The Geometry of Control

Admissibility, Reachability, and Projection Collapse
in *Prisoners of Power*

Flyxion

Independent Researcher

June 2026

Abstract

Arkady and Boris Strugatsky's *Prisoners of Power* is commonly interpreted as a political novel concerned with totalitarianism, propaganda, and the abuse of centralized authority. While such readings capture important dimensions of the text, they leave unexplained the novel's persistent emphasis on uncertainty, hidden structure, failed interventions, and the repeated collapse of seemingly adequate explanations. This essay argues that the novel is more fruitfully understood as an investigation into the geometry of knowledge and control.

Viewed through the lenses of admissibility, reachability, and projection failure, the narrative becomes a sequence of increasingly severe epistemic crises. Maxim Kammerer repeatedly encounters situations in which locally coherent models prove globally inadequate. Every explanatory framework he constructs eventually collapses under the weight of distinctions it fails to preserve. The resulting narrative is not merely political but geometric: it concerns the relationship between representations and the futures they permit agents to reach.

At the same time, the novel resists complete assimilation into any formal framework. Its treatment of loyalty, sacrifice, historical contingency, and intervention under radical uncertainty exposes limitations in current theories of projection and repair. The novel therefore functions not as an illustration of a theory but as a stress-test for one. The goal of this essay is not to show that *Prisoners of Power* confirms a preexisting framework, but to investigate what becomes visible when the novel and the framework are allowed to challenge one another.

Contents

Part I: The Novel as an Epistemic Machine

1. Introduction: Literature as Theory Discovery

The relationship between literature and theory is often conceived in one direction. A theoretical framework is developed independently and subsequently applied to a literary work. The text becomes an example, a case study, or a convenient source of illustrations. Such readings may be useful, but they frequently reduce literature to a passive object of analysis.

There exists another possibility. Certain literary works operate not as examples of theories but as instruments for discovering phenomena that existing theories struggle to describe. In these cases, the text functions less like a specimen and more like an experimental apparatus. Narrative becomes a means of exploring structures that remain difficult to formalize directly.

Science fiction has historically occupied this role. Long before formal work on cybernetics, information theory, artificial intelligence, or complex systems, science fiction authors were already constructing environments in which such phenomena could be observed in exaggerated form. The value of these works does not arise from predictive accuracy. Rather, it arises from their ability to isolate structural relationships that ordinary experience tends to obscure.

The works of Arkady and Boris Strugatsky belong to this tradition. Although frequently discussed as political allegories, their novels consistently display a deeper concern with the relationship between knowledge and action. Their protagonists rarely confront simple moral dilemmas. Instead, they encounter worlds whose structure exceeds the explanatory resources available to them. The resulting conflicts emerge not primarily from malice or ideology but from incomplete understanding.

Prisoners of Power represents perhaps the clearest example of this tendency. The novel follows Maxim Kammerer, a young explorer from the highly developed civilization of the Noon Universe, after his accidental arrival on the planet Saraksh. At first glance, the narrative appears to present a familiar confrontation between freedom and totalitarian control. Yet this interpretation quickly proves inadequate. Every major revelation in the novel destabilizes the explanatory framework that preceded it. The regime is not what it initially appears to be. The resistance is not what it initially appears to be. Even liberation itself becomes difficult to evaluate.

This recurring pattern suggests that the central concern of the novel is not political authority but epistemic authority. The fundamental question is not who

should govern. The fundamental question is how agents construct models of worlds whose most important constraints remain hidden from view.

Critical interpretation has not always recognized this. The history of the novel's reception displays a tendency that the novel itself repeatedly dramatizes at the level of narrative. Faced with a text that resists familiar categories, critics have reached for the nearest available framework. *Prisoners of Power* becomes an example of dystopian fiction, Soviet allegory, Cold War anxiety, or anti-totalitarian critique. It is compared to Orwell, to Lem, to Le Guin, to the tradition of outsider intervention narratives. These comparisons are not incorrect. Yet they share a common structure: they stabilize the text's uncertainty by assigning it to a known category.

This is precisely what Maxim does on Saraksh.

Faced with an environment whose structure exceeds his available frameworks, he reaches for the nearest category. The planet becomes an adventure, then a dystopia, then a propaganda system, then a liberation project. Each category captures something real. Each conceals something essential. The critical reception of the novel reproduces, at the level of interpretation, the epistemic condition the novel is investigating at the level of plot.

The present essay attempts a different approach.

Rather than locating the novel within an existing conceptual landscape, it treats the novel as a source of theoretical pressure. The question is not which existing framework best accommodates the text. The question is what the text reveals that existing frameworks have not yet reached.

The concepts brought to bear in what follows—projection, admissibility, reachability, repair—are not applied to the novel from outside. They emerge as responses to problems the novel has already made visible. The theory arrives after the phenomenon, not before it.

This ordering is methodologically significant.

When theory arrives before phenomenon, the phenomenon becomes evidence for the theory. Confirming instances are foregrounded. Resistant instances are explained away. The text is interpreted until it fits.

When theory arrives after phenomenon, the relationship inverts. The phenomenon becomes a constraint on the theory. The text is allowed to resist. The places where the formal vocabulary fails to capture what the narrative sees are not embarrassments to be minimized. They are the most valuable results the analysis can produce.

The argument proceeds in six parts. The first three examine the novel on its own terms, attending to the specific texture of its epistemic structure before any formal apparatus is introduced. The middle sections develop the formal

vocabulary and bring it into contact with the novel's central conflicts. The final sections examine the limits of the framework and the theoretical questions the novel generates that the framework cannot yet answer.

The goal, throughout, is not to reduce the novel to mathematics. The goal is to place two modes of inquiry into conversation, and to take seriously the possibility that the conversation will change both.

2. The Experience of Epistemic Collapse

Most political novels derive their force from revelation. The protagonist discovers a hidden truth concealed beneath an official narrative. Once this truth becomes visible, the moral and political structure of the story becomes increasingly clear. The reader moves from ignorance toward understanding.

Prisoners of Power operates differently.

Rather than moving from confusion to clarity, the novel repeatedly moves from local clarity to global uncertainty. Every revelation generates a larger mystery. Every explanation succeeds only temporarily before being displaced by another. The reader does not experience the accumulation of knowledge so much as the progressive destabilization of certainty.

This structure is visible from the opening chapters. Maxim arrives on Saraksh carrying assumptions inherited from the civilization of the Noon Universe. He expects contact, cooperation, and rational dialogue. His understanding of social development encourages him to interpret unfamiliar institutions as incomplete versions of familiar ones. The result is not simple misunderstanding but systematic misprojection. He continuously interprets local phenomena through categories whose applicability has not yet been established.

The reader initially shares these assumptions. As a consequence, each collapse of Maxim's model simultaneously becomes a collapse of the reader's model. The narrative repeatedly reproduces the experience of discovering that an apparently sufficient description has omitted distinctions that later prove indispensable.

In this sense, the novel functions as a machine for generating epistemic humility.

The structure is worth examining in its details, because the pattern of collapse is not random. Each framework Maxim adopts is more sophisticated than the one that preceded it. Each captures more of the available evidence. Each lasts longer before failing. Yet each failure is more consequential than the last, because each collapse occurs after Maxim has already acted on the basis of the framework that is about to give way.

This asymmetry between the increasing sophistication of Maxim's models and

the increasing severity of their failures is the novel's primary structural feature. It is not a coincidence of plot. It reflects a deep property of the kind of system Saraksh represents: a system in which increasing local accuracy of representation does not guarantee increasing global adequacy, because the most consequential structures of the system are precisely those that every available framework is least equipped to preserve.

The reader's experience of this structure is itself theoretically significant.

One does not merely observe Maxim's failures from outside. One inhabits them. The novel is constructed so that the reader's explanatory resources are always approximately one step behind the situation. The explanation that seems adequate always turns out to have omitted something. The framework that seemed to account for the evidence always turns out to have collapsed a distinction that the next chapter will reveal to be essential.

This is not a flaw in the novel's construction. It is the novel's primary theoretical contribution at the level of form.

The reader is given the experience, not merely the description, of operating within a projection that proves inadequate. The theoretical content is delivered phenomenologically before it is delivered propositionally. By the time the formal analysis of the following chapters arrives, the reader has already inhabited the condition that analysis is attempting to describe.

3. The Architecture of Concealment

Before the formal apparatus of projection and admissibility is introduced, it is necessary to attend carefully to the specific architecture of concealment that the novel constructs.

The world of Saraksh is organized as a hierarchy of hidden constraints, each layer generating the phenomena that the layer above it misinterprets as fundamental.

The surface layer is the one Maxim encounters first: a totalitarian state, perpetually at war, with omnipresent police, compulsory ideology, and a population living under conditions of privation and surveillance. This layer is real. The suffering is genuine. The coercion is genuine. The institutional apparatus of authoritarian governance is present and operative.

But it is not the deepest layer.

Beneath the political surface lies the broadcast system. The towers scattered across the landscape are not, as their official designation suggests, anti-ballistic missile defenses. They are the primary mechanism through which the cognitive space of the population is managed. The political institutions that occupy the

surface are in part downstream effects of this deeper mechanism. Understanding the surface without understanding the broadcasts is like understanding a shadow without understanding the object that casts it.

Beneath the broadcast system lies the distinction between the immune and the susceptible. The Unknown Fathers, the degens, and eventually Maxim himself inhabit a different cognitive space than the susceptible majority. This distinction organizes the entire political landscape in ways invisible to anyone who does not know it exists. The resistance movement, the structure of governance, the categories of insider and outsider: all are shaped by a variable that most inhabitants of the system cannot observe.

Beneath the immune-susceptible distinction lies the geopolitical situation that the planet's own information environment cannot represent. The Island Empire, the broader context of Saraksh's history, the presence of Progressors from the Noon Universe: these are structures operating at a scale that no native actor's framework is designed to accommodate.

And beneath all of these lies the physical fact with which this analysis must begin: the atmospheric conditions that produce the spherical world illusion, the most fundamental projection failure in the novel, the one that precedes every other distortion and provides the structural template for all that follows.

4. The First Projection: Living Inside the Sphere

Before the reader encounters the Towers, the Unknown Fathers, the mutants, or the hidden mechanisms of political control, the novel introduces a more fundamental distortion. The inhabitants of Saraksh believe that they live on the interior surface of a sphere.

This belief is not merely a cultural superstition. It emerges from the physical conditions of the planet itself. Atmospheric refraction prevents direct observation of the horizon and produces the appearance of a closed world. The illusion is sufficiently compelling that it becomes embedded within the civilization's ordinary understanding of reality. What is remarkable is not that the inhabitants are mistaken. What is remarkable is that the mistake is rational.

The people of Saraksh are not irrational observers ignoring obvious evidence. On the contrary, they are responding appropriately to the information available to them. Their conclusions emerge from observation, inference, and experience. The error arises not from defective reasoning but from the structure of the observations themselves.

This distinction is crucial.

Many accounts of knowledge assume that error originates primarily from

failures of logic. An individual reaches an incorrect conclusion because they reason badly. The world itself remains transparent. The responsibility for failure lies entirely with the observer.

The situation on Saraksh is different. The inhabitants inhabit a world whose structure actively distorts observation. Their reasoning operates upon projections already shaped by physical constraints. Under such circumstances, intelligence alone cannot guarantee correctness. Better reasoning applied to the same distorted evidence may simply produce more sophisticated versions of the same mistake.

The atmospheric illusion therefore functions as the novel's first and most important lesson. Before political control appears, before propaganda appears, before ideological conflict appears, the reader is confronted with a deeper possibility: reality may be systematically misrepresented even in the absence of deception.

The significance of this point extends far beyond the geography of Saraksh. The spherical-world illusion provides a miniature version of the epistemic condition that governs the entire narrative. Again and again, characters operate within environments where observations are genuine yet incomplete, where conclusions are locally justified yet globally incorrect, and where hidden structures shape visible phenomena without becoming directly visible themselves.

The planet itself therefore becomes a metaphor for representation.

The inhabitants do not see the world directly. They see a projection generated by interactions between the world and the medium through which it is observed. The resulting model successfully supports many practical activities. People navigate, build cities, conduct wars, and organize societies while possessing a fundamentally incorrect understanding of the larger structure in which those activities occur.

This possibility should not be dismissed as merely fictional. The history of science contains numerous examples of comparable situations. Ancient astronomy successfully predicted celestial motions while operating within an incorrect cosmological framework. Classical mechanics generated extraordinarily accurate predictions despite lacking knowledge of relativistic effects. Early theories of heredity enabled practical breeding long before the molecular basis of genetics became understood. In each case, useful local competence coexisted with incomplete global understanding.

The lesson is not that knowledge is impossible. The lesson is that successful action does not guarantee adequate representation.

The distinction between competence and understanding becomes one of the novel's recurring themes. Systems may function while resting upon mistaken assumptions. Societies may remain stable while operating under distorted descriptions of reality. Individuals may act effectively while misunderstanding the

mechanisms that make their effectiveness possible.

The atmospheric illusion of Saraksh represents the first instance of this pattern.

It also establishes a structural irony that persists throughout the novel. The reader encounters a civilization whose entire worldview is organized around an unnoticed projection. Yet the reader simultaneously assumes that this revelation places them in a privileged epistemic position. Having recognized the inhabitants' mistake, one is tempted to believe that the deeper structure of the world has now become visible.

The remainder of the novel systematically dismantles this confidence.

Each subsequent revelation reproduces the same pattern at a different scale. The official ideology is exposed as a projection. The resistance's interpretation is exposed as a projection. Maxim's revolutionary ambitions become a projection. Even the apparently straightforward opposition between freedom and control eventually proves inadequate.

The sphere therefore serves not merely as background worldbuilding but as a structural template for the entire narrative. Every major conflict in the novel can be understood as a dispute occurring within a representational framework that later turns out to be incomplete.

What changes from episode to episode is not the existence of projection but the location of the hidden constraint.

From the perspective of admissibility, the significance of the spherical illusion becomes even clearer. The inhabitants possess a representation that preserves many distinctions while destroying others. Their model supports local navigation but fails to preserve the global geometry of their environment. The representation is therefore neither wholly wrong nor wholly correct. It occupies an intermediate position in which certain forms of action remain possible while others become systematically distorted.

This observation suggests a more general principle. Representations should not be evaluated solely in terms of truth or falsity. They should also be evaluated according to the futures they preserve or destroy.

A representation may contain inaccuracies while remaining highly useful. Conversely, a representation may contain many correct facts while obscuring distinctions essential for navigation. The practical significance of a model depends not only upon correspondence with reality but also upon its capacity to preserve reachability.

The civilization of Saraksh survives despite its mistaken cosmology because the distinctions required for everyday activity remain largely intact. The illusion becomes dangerous only when actions require access to structures that the projection has erased.

This relationship between projection and reachability forms the conceptual foundation of the novel. Long before the Towers begin broadcasting, long before Maxim becomes involved in revolutionary politics, the reader has already been introduced to a world in which representation and reality have diverged.

The remainder of the narrative can therefore be read as an exploration of a single question: what happens when an entire civilization attempts to navigate using maps that preserve only part of the territory?

5. Maxim Kammerer and the Confidence of Good Maps

The tragedy of Maxim Kammerer does not arise from ignorance.

Indeed, Maxim is among the least ignorant individuals on Saraksh.

He arrives from a civilization possessing scientific, technological, and social capabilities far beyond those available on the planet. He possesses knowledge inaccessible to nearly everyone he encounters. He is educated, intellectually curious, courageous, and remarkably resistant to ideological manipulation. By conventional standards, he is exactly the kind of individual one would expect to understand the situation most clearly.

Yet the novel repeatedly demonstrates that Maxim's advantages become sources of vulnerability.

The reason is subtle. Maxim does not merely possess knowledge. He possesses confidence in the relationship between knowledge and understanding.

He assumes that increasing information leads naturally toward better models. He assumes that hidden structures, once discovered, can be incorporated into a progressively more accurate description of reality. Most importantly, he assumes that truth and successful intervention tend to align.

These assumptions are reasonable.

They are also precisely the assumptions the novel places under pressure.

Throughout the narrative, Maxim repeatedly discovers facts that overturn previous beliefs. Each discovery improves his understanding of some aspect of the situation. Yet the accumulation of these discoveries does not produce convergence toward certainty. Instead it reveals additional layers of complexity. Every solved mystery exposes a larger unsolved one.

The resulting pattern differs significantly from traditional heroic narratives. The hero's journey is often structured around increasing competence. Obstacles are overcome through learning, adaptation, and perseverance. Knowledge accumulates and eventually enables decisive action.

Maxim's trajectory follows the opposite direction.

Knowledge accumulates. Confidence deteriorates. The world becomes less legible rather than more.

This distinction is essential for understanding the novel's deeper structure. Maxim is not primarily a revolutionary, a soldier, or an explorer. He is an interpreter. His central activity throughout the narrative consists of constructing explanatory models and revising them when they fail.

The true drama of the novel therefore unfolds not on battlefields or within political institutions but within the space of representation itself.

The question is never merely what Saraksh is.

The question is how one determines what Saraksh is when every available description appears incomplete.

5.1. Successive Maps of Saraksh

The structure of *Prisoners of Power* can be understood as a succession of maps. Each map explains more than the one that preceded it. Each map is also eventually revealed to be inadequate.

The significance of this pattern lies in the fact that the maps are not simply false. Every interpretive framework Maxim adopts captures genuine features of the world. The problem is never complete error. The problem is partial adequacy.

This distinction is important because it shifts attention away from truth and falsity toward preservation and omission. The central question becomes not whether a model is correct but whether it preserves the distinctions upon which future understanding depends.

Maxim's first map emerges immediately after his arrival. He interprets Saraksh through the lens of exploration. The planet appears as an opportunity for discovery. Its inhabitants seem likely to become partners in communication and cultural exchange. The situation resembles the adventure narratives familiar from earlier eras of exploration. Maxim imagines himself as a modern Robinson Crusoe temporarily isolated from civilization but ultimately destined to establish productive contact with a new world.

This interpretation is not irrational. The inhabitants are human. Communication proves possible. Friendships do emerge. Yet the map omits the institutional and historical structures that organize the society Maxim has entered. The resulting failure is not a consequence of mistaken observations. It is a consequence of premature framing. Before enough distinctions have been gathered, the entire situation has already been assigned to a category.

The first map collapses.

A second map replaces it. Having encountered forced labor camps, police surveillance, ideological rituals, and omnipresent militarization, Maxim begins to

interpret Saraksh as a totalitarian state. The explanatory structure now appears considerably stronger. Previously inexplicable events acquire coherence. The suffering of ordinary citizens becomes intelligible. Political authority emerges as the central organizing principle of social life.

Once again, the interpretation succeeds. The regime is genuinely authoritarian. Violence is real. Oppression is real. Yet the map remains incomplete. The society's visible institutions are not the primary mechanism through which power operates. The categories of conventional political analysis explain certain features of the system while obscuring others. Maxim has identified an important layer of reality but mistakes it for the deepest layer.

The second map collapses.

A third map emerges. The discovery of the broadcasts appears to provide the missing explanation. Suddenly the irrational enthusiasm of the population, the apparent stability of the regime, and the peculiar status of the degens become comprehensible. The hidden mechanism has been revealed.

This stage occupies a privileged position within the narrative because it resembles the kind of revelation toward which many political novels build. The hidden machinery becomes visible. The protagonist discovers the truth concealed behind appearances. The reader expects the trajectory toward liberation to begin.

The novel refuses this structure.

Instead, the discovery of the broadcasts generates a new simplification. The world becomes divided into manipulators and manipulated, oppressors and victims, controllers and controlled. The explanatory power of this model is considerable. Yet it also encourages a new reduction: the entire complexity of the civilization threatens to become subordinated to a single mechanism.

Here the novel introduces one of its most important insights. The identification of a hidden mechanism does not guarantee the identification of an admissible intervention. Knowledge of cause does not imply knowledge of consequence. The discovery of a constraint does not reveal the geometry of its removal.

The distinction becomes visible through Maxim's encounters with actors who refuse to share his certainty. The mutants, in particular, repeatedly evaluate situations according to criteria different from his own. Their concern is not whether the system is unjust. They already know that it is. Their concern is what futures become reachable after the proposed repair.

From Maxim's perspective, these responses often appear passive or overly cautious. From another perspective, however, they reveal a form of reasoning largely absent from his own approach. Whereas Maxim evaluates the present state of the system, the mutants evaluate the trajectories emerging from interventions upon it.

The difference is subtle but profound.

One asks what is wrong. The other asks what follows.

As the narrative progresses, Maxim gradually acquires additional information regarding the broader geopolitical situation. The Island Empire, initially imagined as a potential source of assistance, proves capable of atrocities comparable to those committed by the regime he opposes. The opposition between freedom and tyranny begins to fracture. Every candidate solution reveals its own anomalies. Every apparent escape route terminates in another constraint.

The result is a progressive erosion of explanatory certainty. What initially appeared to be a contest between truth and falsehood becomes a contest among incomplete descriptions. What appeared to be a struggle between freedom and oppression becomes a conflict among incompatible futures.

At each stage Maxim's understanding increases. At each stage the world becomes more difficult to navigate.

This recurring phenomenon suggests a more precise characterization of Maxim's failure.

His error is not overconfidence in particular conclusions. His error is confidence in the adequacy of the representational frame from which those conclusions are drawn.

Every map is treated as though it contains the distinctions necessary to understand the next decision. Every map eventually proves to have discarded distinctions that later become indispensable.

The resulting pattern may be termed projection confidence: not confidence that one's beliefs are true, but confidence that the categories through which beliefs are organized are sufficient.

The tragedy of Maxim Kammerer is therefore not that he repeatedly reaches incorrect conclusions. The tragedy is that he repeatedly discovers the incompleteness of the spaces within which those conclusions were formed.

The novel's deepest lesson may be that intelligence alone does not guarantee escape from this condition. One may continually revise beliefs while remaining trapped within representational structures whose limitations become visible only after action has already begun.

In this sense, Maxim's journey is not a movement from ignorance to knowledge.

It is a movement from confidence in maps to awareness of mapping itself.

Part II: Projection Failure

6. Admissibility, Projection, and Reachability

The preceding chapters have developed a problem through narrative means.

A civilization rests upon a distorted representation of its physical environment. A political system maintains stability through systematic deformation of cognitive space. An agent with superior knowledge repeatedly constructs models that prove locally adequate and globally incomplete. A liberation event produces consequences indistinguishable, in their immediate effects, from catastrophe.

These observations share a common structure.

In each case, a representation that appears sufficient proves to omit distinctions upon which future navigation depends. In each case, the failure is not located in reasoning but in the representational framework within which reasoning occurs. And in each case, the question of whether a given representation is adequate cannot be answered by examining the representation alone. It requires examining what the representation preserves and what it destroys.

This section introduces a formal vocabulary for these phenomena. The concepts are not imported from outside the problem. They are responses to the problem that the novel has already forced into view.

6.1. State Descriptions and Projection Operators

Let Ω denote the space of possible states of a system. For complex social systems, Ω is typically high-dimensional, encompassing institutional configurations, population belief states, material conditions, historical trajectories, and latent causal structures invisible to any particular observer.

No agent has direct access to Ω.

Every agent constructs a representation by selecting dimensions deemed relevant and discarding others. This operation can be formalized as a projection operator $\pi : \Omega \to \Sigma$, where Σ is a lower-dimensional representation space.

The projection π is not a distortion introduced by error. It is a necessary feature of finite cognition. No agent can maintain a complete description of Ω. Every act of understanding involves selecting which distinctions to preserve.

The critical question is therefore not whether projection occurs. It always occurs. The question is which distinctions survive the projection and which are discarded.

6.2. Admissibility

Definition 6.1. A projection $\pi : \Omega \to \Sigma$ is *admissible* with respect to a decision class $\mathcal{D}$ if it preserves the distinctions necessary for decisions in $\mathcal{D}$ to be executed without systematic distortion.

More precisely, let $d \in \mathcal{D}$ be a decision whose correct execution depends on distinguishing between states $\omega_1, \omega_2 \in \Omega$. The projection π is admissible for d if $\pi(\omega_1) \neq \pi(\omega_2)$.

If $\pi(\omega_1) = \pi(\omega_2)$ while ω_1 and ω_2 require different responses under d, then π is inadmissible for that decision. The agent will behave identically in situations that require different behavior. Failure is geometric rather than logical: the map has collapsed a distinction the territory requires.

Admissibility is therefore always relative to a decision class. A projection may be highly admissible for one class of decisions while being deeply inadmissible for another.

The civilization of Saraksh provides a clear illustration. The spherical-world model is admissible for a wide range of everyday navigational and social decisions. It becomes inadmissible precisely when decisions require knowledge of the planet's true geometry.

6.3. Reachability

Definition 6.2. Let $R(\omega, t)$ denote the set of states accessible from ω within time horizon t under the available repertoire of actions. This set constitutes the *reachability volume* of an agent at state ω.

A representation π preserves reachability if the actions available within Σ can be mapped back to actions in Ω without systematic reduction of $R(\omega, t)$.

A representation fails to preserve reachability when it collapses distinctions between states that require different actions to reach the same target. Under such collapse, the agent believes a target is reachable when it is not, or believes it is unreachable when it is, because the distinctions necessary to identify the correct action sequence have been discarded.

Reachability failure of this kind is pervasive in *Prisoners of Power*. Maxim repeatedly believes that certain futures are reachable from his current position and acts accordingly. The mutants' refusals, the Wanderer's interventions, and the catastrophic aftermath of the Control Center's destruction all represent situations in which the reachability volume Maxim believed he inhabited was larger than the reachability volume that actually existed.

6.4. Projection Collapse

Definition 6.3. *Projection collapse* designates the failure mode in which a representation that successfully supports one class of decisions is extended to a second class for which it was not designed, with the result that distinctions critical to the second class are systematically discarded.

This is the precise structure of Maxim's recurring error. His exploration frame was constructed for initial contact decisions. Extended to political analysis, it collapses distinctions between institutional types. His political analysis frame was constructed for understanding surface power relations. Extended to intervention design, it collapses distinctions between causal mechanisms. His broadcast-control frame was constructed for understanding the mechanism of population management. Extended to civilizational intervention, it collapses distinctions between constraint-removal and stabilization.

In each case, the agent is confident in the adequacy of a frame constructed for a narrower purpose than the one to which it is now being applied. Projection confidence is therefore a systematic phenomenon rather than an individual psychological tendency. It arises whenever successful local application of a representation generates confidence that the same representation is adequate for decisions of wider scope.

6.5. The Broadcast System as Admissibility Architecture

The Towers of Saraksh can now be characterized more precisely than the political analysis of earlier chapters permitted.

The broadcast system is not a content delivery mechanism. It is an admissibility architecture.

It operates by constraining the projection operators available to the population. Under continuous low-intensity broadcast, the representations through which most citizens interpret their social environment cannot preserve certain distinctions. Specifically, they cannot preserve distinctions between the propaganda model of social reality and alternative models that would be constructed by an agent with access to the suppressed cognitive space.

The twice-daily intense broadcast performs a complementary function. It does not install beliefs. It resets affective states in ways that prevent the accumulation of the dissonance that would otherwise motivate revision of the available representational frameworks.

The system therefore operates on the meta-level of cognition rather than the object level. It does not determine what citizens conclude. It determines the space of frameworks within which conclusions can be formed.

6.6. Epistemic Admissibility and Civilizational Admissibility

The analysis now permits a precise formulation of the novel's central tension.

Maxim optimizes for epistemic admissibility. He seeks representations that preserve distinctions between true and false descriptions of reality. His consistent objective is transparency.

The Wanderer optimizes for what might be termed civilizational admissibility. He seeks to preserve the reachability volume of Saraksh as a civilization across long time horizons. His consistent objective is navigability.

These objectives are not the same.

A representation may be epistemically inadmissible while preserving civilizational reachability. The broadcasts produce systematically distorted models of reality among the susceptible population. Yet the civilization continues to function.

A representation may be epistemically admissible while destroying civilizational reachability. Sudden access to accurate information about the actual mechanisms of social control, in the absence of replacement stabilizing structures, may produce disorientation, institutional collapse, and long-term reduction in the reachable futures available to the civilization.

The Wanderer's argument is precisely that Maxim has optimized for the wrong admissibility criterion.

Maxim's response is equally coherent. A civilization maintained through systematic epistemic inadmissibility is navigating toward futures that its members have not chosen and cannot evaluate. The reachability that the Wanderer seeks to preserve is not the reachability of the population. It is the reachability of a trajectory selected by a small immune minority operating under objectives the majority cannot access.

Both arguments are formally valid.

The novel provides no theorem that resolves them.

What it provides instead is the structure of the problem: two admissibility criteria, incompletely correlated, each locally compelling and globally in tension. The following chapters develop the formal tools required to state this tension precisely enough that the question of resolution can at last be asked clearly.

Part III: The Geometry of Intervention

7. Intervention Under Incomplete Observability

Every actor in *Prisoners of Power* who attempts to change the situation on Saraksh fails.

This observation is worth dwelling upon before any theoretical apparatus is introduced, because it is not the kind of observation that conventional political analysis tends to produce. Political analysis typically asks which actor had the correct diagnosis and what prevented them from implementing it. The implicit assumption is that correct diagnosis plus sufficient power produces successful intervention.

The novel systematically refuses this assumption.

The actors who fail include not only those with incorrect diagnoses but also those with sophisticated, partially correct, and even remarkably accurate understandings of the system. Failure is not distributed according to epistemic merit. It is distributed according to the structure of the problem itself.

This suggests that the novel's deepest subject is not any particular actor's inadequacy. It is the conditions under which intervention in complex systems becomes possible at all.

7.1. Load-Bearing Constraints

The central error in Maxim's intervention theory can be stated precisely.

He treats the broadcast system as a pure restriction on reachable cognitive states. Under this model, the system reduces the population's accessible future manifold. Removing the system therefore expands that manifold. The intervention is unambiguously beneficial.

The Wanderer's opposing argument rests on a different structural observation. The broadcast system is not only a restriction. It is also a stabilizing constraint on a civilization that has recently undergone catastrophic nuclear war, that operates under conditions of severe material scarcity, that faces external military threats, and that has organized its social and psychological life around the rhythms and affective states the broadcasts produce.

Under these conditions, the constraint does not merely reduce the population's accessible futures. It also partially constitutes the navigational infrastructure within which any future becomes accessible at all.

The distinction between a restrictive constraint and a load-bearing constraint is not visible from within the framework Maxim employs. His framework preserves

the distinction between freedom and oppression while collapsing the distinction between constraints that limit and constraints that also support.

7.2. The Successive Objective Functions

The pattern of Maxim's failure across the novel can be made precise by examining not the content of his beliefs but the structure of the objective functions he successively adopts.

In the first phase, the objective is survival and contact. The frame collapses the distinction between a society that is less developed in a neutral sense and one whose institutional structure is organized around coercion and concealment.

In the second phase, the objective is resistance to tyranny. The frame captures surface political organization while collapsing the distinction between a tyranny maintained through force and one maintained through neurological manipulation of the population's representational space.

In the third phase, the objective is elimination of the broadcast mechanism. The frame correctly identifies the primary mechanism while collapsing the distinction between a constraint that limits and a constraint that also stabilizes.

In the fourth phase, which is largely implicit in the novel's ending, the objective becomes post-liberation stabilization. This is the objective for which Maxim has no prepared framework at all. He has spent the novel developing increasingly sophisticated analyses of what is wrong with Saraksh. He has not developed any framework for what a navigable Saraksh would look like or how to reach it from the current state.

This asymmetry is not accidental. It is a structural feature of how the novel positions its protagonist. Maxim's analytical resources are organized around diagnosis. The problem of repair under uncertainty receives almost no theoretical attention until the consequences of his intervention have already materialized.

7.3. Why Every Actor Fails

The systematic failure of every intervention in the novel suggests that the problem is not located in any particular actor's limitations. It is located in the structure of the situation itself.

The Unknown Fathers maintain the broadcast system as a permanent feature of governance. Their intervention succeeds in maintaining stability across decades while systematically preventing the population from developing the cognitive resources that would make the control mechanism unnecessary. The Fathers' intervention preserves civilizational continuity at the cost of indefinitely deferring the conditions under which the intervention could be legitimately ended.

The underground attempts to build resistance through information and organization. Their intervention is epistemically admissible in the sense that it operates through accurate information and genuine consent. It fails to achieve civilizational scale because the representational space of the susceptible majority cannot accommodate the distinctions the underground's analysis requires.

Maxim attempts to destroy the control mechanism directly. His intervention is physically successful. It produces consequences that exceed his reachability analysis by an order of magnitude.

The mutants refuse intervention entirely. Their refusal is epistemically sophisticated: they correctly identify that the proposed futures following from available interventions are worse than the current situation from their perspective. Yet refusal is itself an intervention in the sense that it determines which other interventions become possible.

The Wanderer pursues the most sophisticated intervention: gradual systemic reform operating below the threshold of political visibility, designed to improve conditions incrementally without triggering the instabilities that rapid change would produce. His intervention is destroyed by Maxim's action.

The observation that emerges from surveying these failures is not that some actors were right and others wrong. It is that the system resists intervention at every level of sophistication.

7.4. The Intervention Problem Formalized

Let $\mathcal{S}$ denote the state space of a social system and $\mathcal{I}$ a class of interventions available to an agent. An intervention $\iota \in \mathcal{I}$ maps a current state $s \in \mathcal{S}$ to a distribution over future states $\iota(s) \in \Delta(\mathcal{S})$.

The agent's observational access to $\mathcal{S}$ is mediated by a projection $\pi : \mathcal{S} \to \Sigma$, where Σ is the agent's representational space. The agent selects interventions based not on s directly but on $\pi(s)$. If π is inadmissible for the decision class associated with intervention design, the agent's selection procedure will systematically produce interventions whose actual consequences in $\mathcal{S}$ differ from their anticipated consequences in Σ.

This gap between anticipated and actual consequences is not reducible to prediction error in any ordinary sense. It is a structural consequence of the mismatch between the distinctions preserved in Σ and the distinctions upon which the consequences of interventions in $\mathcal{S}$ depend.

The intervention problem under incomplete observability can therefore be stated as follows.

Given an agent with projection π, a system state s, and a target future $s^* \in \mathcal{S}$, does there exist an intervention $\iota \in \mathcal{I}$ such that $\iota(s)$ places high probability mass

on states near s^*, and can the agent identify this intervention from within Σ?

In general, the answer depends on whether the distinctions required to identify the correct intervention are preserved by π. If they are not, no amount of reasoning within Σ will reliably produce the correct selection.

8. Political Conflict as Competing Repair Programs

The preceding analysis has established that every actor in *Prisoners of Power* fails to produce an admissible intervention. What it has not yet explained is why the actors fail in such different ways and toward such different ends.

The actors do not simply disagree about methods. They disagree about what is broken.

8.1. Anomalies and Repair Operators

Definition 8.1. Let a system state $s \in \mathcal{S}$ be evaluated against a reference standard $s^* \in \mathcal{S}$. An *anomaly* is any feature of s that diverges from s^* in a way the evaluating agent regards as requiring correction. The *anomaly set* $\mathcal{A}$ of an agent is the collection of features they identify as requiring repair. A *repair operator* $\rho : \mathcal{S} \to \mathcal{S}$ is an intervention designed to eliminate or reduce the anomalies in $\mathcal{A}$.

The anomaly set is not determined by the state of the system alone. It is determined by the intersection of the system's current state and the agent's projection. Features that are invisible within the agent's projection cannot appear in their anomaly set. Features that are visible but assigned low priority may be structurally important while receiving no attention from the repair operator. And features that the agent regards as anomalies may be load-bearing constraints that the system requires for stability.

Political conflict, on this analysis, is rarely a disagreement about facts. It is a collision between repair operators designed to eliminate incompatible anomaly sets.

8.2. The Anomaly Sets of Saraksh

For the Unknown Fathers, the primary anomaly is civilizational vulnerability. The broadcasts are not a problem to be repaired. They are a repair already implemented.

The underground resistance selects a different anomaly set. For them, the primary anomaly is the suppression of cognitive freedom and political self-determination. The broadcasts are not a solution but the central problem.

The mutants identify yet another anomaly set. Their primary concern is survival and the preservation of their community's distinctive mode of life. They evaluate proposed repairs not by asking whether they address the underground's anomalies or the Fathers' anomalies but by asking whether the futures following from those repairs contain viable trajectories for their own community.

The Wanderer's anomaly set is the most comprehensive and temporally extended. He identifies the primary anomaly as the absence of conditions under which Saraksh can develop into a self-determining civilization. Every feature of the current situation is evaluated against this standard. The broadcasts are anomalous not because they suppress freedom in the present but because they prevent the development of the cognitive and institutional resources that long-term civilizational health requires. Yet their sudden removal is equally anomalous because it destroys the stability that any developmental trajectory requires as a foundation.

8.3. Why Identical Observations Generate Incompatible Repair Strategies

When two agents operate from incompatible projections, the information produced by one agent's analysis is not straightforwardly receivable by the other. The distinctions that make a piece of evidence significant within one framework may not be preserved within the other. Argument that proceeds from premises visible within one projection and invisible within another will appear, from within the second projection, as a non sequitur.

None of the communication failures in the novel are failures of intelligence or good faith. They are failures of projection compatibility.

8.4. Repair-Induced Damage

Every repair operator in the novel produces damage alongside whatever repair it achieves. This follows from the structure of the situation.

When the anomaly set of a repair operator omits features of the system that are relevant to stability, the operator will treat those features as background conditions rather than as variables its intervention will affect. When the intervention proceeds, the omitted features respond to the change in ways the operator did not anticipate.

The Fathers' repair of civilizational vulnerability produces a population permanently dependent on a control mechanism that prevents the development of autonomous capacity. The repair addresses the immediate anomaly while creating a structural condition that becomes increasingly anomalous over time.

The underground's repair of cognitive suppression, were it to succeed at scale, would expose a traumatized population to accurate information without the

institutional infrastructure required to support the transition.

Maxim's repair of the broadcast mechanism produces the withdrawal crisis, the political vacuum, the economic disruption, and the invitation to imperial invasion that the Wanderer had been working to prevent.

The mutants' repair of their community's survival produces a permanent withdrawal from any participation in the broader situation, ensuring that the most epistemically sophisticated non-aligned actors remain unavailable for any intervention that might otherwise benefit from their analysis.

In each case the repair operator is locally coherent. In each case its application produces consequences that the operator's anomaly set did not contain.

The political conflict on Saraksh is therefore not a conflict between those who understand the system and those who do not. It is a conflict among repair programs each of which understands part of the system well enough to intervene effectively within that part, and none of which understands the full system well enough to predict the consequences of its own success.

This is the situation into which Rudolf Sikorski has inserted himself, with more information, longer time horizons, and more sophisticated intervention theory than any other actor in the novel.

His failure, examined in the chapter that follows, is therefore not a personal failure.

It is a demonstration.

9. Rudolf Sikorski and the Limits of Repair

The figure of Rudolf Sikorski occupies an unusual position in the architecture of *Prisoners of Power*.

He appears late. He operates in disguise. His objectives remain opaque until the novel's final pages. When he finally reveals himself and his analysis, the narrative has already ended in the sense that matters most: the Control Center has been destroyed, the broadcasts have ceased, and the consequences are already propagating through the system in ways that cannot be reversed.

This timing is not a dramatic convenience. It is the novel's central structural argument. The most sophisticated actor in the system arrives too late to prevent the least sophisticated actor's intervention from determining the future.

9.1. The Progressor as Theoretical Construct

The Progressor represents the Noon Universe's best answer to the intervention problem. Not a conqueror, not a missionary, not a development economist imposing external models, but a sophisticated analyst operating from within the system

with a long time horizon and a commitment to indigenous development rather than external imposition.

Sikorski is among the most experienced of these agents. His presence on Saraksh is not an accident of the plot but a theoretical statement: the novel is asking what happens when the best available theory of intervention encounters a system of sufficient complexity.

9.2. What Sikorski Understands

Sikorski understands the broadcast system at a technical and functional level that no native actor possesses. He understands the geopolitical situation within a context invisible to anyone operating from within the planet's own information environment. He understands intervention theory.

He knows that rapid systemic change in a fragile civilization produces consequences that outrun the capacity of existing institutions to absorb them. He knows that the removal of control mechanisms without replacement stabilizing structures generates the very crises that the mechanisms were originally installed to prevent.

His strategy reflects this understanding. He does not attempt to destroy the broadcasts, overthrow the Fathers, or accelerate the underground's revolutionary program. He works at the margins of the system, creating conditions that will eventually make the control architecture unnecessary by building the institutional and cognitive infrastructure that would allow the population to function without it.

9.3. The Structure of Sikorski's Failure

Sikorski's plan is destroyed by Maxim.

The mechanism of destruction is not military, political, or ideological. It is projective.

Maxim operates from a framework that does not contain the distinctions Sikorski's analysis requires. He cannot see the load-bearing function of the broadcasts because his framework preserves the distinction between freedom and oppression while collapsing the distinction between constraint and support. He cannot see Sikorski's long-term developmental strategy because it is invisible by design, operating below the threshold of political visibility.

The collision between Maxim's framework and Sikorski's framework is not a collision between ignorance and knowledge. It is a collision between two projections that preserve different distinctions and therefore generate different anomaly sets and different repair operators.

Neither projection is complete. The difference is that Sikorski's projection is

organized around the consequences of interventions across long time horizons, while Maxim's projection is organized around the justice of the present state.

Both orientations track something real. Their incompatibility is not resolvable by simply adding one agent's information to the other's framework, because the information each agent possesses is structured by the projection that organized its collection.

9.4. Intervention, Control, and Irreversibility

Irreversibility transforms the structure of the intervention problem fundamentally.

In a reversible system, the cost of intervention under incomplete observability is bounded by the cost of correction. Agents can afford to act under uncertainty because errors can be identified and addressed.

In an irreversible system, the cost of intervention under incomplete observability is potentially unbounded. A single inadmissible intervention may initiate trajectories that no subsequent action can redirect toward the original target.

Sikorski understands this. His strategy of incremental, invisible intervention is designed precisely to preserve reversibility. Small changes can be observed, evaluated, and adjusted. Large rapid changes cannot.

Maxim does not understand this. His framework does not preserve the distinction between reversible and irreversible interventions as a primary evaluative criterion.

9.5. What Sikorski's Failure Means for Repair Theory

The analysis of Sikorski's situation reveals a limitation in repair theory as currently formulated.

Existing formulations assume that repair operators can be designed independently and applied sequentially. This model does not account for the interaction between repair operators designed by different agents operating from different projections in the same system simultaneously.

On Saraksh, multiple repair programs are active at the same time. Each program is designed to address a real anomaly. Each program produces consequences that affect the conditions under which other programs operate. The programs are not coordinated. They cannot be coordinated, because the projections from which they are designed do not preserve sufficient common structure to support coordination.

The novel suggests that sufficiently complex social systems require a theory of intervention that accounts for the simultaneous activity of multiple incompatible repair programs and the interactions among them.

Current formal frameworks do not provide this theory.

The Wanderer understood the problem more clearly than any other actor in the novel. He did not solve it. Whether the problem is solvable in principle is a question the novel poses without answering. It is also a question that the mathematics of admissibility and reachability has only begun to approach.

Part IV: The Ending and Its Implications

10. The Ending as Reachability Problem

The final pages of *Prisoners of Power* refuse the satisfactions that the preceding narrative has repeatedly made available and then withdrawn.

Maxim has destroyed the Control Center. The broadcasts have ceased. The Wanderer has delivered his accounting of the consequences already in motion. Famine, anarchy, institutional collapse, the withdrawal crisis affecting perhaps a fifth of the population, and the looming invasion from the Island Empire whose suppression the Wanderer had been planning through mechanisms now unavailable.

Maxim listens. He does not recant.

He acknowledges the consequences. He acknowledges that he did not foresee them. He acknowledges that the Wanderer's analysis is probably correct in its broad outlines.

And then he refuses to leave.

He will stay on Saraksh. He will help stabilize the situation. He remains glad that he destroyed the Control Center because now the people can be in charge of their own destiny.

The novel ends on this refusal. Not on resolution. Not on vindication. Not on condemnation. On the coexistence of two formally valid positions that the narrative declines to adjudicate.

10.1. Two Readings of the Same Event

The destruction of the Control Center admits two interpretations that the novel holds simultaneously without resolving.

The first interpretation is Maxim's. The broadcast system was a mechanism of neurological coercion imposed on a population without consent. It systematically prevented the population from forming accurate representations of their own situation. Whatever instabilities follow from its removal are instabilities the population is now at least capable of navigating as agents rather than as managed objects.

The second interpretation is the Wanderer's. The broadcast system was a load-bearing constraint in a civilization that had recently nearly destroyed itself. The instabilities following from its removal are not navigable by a population

that has organized its psychological and social life around the system's rhythms for generations. The removal does not produce self-determination. It produces disorientation at civilizational scale.

Both interpretations are accurate descriptions of the same event. They do not contradict each other at the level of fact. They diverge at the level of which facts are assigned priority within the evaluative framework each interpreter employs.

The formal analysis that follows does not resolve this divergence. It explains why the divergence is not resolvable from within the information available to any actor in the system.

10.2. Formal Reachability and Effective Reachability

To make the disagreement precise, it is necessary to distinguish two quantities that political analysis typically conflates.

Let $\mathcal{A}_C$ denote the admissible region of a civilization under a constraint field C, and let $\mathcal{A}_0$ denote the admissible region after removal of the constraint. Define the formal reachability volume as

$$V_R(\mathcal{A}) = \int_{\mathcal{A}} w(x)\, d\mu(x),$$

where $w(x)$ is a viability weighting function and μ is the induced measure on the admissible manifold.

The naive interpretation of liberation assumes $\mathcal{A}_C \subseteq \mathcal{A}_0$ and therefore $V_R(\mathcal{A}_0) \geq V_R(\mathcal{A}_C)$. Removing a constraint expands the admissible region. The argument appears to close here. It does not close here.

Let $N(x,t)$ denote the navigational capacity of agents occupying state x at time t. Define the effective reachability volume as

$$V_E(\mathcal{A}, t) = \int_{\mathcal{A}} N(x,t)\, w(x)\, d\mu.$$

Proposition 10.1 (Reachability Tradeoff)**.** *There exist systems for which*

$$V_R(\mathcal{A}_0) \geq V_R(\mathcal{A}_C)$$

while simultaneously

$$V_E(\mathcal{A}_0, t_0) < V_E(\mathcal{A}_C, t_0)$$

for some interval $[t_0, t_1]$ immediately following constraint removal.

Proof. Constraint removal enlarges the admissible manifold, so formal reachability does not decrease. However, navigational capacity evolves on a slower timescale

than constraint geometry. Immediately after removal,

$$N(x, t_0) \approx N_C(x),$$

where N_C was adapted to the previous constraint structure. The enlarged manifold $\mathcal{A}_0$ contains regions for which $N(x, t_0) \approx 0$, since agents have not yet developed the capacity to navigate states that were previously inadmissible. The contribution of these regions to effective reachability remains negligible despite their formal availability. Therefore formal reachability increases while effective reachability decreases over $[t_0, t_1]$:

$$V_R \uparrow, \qquad V_E \downarrow . \qquad \square$$

$\square$

Proposition 10.1 formalizes the disagreement between Maxim and the Wanderer with precision that neither character possesses. Maxim argues from formal reachability. The Wanderer argues from effective reachability. They are not disagreeing about the same quantity. They are computing different things and calling both of them liberation.

10.3. The Lamphrodyne Interpretation

Within RSVP field theory, the situation on Saraksh after the broadcasts cease can be characterized as a lamphrodyne event [21].

Let $\Phi(x, t)$ represent the local density of cognitive possibility at position x and time t, and let $\mathbf{v}(x, t)$ represent the transport field carrying navigational capacity through the social manifold. The broadcast system operates as a constraint on Φ, suppressing possibility density in regions of cognitive space corresponding to accurate representations of the political system. The transport field $\mathbf{v}$ has been calibrated over generations to the constraint geometry.

Theorem 10.2 (Lamphrodyne Overshoot)**.** *Let τ_C denote the timescale of constraint removal and τ_v the adaptation timescale of the transport field. If*

$$\tau_C \ll \tau_v,$$

then there exists an interval $[t_0, t_1]$ during which expansion of possibility exceeds transport capacity:

$$\frac{d\Phi}{dt} \gg |\nabla \cdot \mathbf{v}| .$$

During this interval the system enters a state of lamphrodyne overshoot.

Proof. Removal of the constraint produces a discontinuity in the available possibility field,

$$\Delta\Phi > 0,$$

on timescale τ_C. The transport field evolves continuously on timescale τ_v. Since $\tau_C \ll \tau_v$,

$$\mathbf{v}(t_0^+) \approx \mathbf{v}(t_0^-).$$

Immediately after removal,

$$\frac{d\Phi}{dt} \gg \frac{d\mathbf{v}}{dt}.$$

The possibility field therefore expands faster than navigational capacity can be transported through it. Regions become formally reachable before institutions can support navigation toward them. The resulting mismatch constitutes lamphrodyne overshoot. □ □

Theorem 10.2 characterizes the Wanderer's prediction of mass withdrawal psychosis, institutional collapse, and civilizational disorientation as a prediction of lamphrodyne overshoot. What the theorem provides is a characterization precise enough to support the following question: under what conditions does lamphrodyne overshoot remain bounded, and under what conditions does it propagate until it destroys the navigational infrastructure of the system entirely? That question the novel cannot answer. It is an empirical question about the trajectory of Saraksh after the final page, and also the question that makes the novel's ending the beginning of a research program rather than a conclusion.

10.4. Constraint Removal Non-Monotonicity

Most political theories rest on an implicit assumption so obvious it rarely receives explicit statement:

$$C_1 < C_2 \implies V_R(C_1) > V_R(C_2).$$

Fewer constraints imply greater freedom. Greater freedom implies greater reachability. The direction of the inequality is taken as self-evident.

The ending of *Prisoners of Power* directly challenges this assumption, and the challenge can be stated with mathematical precision.

Let $V_R(C)$ denote the formal reachability volume of a civilization under constraint intensity C, and let $S(C)$ denote the stabilizing infrastructure available to the population at that constraint level. Define effective reachability as the product

$$V_E(C) = V_R(C) \cdot S(C).$$

The standard liberation assumption treats S as independent of C: removing a

constraint expands V_R without affecting S. Under this assumption, V_E is monotonically decreasing in C and the Wanderer's position is simply wrong.

The novel proposes a different relationship.

Theorem 10.3 (Constraint Removal Non-Monotonicity)**.** *Suppose stabilizing infrastructure is itself partially produced by the constraint field, so that*

$$\frac{dS}{dC} > 0.$$

Then $V_E(C) = V_R(C) \cdot S(C)$ need not be monotonically decreasing in C. In particular, there may exist a critical value $C^\star$ satisfying

$$\left.\frac{dV_E}{dC}\right|_{C^\star} = 0,$$

at which effective reachability is locally maximal. Consequently,

$$V_E(C^\star) > V_E(0).$$

A partially constrained civilization can possess strictly greater effective reachability than a completely unconstrained one.

Proof. Write $V_E(C) = V_R(C) \cdot S(C)$. Differentiating with respect to C,

$$\frac{dV_E}{dC} = \frac{dV_R}{dC} \cdot S(C) + V_R(C) \cdot \frac{dS}{dC}.$$

Since increasing constraint reduces formal reachability, $dV_R/dC < 0$. By hypothesis, $dS/dC > 0$. The two terms therefore have opposite sign.

At $C = 0$, the constraint is absent and $S(0)$ may be small or zero if stabilizing infrastructure depends on the constraint for its production. At large C, $V_R(C)$ is severely suppressed. Between these extremes, the product $V_R(C) \cdot S(C)$ attains an interior maximum at $C^\star$ satisfying $dV_E/dC = 0$, i.e.

$$\frac{dV_R}{dC} \cdot S(C^\star) = -V_R(C^\star) \cdot \left.\frac{dS}{dC}\right|_{C^\star}.$$

At this point $V_E(C^\star) > V_E(0)$ whenever $S(0)$ is sufficiently small relative to $S(C^\star)$. □ □

Theorem 10.3 does not establish that the Wanderer is correct. It establishes something more important: that Maxim's position is not mathematically trivial. If the broadcast system has been generating stabilizing infrastructure over the decades of its operation — if the cognitive habits, institutional routines, and

social coordination mechanisms of Saraksh's population have formed around the constraint rather than independently of it — then the assumption $dS/dC > 0$ holds, and the constraint removal non-monotonicity theorem applies.

Whether $C^\star$ is close to the actual broadcast intensity, and whether Sikorski's incremental program was designed to reduce C toward $C^\star$ rather than toward zero, are empirical questions the novel leaves open.

What the theorem settles is the logical structure of the disagreement: both positions are consistent with the mathematics.

10.5. The Admissibility–Stability Tradeoff

The preceding analysis treats the disagreement between Maxim and the Wanderer as a disagreement about facts: what the broadcasts actually do, what their removal would actually produce. But there is a deeper disagreement running beneath the factual one, and it is of a different kind.

Throughout the novel, two distinct optimization targets are in play.

Let $A(t)$ denote epistemic admissibility: the degree to which the population's representational frameworks preserve the distinctions required for accurate understanding of their situation. Let $P(t)$ denote population stability: the degree to which the civilization maintains coherent institutional function, physical welfare, and social coordination.

The Unknown Fathers maximize $P(t)$ at the cost of $A(t)$. The broadcasts suppress admissibility in order to maintain stability.

Maxim maximizes $A(t)$ at the cost of $P(t)$. He destroys the mechanism of admissibility suppression even at the cost of the stability it was maintaining.

Suppose a civilization's welfare can be expressed as a utility functional of the form

$$U = \alpha A + \beta P,$$

where $\alpha, \beta \geq 0$ are weighting coefficients reflecting how the civilization values epistemic admissibility relative to population stability.

Proposition 10.4 (Admissibility–Stability Tradeoff)**.** *The disagreement between Maxim and the Wanderer is equivalent to a disagreement about the coefficients (α, β) in the utility functional $U = \alpha A + \beta P$. In particular, no accumulation of shared factual information about the state of Saraksh can resolve a disagreement that is constituted by different utility geometries.*

Proof. Suppose Maxim and the Wanderer share complete information about the current state of Saraksh and agree on the causal consequences of all available interventions. Their disagreement about what to do therefore cannot arise from different beliefs about facts or consequences.

If they nevertheless recommend different interventions, this must reflect different evaluations of the same consequence profile. Under the utility functional $U = \alpha A + \beta P$, Maxim's preference for epistemic transparency implies a high ratio α/β, while the Wanderer's preference for civilizational stability implies a low ratio α/β.

For any finite body of factual information $\mathcal{F}$, if both agents update their beliefs identically on $\mathcal{F}$, their posterior recommendations will still diverge whenever $\alpha_{\text{Maxim}}/\beta_{\text{Maxim}} \neq \alpha_{\text{Wanderer}}/\beta_{\text{Wanderer}}$.

Additional facts therefore cannot close the gap. The conflict is not epistemic. It is variational. □ □

This result illuminates one of the novel's most frustrating structural features: the failure of argument between the factions.

Throughout *Prisoners of Power*, characters who share information do not converge in their recommendations. Readers accustomed to treating political disagreement as a consequence of factual ignorance may find this inexplicable. Proposition 10.4 provides the explanation. When disagreement is variational rather than epistemic, more information does not resolve it. It merely sharpens the contours of the divergence.

The Wanderer's exhaustion in the final scene is not the exhaustion of someone who has been proved wrong. It is the exhaustion of someone who has been unable to communicate that the other party is optimizing a different objective function. That communication failure is not rhetorical. It is mathematical.

10.6. The Projection Horizon Theorem

The three results developed above characterize the structure of the disagreement at the novel's ending. They do not yet explain why the novel refuses to adjudicate between the positions.

The final result of this section provides that explanation.

Theorem 10.5 (Projection Horizon)**.** *Let H denote the prediction horizon available to an observer within the system, and let $\Delta R(T)$ denote the difference in cumulative effective reachability between two interventions over evaluation horizon T. If $T > H$, then the sign of $\Delta R(T)$ cannot generally be determined from information available to any observer within the system.*

Proof. Evaluation of $\Delta R(T)$ requires knowledge of both intervention trajectories across the full horizon T. An observer with prediction horizon $H < T$ has direct access only to trajectory segments of length at most H.

Any estimate of $\Delta R(T)$ from horizon-H information therefore requires a projection operator π_H mapping observable segments to full-horizon trajectory estimates. Different projection operators yield different estimates:

$$\pi_H^{(1)}(\Delta R(T)) \;\neq\; \pi_H^{(2)}(\Delta R(T))$$

in general. Since no observer has access to the unrealized counterfactual trajectory or to exogenous disturbances $\eta(t)$ for $t > H$, no canonical choice of π_H is available.

Therefore $\text{sign}(\Delta R(T))$ cannot be determined from within the system for $T > H$. □ □

Theorem 10.5 captures the deepest point of the novel's ending, and it is worth stating carefully.

The problem is not that Maxim lacks information that the Wanderer possesses. The problem is that the relevant information does not yet exist. It resides in the future. The quantity required to adjudicate their disagreement — the sign of the difference in cumulative effective reachability between the world where Maxim acts and the world where he does not — lies beyond the prediction horizon of any actor inside the system at the moment of decision.

This is why the novel's unresolved ending feels structurally necessary rather than evasive. The Strugatskys have not withheld an answer. They have constructed a situation in which the answer is not yet available to anyone inside the story.

The uncertainty is not psychological. It is not a matter of insufficient research or insufficient courage. It is a structural feature of the problem: the horizon of evaluation exceeds the horizon of observation, and no projection operator bridges the gap without importing assumptions that are themselves contested.

Maxim knows this, at some level, when he refuses to leave at the novel's end. He cannot prove he was right. He also cannot be proved wrong. He stays because the question remains open, and staying is the only form of commitment available to someone who cannot compute the answer.

10.7. The Counterfactual Indeterminacy Theorem

The preceding results establish that constraint removal can decrease effective reachability in the short term (Proposition 10.1), that rapid removal produces lamphrodyne overshoot (Theorem 10.2), that the liberation assumption is not mathematically trivial (Theorem 10.3), that the disagreement between Maxim and the Wanderer is variational rather than epistemic (Proposition 10.4), and that the sign of the long-run reachability difference lies beyond any actor's projection

horizon (Theorem 10.5). They do not establish that Maxim's intervention was wrong.

To establish that, one would need to compare the long-term trajectory generated by the intervention against the long-term trajectory that would have obtained without it. This comparison is not available.

Theorem 10.6 (Counterfactual Indeterminacy)**.** *No actor within the system possesses sufficient information to determine the sign of*

$$\Delta R(T) = R(\tau_M, T) - R(\tau_W, T),$$

where τ_M is the trajectory generated by destruction of the Control Center, τ_W is the trajectory generated by continuation of the Wanderer's incremental program, and cumulative effective reachability over horizon T is defined by

$$R(\tau, T) = \int_0^T V_E(\tau(t), t)\, dt.$$

Proof. Evaluation of $\Delta R(T)$ requires knowledge of both trajectories across the full horizon T. Only τ_M is realized. The counterfactual trajectory τ_W is unobservable by definition. Furthermore, both trajectories depend upon future contingent events,

$$\tau(t) = F(x_0, \eta(t)),$$

where $\eta(t)$ represents exogenous disturbances including military action, institutional adaptation, ecological and material conditions, elite decisions, and technological developments on timescales ranging from months to generations. Since neither Maxim nor the Wanderer has access to future values of $\eta(t)$, and since the counterfactual τ_W is by definition not instantiated, neither can compute $\Delta R(T)$. The quantity required to adjudicate the dispute is therefore not merely unknown but epistemically inaccessible to any actor operating within the system. □ □

Theorem 10.6 establishes that the disagreement between Maxim and the Wanderer is not a disagreement that more information would resolve. Additional information about the current state of Saraksh would not provide access to the counterfactual trajectory.

The novel ends without resolving a question that is formally unresolvable. This is not irresolution. It is precision. The reader who finishes the novel wanting a verdict has misidentified the question.

The question is not what should have been done.

The five results of this section jointly explain why no actor within the system could answer it. Constraint removal is not monotone in effective reachability

(Theorem 10.3), so the direction of the liberation argument is not self-evident. The disagreement between Maxim and the Wanderer is variational rather than epistemic (Proposition 10.4), so shared facts cannot resolve it. The sign of the long-run reachability difference lies beyond any actor's prediction horizon (Theorem 10.5), so the relevant quantity is not computable from within the system. Rapid constraint removal produces lamphrodyne overshoot (Theorem 10.2), so short-term evidence is not diagnostic of long-run outcomes. And cumulative effective reachability across both trajectories is epistemically inaccessible (Theorem 10.6), so even retrospective evaluation remains underdetermined.

The Strugatskys did not construct an ending that refuses resolution because they could not decide. They constructed an ending that refuses resolution because the question it poses is structurally undecidable from within the information available to any actor at the moment of decision.

That is the research program the following chapter begins to sketch.

Part V: Beyond Admissibility

11. Beyond Admissibility

The preceding chapters have developed a reading of *Prisoners of Power* through the concepts of projection, admissibility, reachability, and repair. This reading has produced genuine results. The novel's recurring pattern of model collapse is more precisely characterized as projection failure than as ignorance or naivety. The broadcast system is more precisely characterized as an admissibility architecture than as propaganda. The political conflict is more precisely characterized as a collision among incompatible repair programs. The ending is more precisely characterized as a reachability problem than as a political resolution.

Yet the preceding analysis has also repeatedly encountered phenomena that the framework handles awkwardly or does not handle at all. The purpose of this chapter is to examine those phenomena directly. Not to explain them away. Not to extend the framework prematurely to cover them. But to characterize as precisely as possible the boundary between what the current mathematics can reach and what lies beyond it.

That boundary is itself a theoretical result.

11.1. Loyalty

The first phenomenon that resists formalization is loyalty.

Throughout the novel, characters act in ways that cannot be fully explained by their assessment of outcomes. Maxim's companions in the underground maintain their commitments under conditions where rational reachability analysis would recommend withdrawal. Soldiers follow orders whose consequences they can observe to be catastrophic. Members of the resistance accept imprisonment and death for objectives whose achievement they are unlikely to witness.

One might attempt to absorb loyalty into the framework by expanding the objective function: the agent values consistency of commitment independently of outcomes. This move is formally possible. It is also explanatorily empty. To say that an agent values loyalty is to redescribe the phenomenon without explaining it. The question is why loyalty functions as a stable attractor in human social systems, why it resists revision under conditions where outcome-based reasoning would recommend revision, and what role it plays in the stability and navigability of the social systems within which it appears.

These questions require a theory of motivational structure that the current framework does not provide.

11.2. Charisma

The second phenomenon is charisma.

Maxim himself is an example. He repeatedly persuades individuals to assist him, join him, or trust him in situations where rational assessment of his track record would not support such trust. His persuasiveness is not primarily a function of the arguments he makes. It is a function of something the text conveys as personal presence, energy, and conviction.

Charisma of this kind has observable effects on the reachability volumes of social systems. Leaders who possess it can mobilize collective action that would not occur in their absence, sustain commitment through periods of failure, and produce coordination among agents whose individual assessments would not generate coordinated action. Yet charisma does not appear naturally within the formal vocabulary developed in this essay. It is not a property of a state, a transition, an anomaly set, or a repair operator. It is a property of the relationship between a particular agent and the social field within which that agent operates.

Whether this relationship can be formalized within a geometric framework remains an open question.

11.3. Fear

The third phenomenon is fear.

Fear in *Prisoners of Power* is not simply a negative utility assigned to certain outcomes. It is a cognitive and motivational state that restructures the space of actions an agent is capable of considering. Under sufficient fear, options that are formally available become practically inaccessible.

This contraction is not identical to the contraction produced by broadcast constraint. Broadcast constraint operates externally, shaping the representational space through which the agent processes information. Fear operates internally, altering the relationship between the agent's representations and their motivational structure.

Part of what the Wanderer is predicting when he describes the consequences of Maxim's action is not merely institutional collapse but psychological collapse: a population that has organized its affective life around the management of fear, and whose management strategies have been suddenly rendered unavailable without replacement. The formal framework can identify that practical navigability and formal reachability have diverged. It cannot yet characterize the internal structure of the divergence in terms that would support the design of interventions specifically targeted at restoring navigability.

11.4. Historical Momentum

The fourth phenomenon is historical momentum.

At several points in the novel, events proceed in ways that no actor intended and that the preferences and actions of individual actors do not fully explain. The war against Honti is initiated through a process in which each participant follows a locally rational path toward a collectively irrational outcome. No single actor decides that the war should occur. The war occurs anyway because the interaction of multiple repair programs, each responding to real anomalies within its own framework, produces a trajectory that none of the programs was designed to generate.

Reachability theory can describe the phenomenon in retrospect. What it does not yet provide is a systematic method for identifying, in advance, which interactions among simultaneously active repair programs will produce emergent trajectories that none of the programs was designed to generate.

Historical momentum is not noise. It is the structured interaction of multiple repair programs operating from incompatible projections in a system where the interaction terms are as causally significant as the individual programs themselves. A theory adequate to the phenomenon would need to characterize the interaction structure, not merely the individual programs.

11.5. Moral Commitment

The fifth phenomenon is moral commitment.

Maxim's refusal to recant at the novel's end is not primarily a reachability judgment. He does not say that he has calculated superior outcomes. He says that the people can now be in charge of their own destiny, and that this matters regardless of the immediate consequences. This is a statement about the intrinsic value of self-determination that is not reducible to its effects on reachable futures.

One could attempt to treat moral commitment as a constraint on the objective function. This treatment is formally coherent but leaves the most important question unanswered. Where do the constraints come from? Why do they persist under pressure? Why do some moral commitments function as stable attractors while others dissolve under adversity? And why do individuals sometimes sacrifice their own reachable futures in service of moral commitments that benefit others they will never meet?

11.6. Narrative Agency

The sixth phenomenon is narrative agency.

Throughout the novel, characters act not only in response to their assessment of the current state but in response to their understanding of the story they are in. Maxim repeatedly makes decisions that are influenced by the kind of protagonist he believes himself to be. The Robinson Crusoe frame is not merely an interpretive error. It is a motivational structure that organizes which actions feel available, which feel appropriate, and which feel narratively coherent.

Agents do not simply optimize over reachable futures. They occupy narrative positions that shape which futures they are capable of orienting toward.

Whether narrative can be incorporated into a geometric theory of reachability, and if so through what formal machinery, is among the most interesting open questions the novel raises for the framework.

11.7. The Boundary and Its Significance

The six phenomena examined in this chapter share a common structure. Each is causally relevant to the trajectories of social systems. Each influences which futures become reachable and which do not. Each appears in the novel as a force that shapes outcomes in ways that purely structural analysis of state spaces and transition operators does not fully capture. And each resists clean formalization within the current framework.

This resistance is not a reason to abandon the framework. It is a reason to treat the framework as a first approximation rather than a complete theory.

The boundary between what the mathematics can reach and what lies beyond it is not a defect. It is a map of where the mathematics needs to grow.

The Strugatskys did not possess formal theories of admissibility, reachability, or repair. They possessed something else: the ability to construct imagined worlds complex enough that all of these phenomena appear together, interact with one another, and resist reduction to any single explanatory frame.

The novel sees further than the theory.

That is not a failure of the theory. It is the most useful thing a literary work can do for a theoretical program: identify the terrain that the program has not yet reached.

Part VI: Literature as Theory Discovery

12. Literature as Theory Discovery

This essay began with a methodological claim.

Literary works sometimes function not as illustrations of existing theories but as instruments for discovering phenomena that existing theories struggle to describe. The value of such works for theoretical inquiry lies not in the examples they provide but in the resistance they offer: the places where the narrative sees further than the formalism, where the imagined world contains structure that the available mathematics cannot yet reach.

Prisoners of Power has proved to be such a work.

The novel's treatment of projection failure is more precise than most formal accounts of epistemic error. The Strugatskys understood, without the vocabulary of admissibility theory, that the primary site of cognitive failure is not the reasoning process but the representational framework within which reasoning occurs. Maxim does not fail because he reasons badly. He fails because the categories through which he reasons do not preserve the distinctions the world requires.

The novel's treatment of constraint removal is more nuanced than most formal accounts of liberation. The Strugatskys understood, without the vocabulary of lamphrodyne dynamics [21], that the removal of a load-bearing constraint produces a different situation than the removal of a merely restrictive constraint. They understood that formal expansion of reachability and practical expansion of navigability are different quantities that can move in opposite directions (Proposition 10.1). They understood that the liberation assumption — fewer constraints imply greater freedom — fails when stabilizing infrastructure depends on the constraint for its production (Theorem 10.3). And they understood that the disagreement between Maxim and the Wanderer cannot be resolved by sharing more information because it is constituted by different utility geometries rather than different factual beliefs (Proposition 10.4).

The novel's treatment of intervention under incomplete observability is more rigorous than most formal accounts of political action. The Strugatskys understood, without the vocabulary of counterfactual indeterminacy, that the question of whether an intervention was correct cannot be answered from within the realized trajectory. They understood that the most sophisticated actor in a complex system cannot guarantee admissible outcomes because the interaction of multiple incompatible repair programs operating simultaneously produces emergent trajectories

that no single program was designed to generate.

These understandings were not expressed as theorems. They were expressed as narrative structure, character behavior, and the specific shape of the novel's refusal to resolve its central questions.

12.1. Two Modes of Inquiry

The relationship between narrative and mathematics is not one of translation.

Mathematics achieves precision by abstracting away from particularity. A theorem about reachability tradeoffs applies to any system satisfying the relevant conditions, regardless of whether that system involves Soviet-era broadcast technology, medieval religious institutions, contemporary social media architectures, or colonial governance structures. The abstraction is the source of the generality. But the abstraction is also a loss.

The formal framework developed in this essay cannot represent Maxim Kammerer as an individual. It can represent a projection operator, an anomaly set, a repair program. It cannot represent the specific texture of a young man from an idealistic civilization encountering, for the first time, a world in which good intentions produce catastrophic consequences.

That texture is not decoration. It is information.

The specific way in which Maxim's confidence collapses, the specific way in which he nevertheless refuses to recant, the specific way in which the Wanderer's exhaustion reads as the exhaustion of someone who has understood the problem for longer than Maxim has been alive: these particulars carry theoretical content that the formal framework cannot capture without losing what makes the content significant.

Literature achieves depth by preserving particularity. The novel can hold simultaneously, without resolving, the validity of Maxim's position and the validity of the Wanderer's position, because narrative can sustain contradiction in a way that formal proof cannot. The ending of *Prisoners of Power* is not a failure of resolution. It is the representation of a genuine contradiction whose resolution would require information that is structurally inaccessible.

The two modes are therefore complementary rather than competitive. Mathematics tells us what the structure of the problem is. Narrative tells us what it feels like to inhabit the problem. Neither is sufficient alone. Together they illuminate more than either can reach independently.

12.2. What the Strugatskys Discovered

The claim that a literary work can constitute theoretical discovery requires a standard of evidence. It is not sufficient to observe that a novel contains themes that happen to resemble concepts in a formal theory.

The stronger claim requires demonstrating that the novel exhibits structural features that cannot be fully accounted for by the theoretical frameworks available at the time of its composition, and that the specific structure of these features anticipates distinctions that subsequent theoretical work has found it necessary to introduce.

Prisoners of Power satisfies this standard in at least three respects.

The distinction between formal and effective reachability, which Proposition 10.1 formalizes, is not present in the political theory or cognitive science literature of 1969. The novel embeds this distinction in its plot structure: the entire tragedy of the ending depends on the gap between what the population can now formally access and what they can actually navigate toward. The distinction is not stated. It is enacted, repeatedly, at the level of narrative consequence.

The distinction between constraint removal and load-bearing constraint, which the lamphrodyne analysis formalizes, is present in systems theory literature of the period but not in the political theory literature that would have been the obvious reference for a novel about totalitarianism. The Strugatskys embed this distinction in the Wanderer's argument without reducing it to the political categories their readers would have found familiar.

The structure of counterfactual indeterminacy, which Theorem 10.6 formalizes, is present in philosophy of science literature but not in the narrative conventions of political fiction. Political fiction conventionally resolves by revealing which side was correct. The Strugatskys construct a situation in which this question is not merely unresolved but unresolvable, and they construct it precisely enough that the unresolvability feels like a demonstration rather than an evasion.

12.3. The Critical Reception as Evidence

The history of the novel's critical reception provides indirect evidence for this claim.

Critical interpretation of *Prisoners of Power* has consistently proceeded by locating the novel within existing conceptual frameworks: dystopian fiction, Soviet allegory, Cold War anxiety, outsider intervention narrative, anti-totalitarian critique. These classifications are not incorrect. But the critical literature also consistently reports a residue: something that the available frameworks do not quite capture, a quality of unresolvedness that readers experience as significant rather than merely

incomplete.

This residue is evidence that the novel contains structure that existing frameworks consistently omit. The frameworks locate what is familiar. The residue marks what is not.

The analysis developed in this essay has argued that the residue is not a vague aesthetic quality but a specific theoretical content: the novel's precise treatment of projection failure, load-bearing constraint, and counterfactual indeterminacy in the context of intervention under incomplete observability.

Critical readers have sensed this content without possessing the vocabulary to articulate it. The task of the present essay has been to develop that vocabulary by placing the novel in conversation with formal frameworks capable of reaching the structure the critical tradition has consistently felt but consistently failed to name.

12.4. The Research Program

The analysis has generated several open questions that constitute a research program rather than a conclusion.

The Reachability Tradeoff Proposition establishes that formal and effective reachability can diverge. It does not characterize the conditions under which the divergence is temporary and self-correcting versus the conditions under which it initiates a cascade that permanently reduces effective reachability. That characterization would require a theory of navigational capacity development: how do agents and institutions develop the capacity to navigate newly available states, at what rate, and under what conditions does the development process fail?

The Lamphrodyne Overshoot Theorem establishes that rapid constraint removal produces overshoot. It does not characterize the conditions under which the overshoot remains bounded versus the conditions under which it propagates to systemic collapse. That characterization would require a theory of transport field resilience.

The Counterfactual Indeterminacy Theorem establishes that the comparison between realized and counterfactual trajectories is epistemically inaccessible. It does not characterize what this inaccessibility implies for intervention design. If the consequences of intervention cannot be compared to the consequences of non-intervention, what standards should govern the decision to act?

These questions extend beyond the analysis of *Prisoners of Power*. They are questions about intervention in complex social systems under conditions of incomplete observability, irreversibility, and multiple simultaneously active repair programs. They are relevant to foreign intervention, development policy, institutional reform, ecological governance, and the alignment of artificial systems with human values.

The Strugatskys wrote a novel about a young man stranded on a damaged

planet who destroys a mind control system and produces consequences he did not anticipate. They also wrote, without intending to and possibly without knowing it, one of the most precise fictional treatments of the geometry of intervention under incomplete observability that the twentieth century produced.

Whether that geometry can be fully formalized remains to be seen. What this essay has attempted to demonstrate is that the attempt to formalize it is worthwhile, that the novel is a more precise guide to the relevant terrain than the political allegory reading suggests, and that the conversation between narrative and mathematics, when both are allowed to challenge one another, produces insights that neither can reach alone.

The Strugatskys did not solve the intervention problem. Neither has this essay. What the novel did, and what this essay has attempted to do in response, is make the problem visible in its full complexity: not as a failure of political will or moral clarity, but as a genuine difficulty in the geometry of knowing and acting in worlds whose most important structures cannot be directly observed.

That visibility is the beginning of understanding. It is also, for now, where understanding stops.

Conclusion: The Geometry of Control

Prisoners of Power is not primarily a novel about totalitarianism.

It is a novel about the conditions under which any agent, operating within any system whose most important structures remain partially hidden, can act in ways that reliably improve rather than damage the futures available to the inhabitants of that system.

The answer the novel provides is not encouraging.

Every actor fails. The most sophisticated actor fails last and most painfully, because his understanding is sufficient to anticipate the consequences of others' interventions without being sufficient to prevent them.

The formal analysis developed in this essay has not improved on this answer. It has, however, made the structure of the failure more precise.

Maxim and the Wanderer are not disagreeing about facts or values. They are computing different quantities: formal reachability and effective reachability, present constraint and future navigability, realized trajectory and counterfactual trajectory. The novel's refusal to adjudicate between them is not irresolution. It is the recognition that the quantity required to adjudicate—cumulative effective reachability across the relevant long-term horizon, compared across realized and counterfactual trajectories—is not accessible to any actor operating within the system.

The geometry of control is therefore not the geometry of power alone. It is the geometry of knowledge, representation, and the relationship between what can be observed and what can be reached.

Saraksh is a world in which the atmosphere itself is a projection operator, in which the political system runs on a preference-field architecture, in which every repair program generates anomalies it was not designed to see, and in which the most careful long-term intervention is destroyed by a well-intentioned agent operating from an incompatible and less adequate projection.

It is also, in these respects, a reasonably accurate description of every sufficiently complex social system that has ever existed.

The Strugatskys set their novel on an alien planet.

The planet is not alien.

A. A Formal Sketch of Projection Failure

This appendix collects and extends the formal definitions introduced in the body of the essay.

Definition A.1 (State Space)**.** Let Ω be a measurable space representing all possible states of a social system. Elements $\omega \in \Omega$ include institutional configurations, population belief distributions, material conditions, and latent causal variables not directly observable by any agent.

Definition A.2 (Projection Operator)**.** A projection operator is a measurable map $\pi : \Omega \to \Sigma$ from the full state space to an agent's representational space Σ. The fiber $\pi^{-1}(\sigma)$ for $\sigma \in \Sigma$ represents the set of full states that are indistinguishable to the agent.

Definition A.3 (Decision Class)**.** A decision class $\mathcal{D}$ is a set of functions $d : \Omega \to \mathcal{A}$ mapping states to actions. A decision is correctly implemented if the action $d(\omega)$ is chosen at state ω.

Definition A.4 (Admissibility)**.** A projection π is admissible for decision class $\mathcal{D}$ if for every $d \in \mathcal{D}$ and every pair $\omega_1, \omega_2 \in \Omega$ such that $d(\omega_1) \neq d(\omega_2)$, we have $\pi(\omega_1) \neq \pi(\omega_2)$. Equivalently, π is admissible for $\mathcal{D}$ if it does not identify any pair of states that $\mathcal{D}$ requires to be distinguished.

Definition A.5 (Projection Collapse)**.** A projection π exhibits collapse with respect to decision class $\mathcal{D}$ at state ω if there exists $d \in \mathcal{D}$ and $\omega' \in \Omega$ such that $\pi(\omega) = \pi(\omega')$ but $d(\omega) \neq d(\omega')$.

Remark A.6. Projection collapse is the formal counterpart of what the body of the essay calls projection confidence failure. An agent suffering projection collapse will behave as if states ω and ω' are equivalent when they are not, and will therefore select actions appropriate to one when the situation calls for the other.

B. Reachability Partitions and Social Topology

Definition B.1 (Reachability Relation)**.** Given a social system with state space $\mathcal{S}$ and action repertoire $\mathcal{I}$, define the reachability relation $\sim_T$ by: $s \sim_T s'$ if and only if there exists a sequence of interventions $\iota_1, \ldots, \iota_k \in \mathcal{I}$ such that $\iota_k \circ \cdots \circ \iota_1(s)$ places positive probability on a neighborhood of s' within time horizon T.

Definition B.2 (Reachability Partition)**.** The reachability partition of $\mathcal{S}$ under $\sim_T$ is the partition of $\mathcal{S}$ into equivalence classes. Each class is a reachability component: the set of states mutually accessible from one another within horizon T.

Remark B.3. Different agents acting from different projections effectively inhabit different reachability components even when occupying the same physical state. An agent whose projection collapses distinctions necessary to navigate toward a target state cannot reach that state regardless of whether it is formally reachable. The effective reachability component of an agent is therefore the intersection of the formal reachability component with the set of states the agent's projection permits them to navigate toward.

The factions of Saraksh—the Unknown Fathers, the underground, the mutants, the military, the Wanderer—inhabit effectively different reachability components despite occupying the same physical territory. Their political conflicts can be understood as disputes about which reachability component should be expanded, contracted, or dissolved.

C. Repair Operators on Political Systems

Definition C.1 (Repair Operator)**.** A repair operator is a map $\rho : \mathcal{S} \to \Delta(\mathcal{S})$ from states to distributions over states. A repair operator is designed to reduce the measure of some anomaly set $\mathcal{A} \subseteq \mathcal{S}$ as assessed by the agent who designed it.

Definition C.2 (Repair Compatibility)**.** Two repair operators ρ_1 and ρ_2 designed by agents with projections π_1 and π_2 are compatible if the anomaly sets $\mathcal{A}_1$ and $\mathcal{A}_2$ do not contain states that the opposing operator's repair trajectory passes through.

Remark C.3. The repair operators of the major factions in *Prisoners of Power* are pairwise incompatible in this sense. Each operator's repair trajectory passes through states the other operators identify as anomalies. The resulting dynamic is not convergent. Each repair induces new anomalies that motivate further repair by other operators, producing the oscillatory instability that characterizes the novel's political landscape.

Definition C.4 (Repair-Induced Anomaly)**.** Given repair operator ρ targeting anomaly set $\mathcal{A}$ from projection π, define the repair-induced anomaly set as

$$\mathcal{A}^{\rho} = \{s' \in \mathcal{S} : s' \in \operatorname{supp}(\rho(s)) \text{ for some } s \in \mathcal{S}, \text{ and } s' \notin \pi^{-1}(\pi(\mathcal{S} \setminus \mathcal{A}))\}.$$

This is the set of states produced by the repair that were invisible within the projection used to design it.

D. Timeline of Model Construction and Collapse

The following table maps Maxim's successive interpretive frameworks against the events that generate and destroy each one.

Framework	Generating Event	Collapsing Event
Robinson Crusoe / explorer	Crash landing on Saraksh	Encounter with forced labor camp and police
Totalitarian state	Observation of political institutions, surveillance, compulsory ideology	Discovery of broadcast mechanism and its scope
Broadcast control model	Revelation from underground contact	Encounters with mutants, Wanderer; consequences of tower attack
Liberation project	Commitment to destroying Control Center	Wanderer's account of cascade consequences
Post-liberation stabilization	Novel's ending	Remains unresolved; framework not yet constructed

Each row represents a projection that successfully explains the evidence available at the time of its adoption and fails when the next level of hidden structure becomes visible. The table illustrates that the collapses are not random but follow the hierarchical architecture of concealment described in Part I.

References

[1] Arkady Strugatsky and Boris Strugatsky. *Prisoners of Power* [*Obytaemyy Ostrov*]. Originally published in *Neva*, 1969. English translation by Helen Saltz Jacobson. New York: Macmillan, 1977. Second English translation by Andrew Bromfield. Chicago: Chicago Review Press, 2014.

[2] Arkady Strugatsky and Boris Strugatsky. *Hard to Be a God* [*Trudno byt' bogom*]. Moscow: Molodaya Gvardiya, 1964. English translation by Olena Bormashenko. Chicago: Chicago Review Press, 2014.

[3] Arkady Strugatsky and Boris Strugatsky. *Beetle in the Anthill* [*Zhuk v muraveinike*]. Moscow: Molodaya Gvardiya, 1980.

[4] Neil Postman. *Amusing Ourselves to Death: Public Discourse in the Age of Show Business*. New York: Viking Penguin, 1985.

[5] John Taylor Gatto. *Dumbing Us Down: The Hidden Curriculum of Compulsory Schooling*. Gabriola Island: New Society Publishers, 2002.

[6] Norbert Wiener. *Cybernetics: Or Control and Communication in the Animal and the Machine*. Cambridge: MIT Press, 1948.

[7] W. Ross Ashby. *An Introduction to Cybernetics*. London: Chapman and Hall, 1956.

[8] Herbert A. Simon. *The Sciences of the Artificial*. Cambridge: MIT Press, 1969.

[9] Karl Popper. *Conjectures and Refutations: The Growth of Scientific Knowledge*. London: Routledge and Kegan Paul, 1963.

[10] Thomas S. Kuhn. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962.

[11] Michel Foucault. *Discipline and Punish: The Birth of the Prison*. Paris: Gallimard, 1975. English translation by Alan Sheridan. New York: Pantheon Books, 1977.

[12] James C. Scott. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven: Yale University Press, 1998.

[13] Friedrich A. Hayek. "The Use of Knowledge in Society." *American Economic Review* 35, no. 4 (1945): 519–530.

[14] Elinor Ostrom. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press, 1990.

[15] Amartya Sen. *Development as Freedom*. New York: Anchor Books, 1999.

[16] John H. Holland. *Hidden Order: How Adaptation Builds Complexity*. Reading: Addison-Wesley, 1995.

[17] Stuart A. Kauffman. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford: Oxford University Press, 1993.

[18] Judea Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press, 2000.

[19] David Lewis. *Counterfactuals*. Cambridge: Harvard University Press, 1973.

[20] Robert Stalnaker. "A Theory of Conditionals." In *Studies in Logical Theory*, edited by Nicholas Rescher, 98–112. Oxford: Blackwell, 1968.

[21] Flyxion. *Axioms for a Falling Universe: A Unified Field Theory of the Relativistic Scalar–Vector Plenum*. Independent Researcher, May 2026. Available at: `https://standardgalactic.github.io/alphabet/projects/axioms-for-a-falling-universe.pdf`

[22] Flyxion. *HYDRA: Architecture, Geometry, and the Problem of Coherent Agency*. Independent Researcher, 2026. Available at: `https://standardgalactic.github.io/playfloor/epistemology/hydra_architecture.pdf`

[23] Flyxion. *Hidden Curvature: Constraint Geometry, Projection, and the Reconstruction of Meaning*. Independent Researcher, 2026. Available at: `https://standardgalactic.github.io/playfloor/admissibility/admissibility-field.pdf`

[24] Flyxion. *Repair as a Fundamental Category: Toward an Ontology of Restoration*. Independent Researcher, 2026. Available at: `https://standardgalactic.github.io/alphabet/research/repair_as_fundamental.pdf`