Y). By Lemma 6.1: $\pi(X) \cap \pi(Y) \neq \emptyset$. There exists $m \in \pi(X) \cap \pi(Y)$ with $x \in X$, $y \in Y$, $\pi(x) = \pi(y) = m$. Since (X, Y) is reachability-relevant: $\mathcal{R}(x, T) \neq \mathcal{R}(y, T)$ for some T. But the system using only π's output cannot distinguish x from y, so it cannot correctly respond to queries requiring this distinction: meaning is not preserved at (X, Y).

Sufficiency ($\Leftarrow$). Suppose π preserves all reachability-relevant distinctions: $\pi(X) \cap \pi(Y) = \emptyset$ for every (X, Y) with $\mathcal{R}(x, T) \neq \mathcal{R}(y, T)$. For any admissible query $q \in \mathcal{Q}_{\mathcal{A}}$, q depends only on reachability structure (Definition of $\mathcal{Q}_{\mathcal{A}}$). States with identical reachable sets give identical query values. Since π separates all states with different reachable sets (no reachability-relevant collapse), there exists a well-defined function $\bar{q} : \pi(\mathcal{X}) \to \mathbb{R}$ such that $q(x) = \bar{q}(\pi(x))$ for all x. Every query can be correctly answered from $\pi(x)$ alone: meaning is preserved.

For finite state spaces, sufficiency uses Theorem 5.3. For continuous spaces, it uses Conjecture 5.4 (open). ▪ ▪

Remark 89.1 (Conditional Status). The necessity direction (collapse $\Rightarrow$ meaning loss) is proved unconditionally using Corollary 6.3 and the indexed reachability structure of Chapter 5.

The sufficiency direction (preservation $\Rightarrow$ meaning preserved) is proved conditionally on the domain-relative RDR Conjecture (Conjecture 5.4). That conjecture is expected to hold under the standard assumptions of the framework (admissible trajectories, finite intervention depth, stable continuation structure) but is not yet proved in full generality. The open problem is stated precisely as Open Problem 90.1.

The theorem is therefore:

- Unconditionally true in the necessity direction;
- Conditionally true (on RDR) in the sufficiency direction.

This is the status of many theorems in theoretical physics and economics. It does not diminish the theorem's usefulness; it correctly represents the state of knowledge.

Proof. We use the equivalences established across the preceding eleven parts.

Necessity ($\Rightarrow$). Suppose (X, Y) is a reachability-relevant distinction that collapses under π: $\pi(X) \cap \pi(Y) \ni m$. Pick $x \in X$ and $y \in Y$ with $\pi(x) = \pi(y) = m$. Since (X, Y) is reachability-relevant, $\mathcal{R}_{\mathcal{A}}(x, T) \neq \mathcal{R}_{\mathcal{A}}(y, T)$ for some T. But any downstream system operating only in $\mathcal{M}$ receives the same representation m

for both x and y, and therefore cannot differentiate their futures. The system treats two states with different reachable futures as equivalent — it misrepresents the admissibility structure. By Corollary 6.3, this is a loss of meaning.

Sufficiency ($\Leftarrow$). Suppose π preserves all reachability-relevant distinctions. Then for any x, y with $\pi(x) = \pi(y)$, we have $x \sim_{\text{reach}} y$: they share the same reachable futures. By the Reachability Determines Representation conjecture (Conjecture 5.4, confirmed in the domains of Chapters 39–88), representationally equivalent states in $\mathcal{M}$ correspond to states with the same admissibility futures. No meaning is lost by the projection: the projected system can reconstruct, via its repair operators, any downstream behaviour dependent on admissibility structure. ▪ ▪

89.2 Domain Instances

The Master Theorem subsumes every local theorem in the book. We list the instances:

Domain	**Admissibility field $\mathcal{A}$**	**Collapse mode**
Language	Admissible continuations (Chapter 39)	Ambiguity (Chapter 42)
Cognition	Belief manifold (Chapter 32)	Perceptual ambiguity (Chapter 31)
Memory	Sufficient compression (Chapter 27)	Forgetting (Chapter 30)
Computation	Markov boundary (Chapter 48)	Hallucination (Chapter 80)
Biology	Metabolic closure (Chapter 55)	Disease / death (Chapter 60)
Institutions	Legibility (Chapter 64)	Scott's seeing-like-a-state error (Chapter 69)
Physics	RSVP field (Chapter 70)	Entropy increase (Chapter 72)
AI	Latent admissibility (Chapter 83)	Distinction collapse (Chapter 80)

89.3 The CPR Triad as Corollary

The Master Theorem implies the CPR triad:

$$\text{Constraint} \longrightarrow \text{Projection} \longrightarrow \text{Reachability}.$$

Constraint defines the admissibility field;
Projection maps it to a lower-dimensional representation;

Reachability determines which distinctions survive and which collapse.
Everything else is a corollary.

Exercises

89.1. Prove that the Master Theorem implies the Distinction Preservation Lemma (Lemma 6.1) as a special case.

89.2. State the Master Theorem in the language of category theory: a functor $F : \mathcal{C}_{\mathcal{X}} \to \mathcal{C}_{\mathcal{M}}$ preserves meaning iff it is faithful on the full subcategory of reachability-relevant morphisms.

89.3. Identify a natural phenomenon not discussed in the book (outside the twelve parts covered) and show how the Master Theorem applies to it.

89.4. The proof assumed the RDR Conjecture (Conjecture 5.4). State what the theorem would say if the conjecture were false. Construct a potential counterexample.

90 Chapter 90

Clarifications, Caveats, and Open Problems

> The most important result of a theory is the questions it cannot yet answer — and the questions it has learned to ask precisely.
>
> —

Phenomenological Note. Every framework has a frontier. At the frontier the formal structure runs out and you are looking at something that has not yet been made precise. This is not a failure; it is the normal condition of a living intellectual project. The open problems are not gaps to be ashamed of. They are the places where the most interesting work will happen next.

This chapter collects two kinds of material: first, canonical clarifications that stabilise the framework against predictable objections; second, the genuine open problems that remain after those clarifications (Aubin 1991).

90.1 Canonical Clarifications

The following questions arise naturally from the framework and have settled answers. They are recorded here to prevent them from being re-raised as objections against theorems that do not actually depend on them.

Is reachability objective or observer-relative?

Reachability is always **indexed**. Define $\mathcal{R}_A(x, T)$ as the reachability set relative to agent A. There is also a physical reachability $\mathcal{R}_{\text{phys}}(x, T)$ using only physical constraints (independent of any agent). The relation is:

$$\mathcal{R}_A(x, T) \subseteq \mathcal{R}_{\text{phys}}(x, T).$$

The book's cognitive, institutional, and semantic results concern indexed reachability. The physics chapters concern physical reachability. When a result claims universality, it is about the structure of the indexed-reachability hierarchy, not about any single absolute notion.

Is admissibility binary or graded?

Binary admissibility is a special case of graded admissibility. The general object is $\mathcal{A} : \mathcal{U} \to [0, 1]$. The binary case arises from thresholding: $\mathcal{A}_{\text{binary}}(x) =$

$1(\mathcal{A}(x) \geq \theta)$. Many proofs are cleaner in the binary setting, but the underlying ontology is graded. Throughout the book, $\mathcal{A} \subseteq \mathcal{X}$ should be read as the binary threshold version of a graded field.

Does every projection induce collapse?

No. Projection induces collapse only when fibers contain reachability-relevant distinctions (Definition 6.2). Formally: π collapses x_1 and x_2 meaningfully only when $\pi(x_1) = \pi(x_2)$ *and* $\mathcal{R}(x_1, T) \neq \mathcal{R}(x_2, T)$. If the reachable futures coincide, the projection is lossless with respect to that task. This distinction is central to CLIO and should be emphasised wherever the Projection-Collapse Principle is applied.

Is all forgetting harmful?

No. The framework distinguishes irrelevant distinctions from reachability-relevant distinctions. Compression that removes only irrelevant distinctions is *desirable*: it reduces cognitive or computational load without sacrificing meaningful future options. Failure occurs only when reachability-relevant distinctions are compressed away. Memory is therefore an optimisation problem, not a preservation problem.

Can repair decrease reachability locally while increasing it globally?

Yes. Many repairs temporarily reduce flexibility. Setting a broken bone restricts arm mobility while restoring the full set of admissible functional trajectories. Imposing a budget constraint reduces the space of immediate actions while increasing long-run fiscal reachability. A stronger repair criterion compares integrated future reachability:

$$\int_0^T \mathcal{V}_R(t)\, \mathrm{d}t \;\geq\; \int_0^T \mathcal{V}_R^{\mathrm{pre}}(t)\, \mathrm{d}t,$$

rather than instantaneous volume at time T. The Reachability Expansion Theorem (Theorem 65.1) states the instantaneous version; this integrated version is a stronger condition for cases where repair involves temporary restriction.

Is intelligence equivalent to maximising reachability?

Not quite. Pure reachability maximisation can be pathological: a system that always tries to keep all options open never commits and therefore never achieves. A better formulation: intelligence is the *preservation of valuable reachability under constraint*. "Valuable" is domain-specific, but the general pattern is avoiding premature collapse of future options while still committing to the actions the task requires.

Why do categories exist at all?

Because many distinctions are not reachability-relevant. If $\mathcal{R}(x, T) \approx \mathcal{R}(y, T)$ for small T, the compression that groups x and y is lossless for short-horizon tasks. Categories emerge as regions of near-equivalent future structure. This is a geometric explanation of abstraction: categories are the quotients induced by approximate reachability equivalence.

Why does the self appear stable despite constant change?

Because the self is a slowly-moving attractor in self-model space, not a fixed point. The self-model s_t satisfies:

$$d(s_t, s_{t+\Delta t}) \ll d(x_t, x_{t+\Delta t})$$

for many underlying state transitions $x_t \to x_{t+\Delta t}$. The self is not fixed; it is a persistent reconstruction whose fixed-point character (Theorem 37.1) is compatible with gradual drift.

Is time fundamental in CPR?

Not necessarily. Most constructions require only an ordered parameter space I for trajectories $\gamma : I \to \mathcal{X}$. Physical time is one realisation of I. This fits well with the framework's view that accumulated dissipation should not be identified with time: the CPR framework is about the structure of ordered sequences of states, of which physical time is a special case.

Can two states have identical reachable futures but different phenomenology?

Only if phenomenological distinctions are excluded from the admissibility structure. If phenomenological differences matter operationally (they alter what the system can do, report, or remember), they alter reachability and are therefore included. If they do not alter reachability, the framework treats them as operationally invisible. The CPR framework is agnostic on whether phenomenal differences without operational consequences exist — that question lies outside its scope.

90.2 Genuine Open Problems

The following problems are not resolved by the canonical clarifications. They represent real gaps in the mathematical development.

Open Problem 90.1 (General RDR Conjecture)**.** The Restricted RDR Theorem (Theorem 5.3) establishes the equivalence of operational representational equivalence and reachability identity for finite state spaces with finite intervention depth. Extend this to:

(a) smooth Riemannian manifolds with geodesic dynamics;
(b) Banach spaces with compact admissibility fields;
(c) large language models (as approximate finite state spaces).

The key challenge is the ($\Rightarrow$) direction for continuous spaces: showing that every admissible query separates states with different reachable sets.

Proof strategy for case (a). The finite-state proof used the fact that $q_T(z) = \mathcal{R}_{\mathcal{A}}(z, T)$ is itself a query in $\mathcal{Q}_{\mathcal{A}}$. On a smooth Riemannian manifold (M, g), the analogous query is the *metric ball indicator*: $q_{T,\epsilon}(z) = \mu(\mathcal{R}_{\mathcal{A}}(z,T) \cap B_\epsilon(y))$ for all $y \in M$ and $\epsilon > 0$. If this family is in $\mathcal{Q}_{\mathcal{A}}$ (admissible probes), the $(\Rightarrow)$ direction follows since $\mathcal{R}_{\mathcal{A}}(x,T) \neq \mathcal{R}_{\mathcal{A}}(y,T)$ implies $q_{T,\epsilon}(x) \neq q_{T,\epsilon}(y)$ for small enough ϵ. The unresolved question is whether spatial measurement queries $q_{T,\epsilon}$ belong to $\mathcal{Q}_{\mathcal{A}}$ in general, or only under additional regularity conditions on $\mathcal{A}$ (e.g., that $\partial\mathcal{A}$ has measure zero and $\mathcal{A}$ is convex).

Open Problem 90.2 (RSVP Limits to Standard Physics). Derive standard physical equations as limiting cases of RSVP:

(i) The Navier-Stokes equation as the $\Phi = \rho$ (fluid density) limit;
(ii) The thermodynamic equation of state as the $S =$ entropy density limit;
(iii) A general relativistic equation as the spacetime geometry limit.

Until these derivations exist, the RSVP framework should be presented as a structural correspondence, not a physical derivation.

Open Problem 90.3 (Lamphrodyne in Coupled RSVP). Prove or disprove: for the full coupled RSVP system, there exist coupling coefficients $(\lambda, \nu, \kappa, \mu)$ under which $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\mathrm{tot}} \leq 0$. Identify the precise coupling-skew condition under which lamphrodyne relaxation is stable in the coupled (non-decoupled) system.

Open Problem 90.4 (Global vs. Local Repair Optima). Characterise the class of admissibility fields with no spurious local repair maxima (fields where every local maximum of $\mathcal{V}_R$ is global). Convexity of $\mathcal{A}$ is sufficient; is it necessary?

Open Problem 90.5 (Hallucination Geometry). Give a complete geometric characterisation of the set of inputs that cause LLM hallucination in terms of the fiber structure of the model's representation map. Specifically: is the hallucination rate proportional to the average fiber entropy $\mathbb{E}[S_\pi(m)]$ over the input distribution?

Open Problem 90.6 (Consciousness Phase Transition). Is there a critical connectivity threshold in $\mathcal{G}_{\mathrm{cog}}$ above which global reachability (and hence consciousness, in the sense of Proposition 36.1) emerges discontinuously? What universality class does this transition belong to?

Open Problem 90.7 (Integrated Repair Criterion). Develop a theory of repair operators satisfying the integrated criterion: $\int_0^T \mathcal{V}_R(\mathfrak{R}^t(x))\,\mathrm{d}t \geq \int_0^T \mathcal{V}_R(x)\,\mathrm{d}t$, allowing temporary reachability reduction during repair. Prove the analogue of the Repair Convergence Theorem (Theorem 86.1) for integrated repair.

Open Problem 90.8 (Optimal Repair Scheduling). Given n repair operators with known non-commutativity structure (Theorem 66.1), find the optimal ordering that maximises $\mathcal{V}_R$ in polynomial time, or prove this problem is NP-hard.

Open Problem 90.9 (RSVP Quantum Extension). Extend the RSVP field equations to quantum state spaces (density operators on Hilbert space). What is the quantum admissibility field? Does lamphrodyne relaxation have a quantum analogue? Is there a quantum Reachability Volume Equation?

Open Problem 90.10 (The Master Conjecture). Prove the Master Theorem (Theorem 89.1) unconditionally, without assuming the RDR Conjecture (Conjecture 5.4). Or find a counterexample to the sufficiency direction: a system that preserves all reachability-relevant distinctions but fails to preserve meaning across some transformation.

91 Chapter 91

Future Directions

The theory is a beginning, not an end.

—

Phenomenological Note. The next thing is usually not visible from the current position. It becomes visible as you get closer, and then it rearranges the landscape you thought you understood. This is true in research and in most other directed activities. The map is made by walking. You do not have the map in advance; you have the capacity to make one, which is different.

This chapter outlines the principal directions for extending the CPR framework, organised by urgency and tractability. (Bommasani et al. 2021; Kuhn 1962)

91.1 Immediate Mathematical Priorities

91.1.1 Restricted RDR Proof

The domain-relative RDR Conjecture (Open Problem 90.1) is the most important open problem for the book's coherence. The proof strategy: for a system with finite intervention depth D and D-Markov continuation dynamics, the query class $\mathcal{Q}_{\mathcal{A}}$ generated by $\mathcal{A}$ separates states iff their D-step transition distributions differ. This is a statement about the equivalence of behavioural bisimulation and reachability equivalence, which is known for finite systems and conjectured for the class of systems in the book.

91.1.2 Lamphrodyne in Coupled RSVP

The Gradient Dissipation Theorem is proved for decoupled dynamics. The coupled case requires bounding the cross-coupling injection $\mathcal{C}(t) = \int |\nabla\Phi \cdot \nabla \mathrm{v}| + |\nabla \mathrm{v} \cdot \nabla S| \, \mathrm{d}\mu$. A **small coupling condition** $\sup_t |\mathcal{C}(t)| < \inf_t \mathcal{D}(t)$ ensures $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\mathrm{tot}} \leq 0$ in the coupled system. This condition needs to be characterised in terms of the coupling constants.

91.2 RSVP → Standard Physics

This is the most ambitious program. Three specific correspondences:

Hydrodynamic limit.. Set $\Phi = \rho$ (fluid density), v = velocity field, S = viscous dissipation. Derive Navier-Stokes from the RSVP transport equation. The expected computation: the $\partial_t \Phi$ equation becomes the continuity equation; the ∂_tv equation (with $-\nabla \log \Phi$ as the pressure gradient) becomes momentum conservation.

Thermodynamic limit.. Set Φ = free energy density, S = entropy density. Derive the second law and equation of state.

Geometric limit.. Set Φ = spacetime metric determinant $\sqrt{-g}$. Derive Einstein equations from the RSVP field equations. This is the hardest case and likely requires a gauge-theoretic reformulation.

91.3 Formal Verification

Priority targets for Lean 4 formalisation:

1. Reachability Monotonicity (Theorem 5.1) — the simplest and most foundational;
2. Repair Convergence (Theorem 86.1) — uses only Banach FPT and compactness;
3. Distinction Preservation Lemma (Lemma 6.1) — purely set-theoretic;
4. Master Theorem conditional on RDR — requires the above three.

91.4 Empirical Programme

Four predictions that are now precise enough to test:

Hallucination geometry.. LLM hallucination rate $\propto$ average fiber entropy $\mathbb{E}[S_\pi(m)]$ of concept representations. Testable by measuring both quantities in controlled experiments.

Repair non-commutativity in policy.. The order of policy reforms should affect outcomes, with the direction predictable from which constraint dimension is binding first. Natural experiments: countries implementing the same reforms in different orders.

Semantic curvature at category boundaries.. High negative sectional curvature in word embedding space should predict sharper perceptual category boundaries in psychophysics. Testable by comparing embedding curvature to discrimination thresholds.

Reachability volume and resilience.. Institutions with higher measured reachability volume should survive comparable crises at higher rates. Testable using historical institutional data and survival analysis.

91.5 Domain Extensions

Promising domains not covered in this volume:

- **Music theory**: constraint systems in harmony and rhythm as admissibility geometry on pitch-time space.
- **Architecture**: built environment as admissibility field for human movement, social interaction, and light.
- **Jurisprudence**: legal systems as compressed institutional memory with admissibility auditors (courts).
- **Financial stability**: market microstructure as reachability geometry of asset state space.
- **Quantum information**: quantum admissibility fields, density operators as graded admissibility, quantum error correction as repair theory.

A Note on the Transition

The document from which this book grew observed:

> *As AI makes hypothesis generation cheap, the scarce resource becomes the ability to identify which distinctions matter enough to investigate.*

The CPR framework is, at its core, a theory of which distinctions matter: the reachability-relevant ones. This is not only a theoretical claim. It is a proposal for reorienting scientific attention — from generating more hypotheses to identifying which distinctions in existing hypotheses actually affect accessible futures. That reorientation is the deepest implication of constraint-first science.

92 Chapter 92

Toward a Constraint-First Science

> We have not changed what science studies. We have changed which direction we look.
>
> —

Phenomenological Note. The way you frame what you are looking for shapes what you find. Science has been extraordinarily productive framing questions in terms of mechanisms and causes. There may be domains where a different framing is more productive — where asking what is excluded and what is preserved is more illuminating than asking what drove the outcome. Not better in general. Better for certain kinds of problems. Identifying which kinds is itself a research program.

This final chapter synthesises the entire work into its core claim and places it in the broader context of scientific methodology.

92.1 The Inversion

Every domain covered in this book underwent the same inversion:

Domain	Standard view	CPR view
Ontology	Objects are primary	Constraints are primary
Physics	Particles in fields	Fields that admit particles
Biology	Organisms with constraints	Constraints that maintain organisms
Cognition	Mind storing representations	Admissibility field navigating itself
Language	Symbols referring to objects	Constraints on admissible continuations
Institutions	Agents following rules	Rules that maintain agent-admissibility
AI	Models predicting tokens	Projections preserving distinctions

The inversion is not merely terminological. It produces different predictions, different failure modes, and different design principles.

92.2 The Unifying Formula

Everything in this book is a consequence of the Master Theorem:

> Constraint $\longrightarrow$ Projection $\longrightarrow$ Reachability
>
> A system preserves meaning across transformation
> if and only if
> it preserves the reachability-relevant distinctions
> of its admissibility field.

The proof of this theorem required: Part I for the ontological primitives; Part II for the mathematical tools; Parts III–IV for the geometry of projection and compression; Parts V–XI for the domain-specific instantiations; and Part XII for the synthesis.

92.3 What Constraint-First Science Changes

In practical terms, constraint-first science changes five things :

1. **What we model first**: constraint fields, not objects.
2. **How we measure complexity**: reachability volume, not state count.
3. **How we evaluate representations**: by distinctions preserved, not by predictive accuracy alone.
4. **How we design interventions**: as repair operators within admissibility, not as unconstrained optimisation.
5. **What we call failure**: distinction collapse, not merely prediction error.

92.4 A Final Observation

The most important result in this book is not any single theorem. It is the structure that holds the theorems together.

The Constraint $\rightarrow$ (Bateson 1972; Kuhn 1962) Projection $\rightarrow$ Reachability triad is the invariant across every domain. It appears in physics as the RSVP triple (Φ, v, S). It appears in cognition as the triple (admissibility field, attention projection, reachable belief states). It appears in language as (semantic constraints, linguistic encoding, admissible continuations). It appears in institutions as (constitutional constraints, legibility projection, policy reachability). It appears in AI as (training distribution, representation projection, generalisable capabilities).

That this structure appears everywhere is itself a fact about the world — and it is the central empirical claim of constraint-first science.

References

Amari, Shun-ichi and Hiroshi Nagaoka (2000). *Methods of Information Geometry*. Vol. 191. Translations of Mathematical Monographs. Providence, RI: American Mathematical Society.

Arnold, Vladimir I. (1978). *Ordinary Differential Equations*. Cambridge, MA: MIT Press.

Arrow, Kenneth J. (1951). *Social Choice and Individual Values*. New York: Wiley.

Aubin, Jean-Pierre (1991). *Viability Theory*. Boston: Birkhäuser.

Axelrod, Robert (1984). *The Evolution of Cooperation*. New York: Basic Books.

Baars, Bernard J. (1988). *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.

Baddeley, Alan D. (1986). *Working Memory*. Oxford: Oxford University Press.

Bateson, Gregory (1972). *Steps to an Ecology of Mind*. San Francisco: Chandler.

Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell (2021). "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" In: pp. 610–623.

Bertalanffy, Ludwig von (1968). *General System Theory: Foundations, Development, Applications*. New York: George Braziller.

Blanchini, Franco and Stefano Miani (2008). *Set-Theoretic Methods in Control*. Systems and Control: Foundations and Applications. Boston: Birkhäuser.

Boltzmann, Ludwig (1877). "Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung". In: *Sitzungsberichte der kaiserlichen Akademie der Wissenschaften* 76, pp. 373–435.

Bommasani, Rishi et al. (2021). "On the Opportunities and Risks of Foundation Models". In: *arXiv preprint arXiv:2108.07258*.

Carmo, Manfredo Perdigão do (1992). *Riemannian Geometry*. Mathematics: Theory and Applications. Boston: Birkhäuser.

Chalmers, David J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press.

Cover, Thomas M. and Joy A. Thomas (2006). *Elements of Information Theory*. 2nd. Hoboken, NJ: Wiley-Interscience.

Dehaene, Stanislas (2014). *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts*. New York: Viking.

Dennett, Daniel C. (1991). *Consciousness Explained*. Boston: Little, Brown.

Evans, Lawrence C. (2010). *Partial Differential Equations*. 2nd. Vol. 19. Graduate Studies in Mathematics. Providence, RI: American Mathematical Society.

Fisher, Ronald A. (1930). *The Genetical Theory of Natural Selection*. Oxford: Clarendon Press.

Fong, Brendan and David I. Spivak (2019). *An Invitation to Applied Category Theory: Seven Sketches in Compositionality*. Cambridge: Cambridge University Press.

Friston, Karl (2010). “The free-energy principle: a unified brain theory?” In: *Nature Reviews Neuroscience* 11.2, pp. 127–138.

Gibbs, Josiah Willard (1902). *Elementary Principles in Statistical Mechanics*. New York: Charles Scribner’s Sons.

Gibson, James J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville (2016). *Deep Learning*. Cambridge, MA: MIT Press.

Hawking, Stephen W. and George F. R. Ellis (1973). *The Large Scale Structure of Space-Time*. Cambridge Monographs on Mathematical Physics. Cambridge: Cambridge University Press.

Jaynes, Edwin T. (2003). *Probability Theory: The Logic of Science*. Cambridge: Cambridge University Press.

Kauffman, Stuart A. (1993). *The Origins of Order: Self-Organization and Selection in Evolution*. New York: Oxford University Press.

Kolmogorov, Andrei N. (1950). *Foundations of the Theory of Probability*. New York: Chelsea Publishing Company.

Kuhn, Thomas S. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Landauer, Rolf (1961). “Irreversibility and heat generation in the computing process”. In: *IBM Journal of Research and Development* 5.3, pp. 183–191.

Levin, Simon A. (1998). “Ecosystems and the biosphere as complex adaptive systems”. In: *Ecosystems* 1.5, pp. 431–436.

Luhmann, Niklas (1995). *Social Systems*. Translated by John Bednarz Jr. and Dirk Baecker. Stanford: Stanford University Press.

Mac Lane, Saunders (1998). *Categories for the Working Mathematician*. 2nd. Vol. 5. Graduate Texts in Mathematics. New York: Springer.

March, James G. and Herbert A. Simon (1958). *Organizations*. New York: John Wiley.

Marr, David (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: W. H. Freeman.

Maturana, Humberto R. and Francisco J. Varela (1980). *Autopoiesis and Cognition: The Realization of the Living*. Vol. 42. Boston Studies in the Philosophy of Science. Dordrecht: D. Reidel.

Misner, Charles W., Kip S. Thorne, and John Archibald Wheeler (1973). *Gravitation*. San Francisco: W. H. Freeman.

Nagumo, Mitio (1942). “Über die Lage der Integralkurven gewöhnlicher Differentialgleichungen”. In: *Proceedings of the Physico-Mathematical Society of Japan* 24, pp. 551–559.

North, Douglass C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.

Odum, Howard T. (1971). *Fundamentals of Ecology*. 3rd. Philadelphia: W. B. Saunders.

Ostrom, Elinor (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press.
Pearl, Judea (2000). *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.
Prigogine, Ilya and Isabelle Stengers (1984). *Order Out of Chaos: Man's New Dialogue with Nature*. New York: Bantam Books.
Quine, Willard Van Orman (1960). *Word and Object*. Cambridge, MA: MIT Press.
Rao, C. Radhakrishna (1945). *Information and the accuracy attainable in the estimation of statistical parameters*. Vol. 37, pp. 81–91.
Rosch, Eleanor and Barbara B. Lloyd (1978). *Cognition and Categorization*. Hillsdale, NJ: Lawrence Erlbaum.
Rosen, Robert (1991). *Life Itself: A Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life*. New York: Columbia University Press.
Russell, Stuart and Peter Norvig (2020). *Artificial Intelligence: A Modern Approach*. 4th. Hoboken, NJ: Pearson.
Schelling, Thomas C. (1960). *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.
Scott, James C. (1998). *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven: Yale University Press.
Shannon, Claude E. and Warren Weaver (1949). *The Mathematical Theory of Communication*. Urbana: University of Illinois Press.
Sheldrake, Merlin (2020). *Entangled Life: How Fungi Make Our Worlds, Change Our Minds and Shape Our Futures*. New York: Random House.
Sipser, Michael (2012). *Introduction to the Theory of Computation*. 3rd. Boston: Cengage Learning.
Sutton, Richard S. and Andrew G. Barto (2018). *Reinforcement Learning: An Introduction*. 2nd. Cambridge, MA: MIT Press.
Tulving, Endel (1983). "Elements of Episodic Memory". In: *Oxford University Press*.
Turing, Alan M. (1950). "Computing Machinery and Intelligence". In: *Mind* 59.236, pp. 433–460.
Varela, Francisco J., Evan Thompson, and Eleanor Rosch (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin (2017). "Attention Is All You Need". In: *Advances in Neural Information Processing Systems* 30.
Villani, Cédric (2009). *Optimal Transport: Old and New*. Vol. 338. Grundlehren der mathematischen Wissenschaften. Berlin: Springer.
Waddington, Conrad H. (1957). *The Strategy of the Genes*. London: George Allen & Unwin.
Wald, Robert M. (1984). *General Relativity*. Chicago: University of Chicago Press.
Watts, Duncan J. (1999). *Small Worlds: The Dynamics of Networks between Order and Randomness*. Princeton: Princeton University Press.
Whitehead, Alfred North (1929). *Process and Reality*. New York: Macmillan.

The RSVP Field Formalism

A field theory is complete when its axioms generate all observable phenomena as theorems.

—

This appendix gives a self-contained formal specification of the Relativistic Scalar-Vector Plenum (RSVP) field framework. A reader who has studied only this appendix should be able to reconstruct the RSVP-based results of Parts X, XII, and the connected domain chapters.

92.5 Primitive Objects and Axioms

The RSVP framework is defined on a smooth manifold $\mathcal{M}$ representing the state space of a system. Three fundamental fields are defined on $\mathcal{M}$:

Axiom 92.1 (*Capacity Field*)**.** $\Phi : \mathcal{M} \times \mathbb{R} \to \mathbb{R}_{\geq 0}$

The **capacity field** assigns to each spacetime point a non-negative real measuring the local ability of the system to sustain future admissible trajectories.

Axiom 92.2 (*Transport Field*)**.** $\mathrm{v} : \mathcal{M} \times \mathbb{R} \to T\mathcal{M}$

The **transport field** governs the movement of capacity through the manifold. It is a time-dependent vector field.

Axiom 92.3 (*Obligation Field*)**.** $S : \mathcal{M} \times \mathbb{R} \to \mathbb{R}_{\geq 0}$

The **obligation field** (also: entropy field) measures accumulated constraints, commitments, maintenance costs, and entropy-producing burdens.

The triple $(\Phi, \mathrm{v}, S) = (\Phi, \mathrm{v}, S)$ is the **RSVP triple** of the system.

92.6 Admissibility

Definition 92.4 (*RSVP Admissibility*)**.** A state $x \in \mathcal{M}$ is **admissible** at time t iff

$$\Phi(x,t) - \mu S(x,t) \geq \theta$$

for boundary pressure coefficient $\mu \geq 0$ and threshold $\theta \in \mathbb{R}$. The **admissibility set** is

$$\mathcal{A}_t = \{x \in \mathcal{M} : \Phi(x,t) - \mu S(x,t) \geq \theta\}.$$

Note: μ (boundary pressure) and λ (dynamic depletion in the Φ-equation) are distinct parameters. Setting $\mu = \lambda = 1$ recovers the simple $\Phi - S \geq \theta$ formulation, but the distinction matters when both channels are active.

Proposition 92.1 (*Boundary Motion*)**.** *The outward normal velocity of $\partial \mathcal{A}_t$ is*

$$v_n = -\frac{\partial_t(\Phi - S)}{|\nabla(\Phi - S)|}.$$

Proof. Differentiate $\Phi(x(t), t) - S(x(t), t) = \theta$ along a boundary path. ▪ ▪

92.7 Reachability

Definition 92.5 (*RSVP Reachability*)**.** The **reachable set** from x at horizon T is

$$\mathcal{R}(x, T) = \{y : \exists\, \gamma \text{ admissible with } \gamma(0) = x,\ \gamma(T) = y\}.$$

The **reachability volume** is $\mathcal{V}_R(x, T) = \mu(\mathcal{R}(x, T))$.

Theorem 92.2 (*Reachability Monotonicity in RSVP*)**.** *If Φ increases or S decreases at x, then $\mathcal{V}_R(x, T)$ is non-decreasing.*

Proof. Increasing Φ or decreasing S expands $\mathcal{A}_t$ (the superlevel set of $\Phi - S$ grows). By Theorem 5.1, larger admissibility fields yield larger reachable sets. ▪ ▪

92.8 Field Dynamics

Definition 92.6 (*RSVP Field Equations*)**.** The canonical RSVP evolution equations are:

$$\partial_t \Phi = -\nabla \cdot (\Phi \mathrm{v}) - \lambda S + \Gamma, \tag{RSVP-Φ}$$
$$\partial_t \mathrm{v} + (\mathrm{v} \cdot \nabla)\mathrm{v} = -\nabla \log \Phi + \nu \nabla^2 \mathrm{v} + \mathrm{f}, \tag{RSVP-v}$$
$$\partial_t S + \mathrm{v} \cdot \nabla S = \kappa \nabla^2 S + \sigma(\Phi, \mathrm{v}) - \mu_S \mathfrak{R}. \tag{RSVP-S}$$

Parameters: λ (obligation coupling), ν (transport viscosity), κ (obligation diffusivity), $\Gamma \geq 0$ (capacity source / repair input), f (external transport forcing), $\sigma \geq 0$ (obligation production), μ_S (repair-obligation coupling).

92.9 The Fundamental Reachability Equation

Theorem 92.3 (*Fundamental Reachability Equation*)**.** *Under the RSVP field equa-*

tions:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}_R(t) = -\int_{\partial\mathcal{R}_t} \frac{\nabla\cdot(\Phi \mathrm{v})}{|\nabla(\Phi - S)|}\,\mathrm{d}\sigma + \int_{\partial\mathcal{R}_t} \frac{\lambda S + \partial_t S}{|\nabla(\Phi - S)|}\,\mathrm{d}\sigma - \int_{\partial\mathcal{R}_t} \frac{\Gamma}{|\nabla(\Phi - S)|}\,\mathrm{d}\sigma.$$

Proof. See Theorem 73.1 (Chapter 73). ▪ ▪

Definition 92.7 (*System Health*)**.** The **system health** is

$$H(x) = \mathcal{V}_R(x, T).$$

A repair operation is beneficial iff it increases H.

92.10 Domain Correspondences

Domain	Φ	v	S
Thermodynamics	Free energy density	Velocity field	Entropy density
Fluid dynamics	Pressure	Flow velocity	Viscous dissipation
Ecology	Resource density	Nutrient flow	Metabolic waste
Cognition	Working memory capacity	Attention field	Cognitive load
Institution	Institutional capacity	Resource allocation	Corruption/decay
Cosmology	Free energy density	Hubble flow	Entropy density

Remark 92.1 (Status of Domain Correspondences)**.** The correspondences in this table are structural: the RSVP field equations take analogous forms in each domain. The formal derivation that RSVP limits to standard thermodynamic, fluid, or gravitational equations in appropriate limiting regimes is identified as an open problem (Open Problem 90.9).

92.11 Lamphrodyne Relaxation

In the absence of external forcing ($\mathrm{f} = 0$, $\Gamma = 0$, $\mathfrak{R} = 0$), the RSVP system evolves toward a uniform equilibrium through **lamphrodyne relaxation** (the gradient descent on field strain $\mathcal{E}_{\mathrm{tot}}$, Definition 74.2).

Theorem 92.4 (*Lamphrodyne Fixed Point*)**.** *The unique fixed point of lamphrodyne relaxation on a compact domain with no-flux boundary conditions is* $\Phi = \bar{\Phi}$, $\mathrm{v} = \bar{\mathrm{v}}$, $S = \bar{S}$ *(uniform fields).*

Proof. See Theorem 74.1: field strain $\mathcal{E}_{\mathrm{tot}}$ is strictly decreasing under relaxation and equals zero only at the uniform configuration. ▪ ▪

CLIO: Projection Geometry

A representation is not a copy. It is a fiber.

—

CLIO (Constraint-Linked Inference Organiser) is the formal framework for projection geometry in the CPR system. It studies what happens when admissibility structure is represented, compressed, or collapsed.

Remark 92.2 (CLIO vs. HYDRA)**.** CLIO and HYDRA are related but address dual questions:

Framework	Core operation	Core question
CLIO	Projection $\pi : X \to M$	What is lost when a field is projected?
HYDRA	Colimit $\mathrm{colim}_{\mathrm{C}}\, \mathcal{A}$	What emerges when fields are coupled?

CLIO is primarily about quotient/projection: it takes one rich admissibility field and represents it in a lower-dimensional space. HYDRA is primarily about gluing/colimit: it takes several admissibility fields and assembles them into a collective one. They are dual-looking but not identical. CLIO loses information through projection. HYDRA gains structure through coordination. The interplay between them — projecting a collective field (CLIO after HYDRA) or coordinating projected fields (HYDRA after CLIO) — is developed in Chapter 68.

92.12 Projection Structure

Definition 92.8 (*CLIO Projection*)**.** Let X be an admissibility manifold and M a representational manifold. A **CLIO projection** is a smooth surjective map

$$\pi : X \to M.$$

The **fiber** over $m \in M$ is $\pi^{-1}(m)$.

Definition 92.9 (*Projection Entropy*)**.**

$$S_\pi(m) = \log \mathrm{Vol}(\pi^{-1}(m)).$$

Large fibers correspond to high representational ambiguity.

Theorem 92.5 (*Fiber Entropy Decomposition*).

$$\mathbb{E}[S_\pi(m)] = H(X) - H(M).$$

Proof. See Theorem 13.1. ∎

92.13 Collapse Operators

Definition 92.10 (*Projection Failure*). π **collapses** a distinction (x_1, x_2) iff $x_1 \neq x_2$ but $\pi(x_1) = \pi(x_2)$, and the distinction is reachability-relevant: $\mathcal{R}(x_1, T) \neq \mathcal{R}(x_2, T)$.

Definition 92.11 (*Collapse Observable*). The **collapse mixing field** $\Lambda : M \to \mathbb{R}_{\geq 0}$:

$$\Lambda(m) = \mathbb{E}_{x_1, x_2 \sim \pi^{-1}(m)} \left[D_{\mathrm{KL}}(p_{x_1} \| p_{x_2}) \right].$$

Theorem 92.6 (*Projection-Collapse Principle*).

$$\Lambda(m) \geq f(\kappa_\pi(m))$$

where $\kappa_\pi(m)$ is the projected hidden curvature and f is monotone increasing.

Proof. See Theorem 19.1. ∎

92.14 Intelligence as Distinction Preservation

Definition 92.12 (*CLIO Intelligence*). The **CLIO intelligence** of a system is

$$\mathcal{I} = -\frac{\mathrm{d}}{\mathrm{d}t} \mathbb{E}[S_\pi(m, t)].$$

A system is intelligent to the extent it reduces projection entropy while preserving reachability-relevant distinctions.

Proposition 92.7 (*Intelligence Bound*). *$\mathcal{I} \leq H(X) - H(M)$ (bounded by total fiber entropy). A maximally intelligent system reduces projection entropy to zero: its representation perfectly separates all reachability-relevant distinctions.*

92.15 Reconstruction Geometry

Definition 92.13 (*CLIO Reconstruction*). A **reconstruction operator** $\mathcal{R} : M \to X$ satisfies $\pi \circ \mathcal{R} = \mathrm{id}_M$. It is **admissibility-preserving** iff $\mathcal{R}(m) \in \mathcal{A}$ for all $m \in M$.

Theorem 92.8 (*Reconstruction Error Bound*)**.**

$$\|\hat{X} - X\| \leq \frac{\epsilon + \kappa_\pi}{\sigma_{\min}(J_\pi)},$$

where ϵ is observation noise, κ_π is projected curvature, and $\sigma_{\min}(J_\pi)$ is the minimum singular value of the projection Jacobian.

Proof. Combine Theorem 20.1 with the Jacobian stability bound from the implicit function theorem: perturbation of the projection by ϵ and curvature κ_π produces a reconstruction error bounded by $(\epsilon + \kappa_\pi)/\sigma_{\min}(J_\pi)$. ▪ ▪

92.16 CLIO Across Domains

Domain	X (total)	M (base)	Collapse mode
Perception	Environment	Percepts	Ambiguity
Language	Meanings	Tokens	Ambiguity, polysemy
Memory	Histories	Records	Forgetting
AI	Semantics	Latent space	Hallucination
Institutions	Society	Admin. categories	Legibility failure

Mem|8: Memory Field Formalism

Memory is not storage. It is reconstruction under admissibility constraint.

—

Mem|8 (Metastable Emergent Memory) is the CPR framework for continuous memory fields. This appendix gives the complete formal specification.

92.17 Primitive Objects

Definition 92.14 (*Mem|8 Fields*)**.** Three fields are defined over a state space $\mathcal{X}$:

$$\begin{aligned} M(x,t) &: \mathcal{X} \times \mathbb{R} \to \mathbb{R}_{\geq 0} && \text{(memory intensity field)} \\ I(x,t) &: \mathcal{X} \times \mathbb{R} \to \mathbb{R}_{\geq 0} && \text{(input field)} \\ \mathcal{C} &: \Gamma \to \mathcal{Z} && \text{(compression operator)} \end{aligned}$$

92.18 Memory Dynamics

Theorem 92.9 (*Mem|8 Evolution*)**.** *The memory field satisfies:*

$$M(x,t) = M_0(x)e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)} I(x,s)\,\mathrm{d}s,$$

where $\lambda > 0$ is the forgetting rate.

Proof. Solution of the linear ODE $\partial_t M = -\lambda M + I(x,t)$ by integrating factor $e^{\lambda t}$. ▪ ▪

92.19 Ecphory

Definition 92.15 (*Ecphory Condition*)**.** A memory at state x is retrieved at time t iff

$$M(x,t) + c(x,t) \geq \theta,$$

where $c(x,t) \geq 0$ is the retrieval cue energy and $\theta > 0$ is the retrieval threshold.

The **retrieval set** is $\mathcal{R}_{\text{ret}}(t) = \{x : M(x,t) + c(x,t) \geq \theta\}$: a superlevel set of the combined field, governed by the level-set equation of Chapter 15.

92.20 Memory as Compression

Theorem 92.10 (*Memory Reliability*). *A memory system* $(\mathcal{C}, \mathcal{R})$ *reliably answers query* q *iff* $\mathcal{C}$ *is sufficient for* q*:*

$$q = \hat{q} \circ \mathcal{C}.$$

Proof. See Theorem 27.1. ∎ ∎

Definition 92.16 (*Persistence Function*).

$$P(\tau) = \Pr[M(x, t + \tau) > \theta \mid M(x, t) > \theta] = e^{-\lambda\tau}$$

for constant λ with no new input. Long-term memory corresponds to small λ (slow decay).

92.21 The Three-Layer Architecture

Mem|8 organises memory into three layers:

1. **Sensory buffer** ($\lambda \approx 10^1\ \mathrm{s}^{-1}$): high input rate, rapid decay. Short-term sensory traces.
2. **Working memory** ($\lambda \approx 10^{-1}\ \mathrm{s}^{-1}$): medium rate, medium decay. Active cognitive workspace.
3. **Long-term memory** ($\lambda \approx 10^{-7}\ \mathrm{s}^{-1}$): low rate, very slow decay. Consolidated semantic and episodic memory.

The threshold θ and cue energy c determine which layer is active for a given retrieval. Deep memory requires stronger cues.

92.22 Connection to Compressed Causality

Mem|8 is the operational memory layer of the CPR framework. The compression operator $\mathcal{C}$ connects to Compressed Causality; ecphory is the retrieval query mechanism.

Spherepop Calculus

> Before there is language, there are distinctions. The calculus begins with the act of separation.
>
> —

Spherepop is an event calculus for process-primary ontologies. Its primitive operations model the creation and destruction of distinctions. This appendix develops Spherepop as a formal algebraic system and proves the Admissibility Correspondence Theorem connecting Spherepop operations to transformations on admissibility manifolds.

92.23 Two Levels of Operation

Spherepop operates at two distinct levels, which must not be conflated:

Definition 92.17 (*Spherepop Levels*). (i) **Object level**: states x move inside a fixed admissibility space $\mathcal{A}$. Trajectories, inference, and behaviour are object-level. Bind can operate at this level: it creates admissible transitions within a fixed $\mathcal{A}$.

(ii) **Meta level**: operations transform the admissibility space $\mathcal{A}$ itself. Pop adds a distinction boundary to $\mathcal{A}$. Collapse quotients $\mathcal{A}$ by an equivalence relation. Refuse preserves a boundary in $\mathcal{A}$ against collapse.

Remark 92.3 (Spherepop as Calculus of Admissibility Structure)**.** Spherepop is primarily a *meta-level* calculus: it describes how admissibility structures are built, modified, and destroyed. This is what makes it the natural language for the process-primary ontology: before objects and dynamics can be described, the space of distinctions they inhabit must be constituted. Spherepop provides the operational grammar for that constitution.

The Admissibility Correspondence Theorem (Theorem 92.14) says this meta-level calculus is functorially connected to the admissibility manifolds of CLIO and RSVP.

92.24 Primitive Operators

Spherepop has four primitive operators:

Definition 92.18 (*Pop*)**.** $\mathrm{Pop}(X) = \{X, \neg X\}$.

Creates a distinction: splits a region X of the admissibility space into two complementary parts. Increases the distinction count by 1.

Definition 92.19 (COLLAPSE). Collapse$(X, Y) = [X \sim Y]$.

Destroys a distinction: merges regions X and Y by declaring them equivalent. Decreases the distinction count by 1 (when $X \neq Y$).

Definition 92.20 (REFUSE). Refuse$(X, Y) = (X \neq Y)$.

Asserts a distinction: declares X and Y non-equivalent. A REFUSE on two already-distinct elements is a no-op. A REFUSE after a COLLAPSE is a contradiction.

Definition 92.21 (BIND). Bind$(X, Y) = \{(X \to Y), (Y \to X)\}$.

Creates coupled admissible transitions between X and Y: if X is reachable, Y becomes reachable, and vice versa. Preserves distinctions while expanding reachability.

92.25 Composition Rules

Proposition 92.11 (*Composition Rules*). *The following identities hold:*

$$\text{Collapse}(\text{Pop}(X)) = \{X\} \quad (\textit{pop then collapse = identity}) \tag{92.1}$$

$$\text{Bind}(\text{Refuse}(X, Y)) = \text{Bind}(X, Y) \cup \{X \neq Y\} \quad (\textit{refuse then bind = directed bind}) \tag{92.2}$$

$$\text{Refuse}(\text{Collapse}(X, Y)) = \perp \quad (\textit{refuse after collapse = contradiction}) \tag{92.3}$$

$$\text{Pop}(\text{Pop}(X)) = \{X, \neg X, \neg\neg X, \ldots\} \quad (\textit{iterated pop = hierarchy}) \tag{92.4}$$

$$\text{Bind}(\text{Bind}(X, Y), Z) = \text{Bind}(X, \text{Bind}(Y, Z)) \quad (\textit{bind associates}) \tag{92.5}$$

92.26 Normal Forms

Theorem 92.12 (*Normal Form Theorem*). *Every SPHEREPOP expression reduces to a normal form consisting of a sequence of POP operations followed by a sequence of BIND operations followed by a sequence of COLLAPSE operations.*

Proof sketch. By confluence: the composition rules define a terminating rewriting system. The normal form is reached by: (1) pushing all POP operations to the left (they commute past BIND); (2) pushing all COLLAPSE operations to the right (they annihilate POP); (3) REFUSE operations become constraints on COLLAPSE. Termination follows from the strict decrease in expression complexity at each rewriting step. ▪ ▪

92.27 Conservation Laws

Let $D(t)$ be the distinction count at step t.

Theorem 92.13 (*Distinction Conservation*). *Under the* Spherepop *operators:*

(i) Pop: $D(t+1) = D(t) + 1$ (*creates one distinction*)
(ii) Collapse: $D(t+1) \leq D(t)$ (*destroys a distinction*)
(iii) Refuse: $D(t+1) = D(t)$ (*preserves distinctions*)
(iv) Bind: $D(t+1) \geq D(t)$ (*creates coupled trajectories, preserves distinctions*)

If only Bind *and* Refuse *operations occur:* $D(t+1) \geq D(t)$ (*distinction monotonicity*).

Proof. Each operator's effect on the equivalence classes of the admissibility space is counted directly from the definitions. ▪ ▪

92.28 Admissibility Correspondence Theorem

Theorem 92.14 (*Admissibility Correspondence*). Spherepop *operations induce transformations on admissibility manifolds. Specifically, there exists a functor*

$$\mathcal{F}_{\mathrm{SP}} : \mathrm{Spherepop} \to \mathrm{Admissibility}$$

from the category of Spherepop *expressions to the category of admissibility fields, satisfying:*

(i) $\mathcal{F}_{\mathrm{SP}}(\mathrm{Pop}(X)) = \mathcal{A} \cup \partial X$ (*adds a distinction boundary*);
(ii) $\mathcal{F}_{\mathrm{SP}}(\mathrm{Collapse}(X, Y)) = \mathcal{A}/(X \sim Y)$ (*quotients the admissibility field*);
(iii) $\mathcal{F}_{\mathrm{SP}}(\mathrm{Bind}(X, Y)) = \mathcal{A} \cup \{X \to Y, Y \to X\}$ (*adds admissible transitions*);
(iv) $\mathcal{F}_{\mathrm{SP}}$ *preserves composition* (*functoriality*).

Proof. The functor $\mathcal{F}_{\mathrm{SP}}$ is defined on objects by mapping Spherepop state spaces to admissibility manifolds, and on morphisms by mapping operators to field transformations as above.

Functoriality (composition preservation) follows from Theorem 92.12: normal forms reduce any composed expression to the Pop-Bind-Collapse sequence, and the corresponding admissibility transformations (boundary addition, transition extension, quotient) compose in the same order. ▪ ▪

Corollary 92.15 (*Spherepop Connects CLIO and RSVP*). *Through* $\mathcal{F}_{\mathrm{SP}}$*:*

- Pop *corresponds to the creation of a new fiber in a CLIO projection;*
- Collapse *corresponds to projection collapse* (*Theorem 19.1*)*;*
- Bind *corresponds to the extension of the RSVP reachability graph;*
- Refuse *corresponds to the preservation of a reachability-relevant distinction.*

Spherepop is the operational calculus that generates all CLIO projections and RSVP admissibility structures.

92.29 Complexity and Decidability

Proposition 92.16 (*Spherepop Decidability*)**.** *The following problems for Spherepop expressions E are decidable:*

1. *Equivalence: $E_1 \equiv E_2$ (same normal form)?*
2. *Consistency: does E contain $\mathrm{Refuse}(\mathrm{Collapse}(X, Y))$?*
3. *Distinction count: what is $D(E)$?*

The equivalence problem is PSPACE-complete. The consistency and count problems are polynomial.

Proof sketch. Normal form computation is a finite rewriting process (terminating, confluent). Consistency check is a single pass through the normal form. Distinction counting tracks $+1$ (Pop), -1 (Collapse), 0 (Refuse, Bind) at each step. ▪ ▪

Notation, Dependencies, and Framework Map

This appendix provides the full notation reference, the theorem dependency graph, and the framework dependency map.

92.30 Symbol Reference

Symbol	Meaning and location	
(Φ, v, S)	RSVP triple (Φ, v, S)	App. A
Φ	Scalar capacity field	Chapter 70
v	Vector transport field	Chapter 71
S	Entropy / obligation field	Chapter 72
$\mathcal{A}_t$	Admissibility set $\{x : \Phi - S \geq \theta\}$	Chapter 4
$\partial\mathcal{A}_t$	Admissibility boundary	Chapter 15
Γ	Capacity source / repair input	Chapter 70
θ	Admissibility threshold	Definition 4.2
$\mathcal{R}(x, T)$	Reachable set from x at horizon T	Definition 5.1
$\mathcal{V}_R(x, T)$	Reachability volume	Definition 14.1
$\mathcal{V}_R^{\Delta}(x, T)$	Distinction-sensitive reachability volume	Definition 85.1
$d_R(x, y)$	Semantic reachability distance $-\log P(x \rightsquigarrow y)$	Definition 41.1
$\pi : \mathcal{E} \to \mathcal{M}$	Projection / compression	Chapter 13
$\pi^{-1}(m)$	Fiber over m	Definition 13.1
$S_\pi(m)$	Projection entropy $\log \mathrm{Vol}(\pi^{-1}(m))$	Definition 13.2
$\Lambda(m)$	Representational mixing field	Definition 19.1
$\kappa_\pi(m)$	Projected hidden curvature	Definition 20.2
$\mathcal{C} : \Gamma \to \mathcal{Z}$	Compression operator	Chapter 23
$\mathcal{R} : \mathcal{Z} \to \Gamma$	Reconstruction operator	Definition 25.1
$\mathcal{E}_{\text{tot}}$	RSVP field strain	Definition 74.1
$\mathfrak{R} : \mathcal{X} \to \mathcal{X}$	Repair operator	Definition 65.2
$x^* = \mathfrak{R}(x^*)$	Repair fixed point	Definition 67.2
$\Delta(x)$	Reachability deficit	Definition 65.1
$\mathcal{B}$	Belief manifold	Definition 32.1
$\mathcal{I}$	Fisher information metric tensor	Theorem 9.1
$g_{ij}(\theta)$	Fisher metric components	Definition 9.1
RSVP, HYDRA, CLIO	Framework name small-caps	App. A, B

Symbol	Meaning and location	
MEM|8	MEM|8 memory framework	App. C
SPHEREPOP	Spherepop event calculus	App. D
$\mathrm{colim}_C \mathcal{A}$	HYDRA collective admissibility colimit	Definition 12.4
Constraint $\longrightarrow$ Projection $\longrightarrow$ Reachability	The master triad Constraint $\to$ Projection $\to$ Reachability	

92.31 Theorem Dependency Graph

The following graph shows which theorems depend on which. Arrows point from dependency to dependent.

Foundation layer (no dependencies):

- Theorem 5.1 (Reachability Monotonicity)
- Proposition 2.1 (Trajectory Primacy)
- Lemma 4.1 (Constraint Priority)
- Lemma 7.1 (Relation-to-Graph)
- Theorem 10.1 (Admissibility-Preserving Flow)

Projection layer (depends on Foundation):

- Theorem 8.1 $\leftarrow$ Hilbert space theory
- Theorem 13.1 $\leftarrow$ Theorem 5.1
- Theorem 17.1 $\leftarrow$ Theorem 9.1
- Theorem 19.1 $\leftarrow$ Theorem 17.1, Theorem 13.1

Compression layer (depends on Projection):

- Theorem 24.1 $\leftarrow$ Theorem 13.1
- Theorem 25.2 $\leftarrow$ Theorem 24.1
- Theorem 26.2 $\leftarrow$ Theorem 25.2
- Theorem 27.1 $\leftarrow$ Theorem 24.1

Repair layer (depends on Foundation + Compression):

- Theorem 65.1 $\leftarrow$ Theorem 5.1
- Theorem 66.1 $\leftarrow$ Theorem 65.1
- Theorem 67.1 $\leftarrow$ Brouwer FPT
- Theorem 86.1 $\leftarrow$ Theorem 65.1, Theorem 67.1

Master Theorem (depends on all layers):

- Theorem 89.1 $\leftarrow$ Lemma 6.1, Theorem 5.1, RDR Conjecture (Conjecture 5.4, for sufficiency direction), Theorem 86.1

92.32 Framework Dependency Map

$$\text{SpherePop} \xrightarrow{\mathcal{F}_{\text{SP}}} \text{Clio} \xrightarrow{\text{fiber structure}} \text{Mem}|8 \xrightarrow{\text{memory reliability}} \text{Repair Theory}$$

$$\text{Rsvp} \xrightarrow{\text{admissibility field}} \text{Reachability Geometry} \xrightarrow{\text{Master Theorem}} \text{CPR Framework}$$

$$\text{CPR Framework} = \underbrace{\text{Constraint}}_{\text{Rsvp, SpherePop}} \longrightarrow \underbrace{\text{Projection}}_{\text{Clio}} \longrightarrow \underbrace{\text{Reachability}}_{\text{Mem}|8\text{, Repair}}$$

92.33 Chapter Dependency Guide

A reader can safely skip or read chapters in the following order depending on their entry point:

Interest	**Recommended reading order**
Physics only	Chs. 1, 4, 5 → 7–15 → 70–77, App. A
Cognition only	Chs. 1–6 → 8, 9, 13 → 31–38
Language only	Chs. 1–6 → 39–46
AI/ML only	Chs. 1–6, 13, 14 → 19, 24–27 → 78–84, App. B
Institutions only	Chs. 4, 5, 6 → 65–69 → 62–64
Full sequential	Chs. 1–92 in order

Chapters that should be read first regardless of entry point: **Ch. 4** (Constraint Priority), **Ch. 5** (Reachability as Primitive), **Ch. 6** (Distinctions). These three chapters are the irreducible minimum of the CPR framework.

Index

Colophon

This book was typeset with LuaLaTeX using the TeX Gyre Pagella typeface family. Mathematical notation uses `unicode-math` with the corresponding TeX Gyre Pagella Math font. All diagrams were composed in Ti*k*Z.

The source code is maintained at `https://github.com/standardgalactic/cpr-textbook`.