Title
Kant, Autonomy, and the Self-Organizing Cosmos

ISBN
9798186178675

Author
Spano, Christopher

Publication Date
2026-05-27

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Kant, Autonomy, and the Self-Organizing Cosmos

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Philosophy

by

Christopher Spano

June 2026

Dissertation Committee:
Dr. Erich Reck, Co-Chairperson.
Dr. Alexandra Newton, Co-Chairperson.
Dr. Andrews Reath
Dr. Bai-Lian Li

The Dissertation of Christopher Spano is approved:

Committee Co-Chairperson

Committee Co-Chairperson

University of California, Riverside

Acknowledgements

I am deeply grateful to my family and family for their support throughout this journey.

I owe special thanks to Pierre Keller, whose guidance helped shape this work. I also honor the memory of Joseph Margulis, my undergraduate mentor, whose inspiration and passion for philosophy stays with me.

I also would like to thank my current co-chairs, Erich Reck and Alexandra Newton, for their steady mentorship, and to Andrews Reath for coming to the project late and giving insightful advice. Finally, I thank Larry Li, whose assistance helped foster my obsession with plants, ecology and complex systems science.

ABSTRACT OF THE DISSERTATION

Kant, Autonomy, and the Self-Organizing Cosmos

by

Christopher Spano

Doctorate of Philosophy, Graduate Program in Philosophy
University of California, Riverside, June 2026
Dr. Erich Reck, Co-Chairperson
Dr. Alexandra Newton, Co-Chairperson

This dissertation resolves the apparent conflict between human autonomy and nature's self-organization—a tension rendered urgent by complex systems science. Drawing on Kant's "Ideas" (normative frameworks that bi-constitute subjects and their purposive worlds), I argue that nature's self-organizing character follows from the Idea of autonomy itself. Because the highest good requires a purposively unified cosmos, practical reason compels us to postulate nature as self-organizing. Theoretical reason can never establish this; practical reason affirms it as a necessary condition of moral agency.

I then demonstrate that for Kant, judging organisms, ecosystems, or even convection cells as self-organizing agents involves treating them as if guided by an immanent Idea following a four-part schema: coherent identity, reciprocal causality between parts and whole, selective engagement with an environment, and a coherency test whereby success reinforces or failure revises the agent's organizational form. Science

since Kant has vindicated and articulated this schema. Through Ernst Cassirer, I show this bi-constitutive logic undergirds 20th century science, but also extends across culture. I then provide a critique of complex systems science, examining how graph theory, dynamical systems theory, criticality, and panarchy do not eliminate teleological judgment but rather presuppose and articulate Kant's schema further.

The dissertation culminates in a Kantian environmental ethics. Against readers of Kant that can at best establish indirect duties to nature, I argue that the highest good—the ultimate end of our moral vocation—constitutively requires a purposively unified natural world, a maxim that degrades nature's self-organization yields a contradiction in will, generating a direct but imperfect duty to preserve the multi-scalar architecture of natural agency. This duty is internal to the Formula of Humanity: to will the permissible happiness of embodied others is to will a self-organizing cosmos. We are thus obligated to harmonize human autonomy with the integrity of the self-organizing cosmos.

Table of Contents

Figures and Tables

Preface

We stand at a unique point in history in understanding our place in the universe. Ever since Kant, we have understood the bedrock of our human dignity to be our capacity for autonomous action. We can conduct our own lives by freeing ourselves from what others want us to do and rejecting impulses that tell us to do things that we do not identify with. We are, seemingly, free insofar as we can free ourselves from what our contingent desires or our environmental/social circumstances push us to do. The principles that emerge from neither source seem to be authentic moral principles, as what we 'ought' to do does not emerge from the contingencies of what 'is.' Therefore, to be truly free is to be moral, in community with other free individuals, where our actions are guided by responsibility for the other and for the common world we inhabit.

Simultaneously, however, a revolutionary shift is occurring in our scientific vision of nature. From the formation of galaxies to our ecosystem, leading scientists argue that the cosmos is not a dead machine, but a spontaneous, creative process. Order is not imposed from above, nor is it merely a product of linear, billiard-ball style collisions. The universe is *self-organizing* and active, almost as if it were itself a self-legislating, autonomous agent.

This collision creates the drama I would like to unfold in this dissertation. Is our freedom a continuation of this cosmic self-organizing process? Or is this cosmic self-organization merely an illusion, projecting the structure of our own freedom onto nature?

This dissertation proposes that the self-organization of autonomous human agency mediated by *reason* and the self-organization observed in natural systems share a common structure. Here, “reason” is understood not as a detached mental faculty, but as the social practice of giving, asking for, and responding to justifications for our actions. Reason is central to our agency, to what we *do*. What unites every action is a particular reason for doing it, even if such reason is unconscious: we are walking *because* we are going to the grocery store; we are listening to music *because* we want to unwind; our heart beats faster *because* we fear giving a speech.

Similarly, attributing a reason to something is central for unifying not only our own actions, but the actions of any agent. We are familiar with this process: we see everyday objects, like chairs, as unified wholes rather than mere conglomerates of wood. In the same way, we perceive human beings, organisms, galaxies, and hurricanes as belonging to organized wholes, defined by their capacity to do something. Human beings reason; organisms metabolize; galaxies gravitate; hurricanes spiral. We identify these entities by attributing agency to them: they appear to have a reason to exist, which is to do something. Organisms, chairs, and hurricanes are unified because they function toward an end, even if their ultimate end is merely self-persistence. These non-human agents are unaware of the reasons that unify their actions. Nevertheless, there is a striking structural similarity between reasoning human agents and self-organizing systems: they are organized by a telos—a “reason to do something.” Yet again, however, a critical question remains: is this similarity merely a projection of our own agency onto the non-human world? How can we know that they are truly agents?

To answer this, we must unpack what makes our own agency unique. Our reasoning is a practice that contains norms for what counts as a ‘good’ reason. Plato and

Kant called these normative contexts of significance 'Ideas.' Justice, Truth, and the Good are not empty abstractions, but guides for action. These Ideas function as self-imposed goals, directing what we ought to do and think. They are the very conditions that make coherent thought and action possible, as they generate what counts as a 'good' reason and a 'good' action. In this sense, Ideas guide our practices, forming the normative horizon within which those practices become intelligible. Ultimately, by establishing this horizon of intelligibility, Ideas are what allocate meaning to everything we do.

Ideas are not mental constructs. As normative contexts of significance, Ideas serve to give meaning to *both* our inner mental lives and to external nature itself. Our mental lives are governed by Ideas just as much as nature is. Take, for instance, the Idea of science: it not only constitutes us as scientists—by guiding our methods, aims, and standards of justification—but also constitutes nature as a rationally intelligible and law-governed whole, describable through various theories. This simultaneous process, in which our Ideas shape the understander and what is understood, the agent and what is acted upon, is what I call the process of 'bi-constitution' that forms the backbone of human agency.

We can never know these normative contexts (the Ideas) entirely because they constitute the very framework of our understanding, rather than being objects within it. They guide our practices but always escape explicit articulation. They prompt an endless process of articulation, that changes when our historical circumstances change (such as the discovery of systems science). However, this is not blind relativism: the truth of the Ideas can either be grounded in necessary postulates that make action possible, or *felt* when we manage to act successfully when guided by them. The Idea that nature is a unified whole, for example, cannot be known from our humble perspective in the cosmos. Yet this is a necessary Idea since it *guides* interdisciplinary work within the sciences, connecting disparate theories and sciences. When we connect two frameworks—say, biology and chemistry—we can feel the truth of this Idea: it seems like the universe really is a unified rational whole.

Yet we are not only guided by Ideas—we are also remarkably adept at judging things in nature *as if* they, too, were guided by them. To perceive a forest, a cell, or a hurricane as self-organizing is to see it *as if* guided by an immanent 'Idea'—a principle of wholeness that coordinates its internal parts and its external environment to maintain its existence. The tree's form organizes its roots and leaves while rendering soil and sunlight purposive for its life. To judge these things as guided by Ideas is to claim that they are wholes that do something; they are agents. Without this judgment, knowledge would be impossible. We would never be able to see entities in nature *as* united wholes, but only *as* insignificant hunks of matter. Yet we can never *know* whether these wholes are truly agents. Rather, we can only *feel* the success of such attribution when we are able to do science successfully with this attribution. This feeling is not mere desire satisfaction, but the feeling of a harmony between our own cognitive faculties and the cosmos.[1] Therefore, we can apprehend the truth of this attribution pragmatically—through successful action—without ever possessing certain knowledge of whether the entity is 'really' an agent or not. In this framework, truth is not a static theoretical correspondence between

[1] This distinction will be developed further in chapter 2

mind and world, but the moment a normative "ought" (our guiding judgment or Idea) leads to successful action in the world of what "is."

Thus, the truth of the pan-agential view of nature is established, in part, by pragmatic success. Our era is witnessing the most compelling demonstration of this very principle, as contemporary science achieves its overwhelming successes precisely by attributing self-organizing agency to the cosmos. So-called 'complex systems' science uses mathematical modeling tools to capture self-organizing agency from quantum entanglements to convection cells, organisms, ecosystems, and beyond. The proven success of these insights has shown what many have believed for millennia: the universe is no passive combination of atoms, but a dynamic agent.

If we are justified in judging nature *as if* it were composed of self-organizing agents across all scales—if the pragmatic success of complex systems science warrants a feeling that cells, forests, and climates genuinely share the same basic form of purposive self-organization that we recognize in ourselves—then we must confront our own position within this purposive order. We are not spectators standing outside nature, occasionally deploying the agential lens for scientific convenience. We are ourselves self-organizing agents, constituted by the same logic of reciprocal constraint that we discern in the natural world. Our own agency, guided by Ideas that unify our actions and orient us toward ends, is a paradigmatic instance of a broader phenomenon, where cosmos presents itself as a dynamic ecology of purposive agents. This realization is not merely scientific or metaphysical. It is, at its core, an ethical revolution. For if nature is a community of agents rather than a machine of inert parts, then our relation to it cannot be one of mere utility or detached observation. We stand within a web of purposive beings, and that standing generates obligations. Understanding how we fit in nature, therefore, is not the ultimate goal of this piece. The real question is: assuming this pan-agential lens is true, how ought we to position ourselves in the cosmos? How ought we to self-organize? Are we not obligated not just to other humans, but to the self-organizing cosmos as such?

I argue we are. Our duty transcends the prudent use of nature and the cultivation of our own benevolent dispositions; it demands that we integrate our actions with nature's foundational self-organization—a harmony Kant called the highest good. This ideal is the state in which human freedom aligns with nature's own integrity, where the kingdom of nature and the kingdom of ends are united in a single moral world. Consequently, we incur a direct obligation to preserve and promote the world-constituting activities of all natural agents, from microbes to climates. These agents are not mere resources to be managed or obstacles to be overcome. They are participants in the same purposive order that we participate in, their functional integrity constitutive of the moral cosmos we are called to realize. To will the highest good is to will a world in which self-organizing nature and our autonomous humanity are harmoniously integrated. We cannot will this end while simultaneously willing the disintegration of the natural systems that make it possible. The obligation to respect the purposive unity of nature is therefore not an addendum to Kantian ethics, but a direct implication of its most fundamental commitments.

This dissertation aims to unpack the preceding claims by answering three integrated questions:

1. How do we 'self-organize' according to autonomy/freedom?
2. How does nature self-organize?
3. Synthesizing 2 and 3: How, then ought we to live?

Through a reading of Kant, I will answer all these questions. The first chapter will unpack in more detail our historical and scientific situation, motivating the need to investigate Kant. I will then proceed in chapter 2 to demonstrate Kant's theory of human agency, introducing the 'bi-constitutive' function of Ideas to both give meaning to our internal lives and the external world. In chapter 3, I argue how to see agents 'as if' they are Ideas is to attribute a certain bi-constitutive structure to them: to judge something to be a self-organizing agent is to judge it simultaneously as organizing an internal structure (e.g., its organs) and an external environment. Chapter 4 exposes how such insights were becoming known to neo-Kantian Ernst Cassirer throughout the early 20th century in all the sciences. Chapter 5 argues how contemporary complexity sciences prove to be a significant marker of success of judging nature to be agential, pointing to the truth of the agential thesis. Finally, chapter 6 argues for a Kantian inspired environmental ethic, where our ultimate calling is to create a moral world where the self-organization of nature and our human autonomy are united. This dissertation moves from a stance of awe to one of obligation: awe for our autonomous capacity to engage with Ideas and wonder at the self-organizing, agential cosmos, and obligation to realize a democratic kingdom where that autonomy aligns with the integrity of the natural world.

Chapter 1: The Problem of Self-Organization

I. Introduction

We live in the century of complexity. Scientists across physics, biology, ecology, and social science argue that order in the universe is not pre-set or imposed from above, but emerges spontaneously. From ecosystems to the very laws of physics, organization arises bottom-up from decentralized interactions, creating configured wholes. Crucially, these emergent wholes are not merely sums of their parts, but exert top-down influence, organizing and constraining the very interactions that produced them. Scientists now describe a world that is intrinsically self-organizing, from branching patterns in rivers and neurons to adaptive dynamics in ecosystems and economies. Systems generate their own form through local interactions, producing global stability and structure without external control. A school of fish, for instance, emerges autonomously via rules among individuals, yet is constrained by physics, predators, and motion. In its most radical implication, these scientists suggest that even the laws of physics may be understood as stable, emergent products of deeper self-organizing processes. If true, this would mark a fundamental shift in our understanding of nature—from a reductionist world of passive parts obeying externally imposed laws to an active, creative, and self-constructing cosmos that constructs its *own* laws.

When taken to its logical conclusion, however, this view leads to a deep philosophical tension. The very practices we use to discover self-organization—our scientific methods, theoretical paradigms, and communal reasoning—seem themselves to be complex, adaptive, and historically evolving. They, too, exhibit the hallmarks of complex, self-organizing systems. We therefore find ourselves in a puzzling loop: We are using self-organizing minds to study a self-organizing world. This loop forces a skeptical question upon us: How do we know that the patterns we "discover" in nature are not merely projections of our own cognitive and social organization? Are we finding the world's complexity, or are we imposing our own?

This chapter has two aims. First, in Sections II and III, I will elucidate the rise and appeal of the complexity sciences. Section II shows how self-organization is understood as a process of reciprocal creation between parts and wholes, and introduces the mathematical tools—from network theory to panarchy—that allow scientists to model it. Section III then documents a parallel "practical turn" in the philosophy of science, which reveals that scientific practice itself is a paradigmatic example of a self-organizing system. I argue that the conjunction of these developments—a self-organizing nature and self-organizing scientific practices—has prompted a quiet return to the problems that Immanuel Kant faced. Kant, after all, coined the term "self-organization" and situated his analysis of organisms within a broader investigation of the normative structure of reason.

Second, in Section IV, I will argue that the mutual interest in self-organizing scientific practice and self-organizing nature creates a philosophical impasse. When both the knower and the known are understood as self-organizing, we are left with two unsatisfying options: complexity realism, which tries to ground the mind's order in the world's order, and complexity idealism, which tries to ground the world's order in the mind's order. I show that although each has positive insights, both are insufficient—the former because it undermines the autonomy of our scientific practices, the latter because it cannot secure the reality of nature's self-organization. Their shared failure lies in assuming unidirectional explanation: that order must flow from mind to world or world to mind.

The inadequacies of these positions suggest thinking outside the mind/world dichotomy. I conclude by claiming that Kant has the resources and tools to explain the conjunction of a self-organizing mind and a self-organizing world. Both, for Kant, are mediated by the practice of giving and responding to reasons, which cannot be reduced to the mental or the physical. Instead, reason brings normative standards to how we ought to think, feel, and act, while also unveiling the structure of the external world. Both the organization of the mind and the organization of the world are mediated by a singular structure: reason. The structure of reason and how it articulates the self-organizing structure of nature will be the focus of Chapters 2 and 3.

II. The Appeal of Self-organization to Scientists

A. New Vision of Order

Before examining the technical details of this new scientific understanding of the cosmos (see also, chapter 5), let us first grasp the intuition that guides it. For centuries, the dominant scientific story described a universe of passive parts, moved by external forces that obey timeless laws. Order was merely the sum of countless random interactions playing out within a rigid, unchanging framework of rules. Yet contemporary biologists, physicists, and economists are telling a fundamentally different story. They propose that much of the order we see in nature does not simply unfold from pre-existing rules imposed on inert matter. Instead, order generates and sustains itself from within, continuously shaping both its own form and the very conditions that make it possible. This is the phenomena called 'self-organization' (Kauffman, 1993; Thompson, 2007).

Imagine a flock of seagulls. There is no leader or alpha bird. Each bird follows simple, local rules where they, e.g., align with their neighbors, avoid collisions. From these decentralized rules, the coordinated pattern of the flock emerges, as if the flock was governed by a singular law (Reynolds, 1987; Camazine et al., 2003). These scientists propose that complex global order can arise spontaneously from simple local interactions.

But the deeper insight lies in what happens next. The emergent flock is not just an outcome, but becomes a new, active entity. Its collective motion changes the airflow, the perceived threats, and the navigational options for every bird within it. The whole now shapes the behavior of the parts that created it. Instead of merely reacting to its neighbor, the bird is now flying within a coherent form that guides its possibilities. The parts create the whole, and the whole, in turn, governs the parts.

This logic of reciprocal creation—a loop between bottom-up emergence and top-down constraint—is the signature quality of self-organization. It appears nearly everywhere throughout the cosmos, from the creation of the universe to the coordination of human social systems. Our immune system, for example, conducts itself not from a central command center, but from countless cells communicating locally (Maturana & Varela, 1980). Yet the resulting immune response is a coherent whole that then regulates the very cells that produced it. A forest emerges from the interactions of soil, water, plants, and animals (Holling, 2001; Gunderson & Holling, 2002). Once established, the forest's own microclimate and nutrient cycles create the conditions that determine what can live and thrive there (especially itself and its offspring) for generations.

This is a radical departure from the reductive metaphor. It reveals a world where order is not stamped onto passive matter but arises from an ongoing, circular feedback loop. The "system" and its "environment," the "whole" and its "parts," are not separate (Kauffman, 1993; Thompson, 2007). They are mutually constituting elements in a single, dynamic process of self-creation. This is not just a new scientific fact but constitutes a new framework for seeing the universe as intrinsically creative, adaptive, and self-constructing across scales (Gunderson & Holling, 2002; Juarrero, 2023).

B. An Emerging Science of Emergence

To move from the intuitive recognition of self-organization to a rigorous science requires a new kind of mathematics. Traditional calculus, optimized for describing trajectories of singular objects or aggregate averages, struggles with the decentralized, interactive logic of complex systems. In response, complexity science has integrated an array of complementary tools that together model the structure and behavior of self-organizing wholes. These tools are not exhaustive but are the most representative of the paradigm of contemporary complexity science (for a more detailed exposition, see Chapter 5).

1. Mapping the Network of Interactions: Graph and Network Theory.

If order emerges from interactions, we need a way to map what is interacting with what. Graph and network theory provide this map. Imagine drawing a diagram where each component—a neuron, a species in a forest, a person in a community—is a dot (a node). Then, draw a line (an edge) between two dots if they influence each other. This picture, a graph, is used to model systems and to unpack their architecture (Barabási, 2016). It lets researchers move beyond simply saying a system is "interconnected" to analyzing the exact pattern of connections. They ask: Are there feedback loops? Are there tightly-knit clusters? Are a few components central hubs, with many lines converging on them, while most are only loosely connected? This map of relationships is the blueprint for understanding how influence and constraint flow through a system, with network theory adding statistical tools to characterize these overall patterns (Newman, 2006).

2. Simulating Behavior: Dynamical Systems Theory.

A map is static. To understand the system's dynamics, we need to be able to model the rules its development. How does each component behave, and how is that behavior changed by its neighbors? Dynamical systems theory provides the language for these rules, using precise equations to describe how the state of each component evolves over time (Strogatz, 2015). For example, a node might represent a predator population. An equation can describe how its numbers naturally grow or shrink, and how that change is amplified or dampened by the numbers of its prey (its connected neighbors). This framework introduces the idea of attractors, which are the stable end-states that the system tends to settle into, like a pendulum coming to rest or an ecosystem reaching a balance. The set of all starting points that lead to a particular attractor is its basin of attraction. This allows scientists to simulate the system's trajectory, exploring not just what it is, but what it could become under different conditions.

3. The Optimal State for Adaptation: Criticality.

A striking discovery is that the most adaptive systems—brains, immune systems, thriving ecosystems—are not maximally stable nor completely chaotic. They operate in a special, poised state called criticality, often described as the "edge of chaos" (Langton, 1990; Kauffman, 1993). Critical moments are those such as when an icy slush is neither fully liquid nor solid, or when a sandpile grows so steep that the next grain could trigger either a minor slip or a catastrophic avalanche (Bak, Tang, & Wiesenfeld, 1987). In this state, the system exhibits fractal patterns, where the same types of fluctuations and correlations appear at every scale, from the very small to the very large. This unique architecture makes the system sensitive enough that a small, local event (a single thought, a new species, a market signal) can trigger a cascade of reorganization throughout the entire network. This sensitivity—while catastrophic for most systems that enter it—is an ideal place to 'hover' near for evolutionary success. For example, the sensitivity garnered near criticality is ideal for learning, innovation, and adaptation. Too close to criticality, however, leads to complete destabilization and collapse of the system (Chialvo, 2010).

4. Orchestrating Across Scales: Panarchy and Resilience.

Criticality describes the optimal *internal* state of an adaptive system. But such a system does not exist in isolation. It is embedded within, and composed of, other systems at different scales. A cell is part of an organ, which is part of an organism, which is part of an ecosystem. How do these nested levels influence each other? The theory of panarchy models these cross-scale interactions (Gunderson & Holling, 2002). It views systems as sets of interconnected sub-systems, operating at different speeds and sizes. Fast, small cycles (like a seasonal insect bloom) can "revolt," injecting novelty and disruption that can revitalize or destabilize a larger, slower system (like a forest). Conversely, large, slow systems (like a long-term climate pattern) "remember" established configurations and provide stability, guiding the recovery and reorganization of smaller systems after a disturbance (Holling, 2001). A system's overall resilience—its ability to absorb shocks and adapt—is not a property of one level but emerges from this continuous exchange across all scales. This framework is essential for understanding the conditions that sustain health across scales, from the individual organism to the ecosystem.

Together, these theoretical edifices transform a vague intuition into a determinate science. They provide the tools to model and predict the future of self-organizing systems. But do these tools represent merely a new set of techniques, or do they necessitate a fundamental shift in our scientific worldview?

C. The Paradigm Shift: From Reductionism to a Science of Complexity

For proponents of the complexity paradigm—which emerged in the mid-20th century—the answer is clear: we are experiencing a revolution. Summarizing the sentiment of the scientific community, Stephen Hawking (2000) claimed "the next century [the 21st] will be the century of complexity." Hawking was not wrong. Although they might not all be consciously aware of it, it seems as if science in all domains—from the physical sciences to the humanities to computer and data sciences— are undergoing a subtle revolution. Systems scientists are with great success applying the wide-ranging scope of their models to economics, anthropology, thermodynamics, chemistry, neuroscience, ecology and even the study of scientific practice itself. According to M. Mitchell Waldrop, these scientists "all share the vision of an underlying unity, a common theoretical framework for complexity that would illuminate nature and humankind alike. They believe that they have in hand the mathematical tools to create such a framework" (Waldrop, 1992, 13). Ilya Prigogine and Isabelle Stengers call it a "conceptual revolution" towards the "multiple, the temporal, and the complex" (Prigogine & Stengers, 1984, xxvii-xxviii). The aim of science—once defined as reducing all phenomena to universal, linear causal laws—now appears not as the whole story of nature, but as a special, simplified case within a deeper, nonlinear, and self-creating reality. This shift in our fundamental understanding of nature is why Cliff Hooker calls it "a recent, ongoing, and profound revolution" (Hooker, 2011, 3).[2]

Yet revolutions in science, as Thomas Kuhn famously argued, do not arrive fully formed as abstract theories. They begin with notable achievements—what Kuhn called "exemplars"—that demonstrate a new way of seeing and solving problems, attracting adherents who then extend and refine the approach (Kuhn, 1962). The complexity revolution follows this pattern. Early on, there was no unified "theory of complexity." Instead, there were striking models that revealed order arising spontaneously in systems previously thought to be merely random or reducible.

One such exemplar was the Belousov–Zhabotinsky reaction, discovered in the 1950s. When certain chemicals are mixed in a shallow dish, they do not settle into a uniform state but spontaneously form propagating visible, macroscopic order emerging from local molecular interactions with no external template (Belousov, 1959; Zhabotinsky, 1964). This was deeply puzzling under the prevailing paradigm, which expected such systems to trend toward thermodynamic equilibrium. Yet here was a chemical system that sustained itself far from equilibrium, exhibiting coherent, periodic structure. It became a touchstone for researchers like Ilya Prigogine, whose Nobel Prize-

[2] As Mitchell puts the situation, to self-understand the complexity approach, one must ask, "How do we move beyond the traditional paradigm of reductionism toward a new understanding of seemingly irreducibly complex systems?"(Mitchell, 2009, xi).

winning work on dissipative structures formalized how order can arise from energy flow in open systems (Prigogine & Nicolis, 1967).

Another foundational exemplar emerged from theoretical biology. Stuart Kauffman's NK models, developed in the 1960s and 1970s, demonstrated how networks of interacting genes could spontaneously organize into ordered regimes without centralized control (Kauffman, 1969). These models showed that even random networks, when structured with the right connectivity, self-organize into attractors that resemble distinct cell types, suggesting that order in living systems need not be meticulously designed by natural selection alone but could arise from the dynamics of interaction itself.

Crucially, neither the Belousov–Zhabotinsky reaction nor Kauffman's NK models began as applications of a pre-existing complexity theory. They were puzzles that resisted reductionist explanation. Their resolution required new concepts—far-from-equilibrium thermodynamics, self-organization, attractor landscapes—and new methods, such as computer simulation and nonlinear dynamics. Over time, as more such exemplars accumulated across chemistry, biology, physics, and eventually economics and ecology, a paradigm consolidated. What began as isolated anomalies became a coherent research tradition, unified not by a single theory but by a shared family of models, techniques, and metaphysical commitments. The mathematical tools described below—network theory, dynamical systems, criticality, panarchy—are the outcomes of this trajectory.

As we have seen, the complexity revolution is defined less by an interlocking set of theories and more by a series of remarkably successful mathematical modelling techniques. Yet scientists struck by their pervasiveness take them to be more than useful tools, but as illuminating a new paradigm for understanding nature. Through this lens, the world appears as a dynamic orchestration of creativity, where order and harmony are forged against entropy's slow dissipation. Yet a fundamental question remains: do these techniques ground a truly revolutionary cosmology, or are they simply powerful instruments for investigating nature? To answer this, we must first understand the core concepts of the paradigm itself.

Central to this paradigm is the notion of 'complexity.' A 'complex system' is a technical term. Systems are sets of interacting components that unify to form a whole. While some 'simple' systems may be composed of parts that can be 'added' to form the whole, systems that are complex are driven by internal dynamics and feedback loops that form— as Melanie Mitchell uses the well-worn adage—a "whole [that] is more than the sum of its parts" (2009, x). These structures generate themselves spontaneously without any external control, as it would if it were 'designed' by some creator (e.g., God, humans). Such systems usually (but not always) possess the following properties: they are open, exchanging information, energy, and matter with their environment; they are resilient because they can adapt and adjust to external disturbances; they are 'autonomous'[3] in that they self-generate properties for themselves without any intervening external factor determining these properties; and they behave nonlinearly, having an output not directly proportionate to the input as it would be in a linear system (Anderson, 1972; Cross & Hohenberg, 1993; Eigen, 1971; Haken, 2006; S. A. Kauffman, 1993).

[3] Not in the robust sense of human autonomy which I will unpack in chapter 2

For any complexity researcher, one of the greatest points of pride is not necessarily breaking free from the constraints of reductionism entirely, but rather embedding the explanatory power of reductionism within a broader, multi-level understanding of reality. Most forms of reductionism share three key tenets: 1) complicated systems can be fully explained by their constituent parts and the fundamental laws governing them; 2) natural laws are universal, admitting of no exceptions; 3) science is 'unified' based on the potential to reduce all sciences to physics. Complexity scientists do not wholly reject these tenets so much as they reframe their scope and significance.

Complexity scientists challenge the first tenet: they do not view complex systems as fully explained by reducing them to their component parts. Instead, they regard emergent properties—where higher-level behaviors arise from the interactions of lower-level components in a way that cannot be deduced or predicted from those components alone—as, in the words of William Wimsatt, "the rule, rather than the exception" in nature (Wimsatt, 2007, 302). In this view, reducing a system to its constituent parts remains essential for a simplified description, but it is insufficient for a complete explanation of the system's behavior. A full explanation requires tools that go beyond reductionism—namely, mathematical models that can adequately capture the dynamics of emergence. Although these emergent processes share no single, universal, mechanism, patterns of self-organization can be observed, modeled, and understood across vastly different scales, from the flow of thermal fluids and the formation of complex molecules to the rise of life and the dynamics of human social structures.

Complexity scientists also thereby challenge the second tenet. Under the reductionist paradigm, scientific laws are typically necessary (if conditions X hold, then Y must follow) and universal (laws are not context-dependent, applying across scales and domains). Complexity scientists still believe in laws, although they are no longer necessary nor universal. Laws are rather contingent (context-dependent) and non-universal (emergent only at certain scales).[4] Self-organizing systems exhibit order without external or pre-existing laws. Outcomes (such as in evolutionary biology or cosmology) depend on path-dependence and chance events. Even 'classical' necessary and universal laws (such as Newton's laws or the laws of general relativity) are conceived of as arising out of self-organizing processes (Laughlin, 2005; Mersini-Houghton, 2022; Verlinde, 2006).

Finally, complexity scientists challenge the third tenet of reductionism. While complexity scientists reject the idea that physics is the ultimate, privileged foundation for all sciences, they do not conclude that the sciences are therefore disunified. Instead, they are seen as studying different aspects and scales of a singular, self-organizing cosmos. According to Stuart Kauffman, organisms, ecosystems, economies, and societies all have "self-organization" as "the root source of order" (1995, 2). This provides a shared ontological ground. Consequently, multiple explanatory frameworks—from physics and chemistry to ethology and ecology—can be coherently applied to the same phenomenon, such as an organism. Each field examines a specific level of organization, revealing how bottom-up processes (e.g., as studied in biochemistry) and top-down constraints (e.g., as

[4] We will see that Kant has a different understanding of universality—not as exceptionless rules, but as universally communicable rules in a singular ideal self-consciousness

studied in ecology) interact. This scientific pluralism is made coherent by an underlying unity, not of fundamental particles, but of dynamically evolving forms and processes. The complex systems approach thus unifies the sciences through a common conceptual and mathematical toolkit for modeling emergence, recognizing that different perspectives are all valid views of a dynamic, self-creating reality.

The most attractive feature of the complexity paradigm over the reductionist paradigm is its explanatory scope. Reductionism alone proves insufficient to explain a wide range of complex phenomena for which the toolkit of complexity science offers compelling accounts. As Melanie Mitchell enumerates, these include:

> the seemingly irreducible unpredictability of weather and climate; the intricacies and adaptive nature of living organisms and the diseases that threaten them; the economic, political, and cultural behavior of societies; the growth and effects of modern technology and communications networks; and the nature of intelligence and the prospect for creating it in computers (Mitchell, 2009, x).

In seeking to explain emergent phenomena by reducing them exclusively to the properties of local interactions, the reductionist approach, as Mazzocchi (2008) argues, inevitably struggles. It lacks the conceptual and formal resources to capture the higher-order organization, feedback loops, and adaptive dynamics that define such systems.

It should be implied from the foregoing that although the complexity approach is non-reductionist, it is not *anti*-reductionist. Any non-reductionist paradigm would be severely mistaken if it ignored all the successes of the reductionist paradigm. The complexity approach embeds local, lower level, 'reductive' explanations within a broader, more comprehensive framework. This allows for non-reductive explanations to co-exist with reductive explanations. A complexity scientist does not ignore the parts. Rather, they are required to shift perspectives from whole to parts and back again (Morin, 1990).

Consider a simple cause-and-effect interaction: a billiard ball strikes another billiard ball. These types of interactions can be explained very easily within the reductionist paradigm. In the case of the billiard balls, the interaction can be fully described by fundamental physical laws, such as Newton's laws of motion, the conservation of momentum, the behavior of atoms at the point of contact, and the relevant forces, velocities, and masses involved.

A non-reductive explanation, however, can still draw upon explanations using these resources. However, this would not exhaust the explanation. Additionally, it would deem relevant higher-order systems in which the billiard balls are embedded. For example, if the billiard ball was struck during a game of pool, then the rules of the game of billiards, the player's intention, and the current context of the game all are relevant for explanations as to why the billiard balls moved in a particular way. Human culture and the games we play are just other kinds of self-organizing systems on different scales. For complexity scientists, it is unproblematic for an explanation to move from the 'whole' (the game of billiards) back to the 'parts' (the interaction on the physical level) and back to the whole to fully understand what is happening when the two balls collide.

Even the fundamental assumptions used by the reductive approach can be understood as having emerged from self-organizing interactions. For instance, the complexity approach attempts to understand how Newton's laws emerge from more complex interactions. In the case of the billiard balls, self-organizing systems of mass and energy warp spacetime, influencing gravitational forces such that they can be locally approximated by Newtonian mechanics. Going a step deeper, it is even possible that the very fabric of spacetime, as described in general relativity, could emerge from self-organizing quantum entanglements (Mersini-Houghton, 2022). From a complex systems perspective, non-linear, self-organizing systems (like a game of pool or the structure of spacetime) can encompass local interactions that are approximately linear and governed by simple laws. However, it is crucial to remember that these (approximately) linear mechanisms and laws are organized and made possible by the patterns of an overarching, self-organizing non-linear system.[5]

The vast explanatory scope—combining explanations of emergent properties with reductive explanations which are used pragmatically for their methodological simplicity—makes the complex systems paradigm extraordinarily attractive. Yet this paradigm is not perfect. It still faces challenges in terms of predictability. Since the systems it studies are typically nonlinear, where small inputs can lead to massive outputs (as seen, for example, in the so-called 'butterfly effect'), predicting the course of such systems is extraordinarily difficult. This is known as the 'sensitive dependence on initial conditions' (Mitchell, 2009, 20ff.). If judged based on the merits of the reductionist paradigm, this method is a massive failure: it cannot make accurate predictions. How then can we know if it is successful?

For complex systems scientists, we must remember that certain systems can be simplified for the sake of prediction. Hence, it achieves the same predictive success as the reductionist paradigm. Yet for those complex systems that cannot be simplified, the

[5] The exception to this claim, as far as we know, is the conservation laws, which are linear.

complexity approach still offers other criteria of success, namely, crucial qualitative insights into their processes and patterns. This enables, for instance, complexity scientists to explore properties of systems, such as critical nodes in the system, as well as their resilience and their adaptability to new disturbances and circumstances (Holling, 2010; Holling & Gunderson, 2002). By modeling the complexity of phenomena with various computational tools, we can hypothesize the conditions under which a system may become more resilient or less adaptable to specific disturbances. This makes complex systems not merely descriptive but also prescriptive, equipping us with the tools to guide city-planning, policies that enhance the resilience of ecosystems, and monetary strategies that help prevent economic collapse (see e.g., Berkes & Folke, 1998; Pickett et al., 2004; Wu et al., 2020).

III. Self-organization in biology and in science: the Kantian Turn

The rise of the complexity paradigm is not merely a technical shift within the sciences but represents a profound challenge to our basic philosophical categories of cause, law, purpose, and explanation. How are we to understand a world that appears intrinsically creative and self-ordering, yet does so without a designer? And if our own scientific practices are dynamic, adaptive, and historically evolving, what does this imply about the status of the knowledge they produce? In brief, how do we reconcile the self-ordering of our human cognitive practices (like science) and the self-organization of nature?

Faced with these questions, a peculiar trend has emerged in 21st-century philosophy: a return to the problems that Immanuel Kant faced. Kant is uniquely relevant because he stood at a similar crossroads. He sought to reconcile a critical analysis of the practice of giving and responding to reasons with the sciences of his time. Simultaneously, he coined the term "self-organization" (*Selbstorganisation*) to articulate a non-reductive yet naturalistic understanding of natural agents. This dual legacy—Kant as theorist of reason and Kant as theorist of self-organization—has prompted renewed engagement across two fronts. On one hand, philosophers of biology are re-engaging with Kant's concept of self-organization to make sense of emergence, teleology, and natural agency. On the other hand, philosophers of science, influenced by the "practical turn," are examining the norms, interests, and social dynamics that structure scientific inquiry—a move that echoes Kant's focus on the normative governance of reason.

This section will survey this dual "Kantian turn." Although most complexity scientists are not Kantian, they move in directions that are compatible with Kant and suggest the fruitful application of Kant's ideas. To establish this claim, I first will outline how Kant's analysis of organisms as self-organizing beings has been revived and revised in contemporary philosophy of biology to make sense of emergence, teleology, and non-human natural agency. Second, it will show how the study of scientific practice—popularized by the work of Thomas Kuhn— has led to a vision of science as a complex, self-correcting, normatively driven and historically adaptive system—a vision that seems to echoe Kant's study of the normative presuppositions of our theoretical inquiry. Together, these developments demonstrate that the conceptual tools Kant developed to navigate the limits of reason and the puzzles of natural purpose could be invaluable tools to make sense of our historical situation and solve its deepest problems. For the revival of natural agency and scientific practice has brought forth a conflict between the mind's ordering activity and the world's intrinsic order (see next section). Kant is useful because he can resolve this perceived dilemma[6] that dominates current understandings of the complex world.

The existence of complex systems pose specific conceptual challenges to our understanding of nature. Such systems exhibit order without centralized control, characterized by nonlinear causality, feedback loops, and emergent properties. This very autonomy raises a pressing metaphysical puzzle: how do we reconcile their apparent purposiveness (e.g., adaptive behaviors) with the absence of an intentional designer, i.e., God?[7] While contemporary science can mathematically model how self-organization occurs, it leaves the conceptual grounding of this purposive unity vague. This is precisely the dilemma Kant sought to dissolve, making his work pertinent for this line of inquiry.

Kant's framework in the *Critique of the Power of Judgment* (2000) not only provides his own conceptual solution but also supplies the essential taxonomy for the major positions that dominate today's debates. To explain the experience of purposive unity, we

---

[6] The tension between mind and world was not a dilemma at all for Kant. Rather, the fundamental tension that Kant aims to dissolve is that between freedom and nature (see chapter 2).
[7] Complex systems have been a major topic in the functions literature. Robert Cummins (1975) and Tim Lewens (2004) emphasize that functions are just causal contributions to the capacities of a larger system. This contrasts with 'selected-effect' theories of function, where a trait's function is what it was selected for in evolutionary history (Millikin 1989; Neander 1991; see also Wright 1973). Organizational approaches share with causal theories the focus on complex systems, but instead of focusing merely on causal roles, they add a normative dimension: functions are tied to the closure of constraints of a system, where parts of a system mutually sustain each other in a way that *ought* to maintain the whole (Moreno & Mossio, 2015). The latter is closer to Kant's view, but with important differences: see Chapter 3.

can either conclude that final causes are 'real' (what Kant identified as explaining things teleologically) or that they are mere figments of our imagination, awaiting to be reduced to combinations of existing properties of the system's parts (what Kant identified as the 'mechanistic' reduction of final causes to efficient causes). Contemporary philosophers frame the former as 'strong emergentism' (Chalmers, 2006), while the latter falls under various forms reductionism (Bechtel & Richardson, 1993). Kant's own solution—that we *necessarily* experience organic beings 'as-if' they final causality, without requiring final causality to exist 'ontologically'—is closer to what is known as 'weak emergentism,' which characterizes higher-level phenomena (e.g., life, mind) as epistemologically novel but ontologically reducible. The 'emergence' of order is a condition for the possibility of judging living systems. Kant, however, would reject the inherent realism in most weak emergentist claims—that is, Kant does not want to 'ontologically' reduce emergent properties to reductive mechanical causes, since nature's intrinsic structure, for Kant, is one we cannot know (see Chapters 2 and 3).[8]

Yet Kant is relevant to present-day philosophical issues surrounding self-organization for a more fundamental reason. Kant, for first time in philosophical history, coins the term 'self-organization' to designate organisms that order themselves apart from external design. Organisms are 'innerly purposive' independent from any 'external purpose' set to them by some outside agent (e.g., God, other humans). Yet they should not be seen as if they were 'really' organized to fulfill a purpose. This conviction escapes

[8] Weak emergentists do—albeit they might not explicitly put it this way—recognize that there is an interface between our knowing practices and how things are, which is in loose agreement with Kant.

our cognitive capacities as finite agents. Instead, Kant argued that judging beings to be self-organizing is necessary for our cognition, but only to 'regulate' our scientific explanations. We can never know that such self-organization is 'constitutive' of matter itself. Unlike Kant, contemporary scholars now often regard organisms as 'constitutively' purposive, meaning that they exhibit a kind of purposiveness that escapes the limits of our judgment (Gambarotto & Nahas, 2022). However, like Kant, they still reject the introduction of any special causal powers or appeals to intelligent design.

By reintroducing the concept of self-organization, twenty-first-century philosophy of biology follows a path implicit in Kant's analysis. In contrast to the reductionist tendency that dominated much of 20th century biology, which focused on genes and molecules as the primary units of explanation, the concept of 'organism' is making a comeback as the central explanatory concept in biology (Bateson, 2005; Huneman, 2010; Nicholson, 2014). This movement has been heavily influenced by developments in evolutionary developmental biology (Goodwin, 1982), epigenetics (Jablonka & Lamb, 1995), phenotype plasticity (West-Eberhard, 2003) and niche construction (Odling-Smee et al., 2013). These perspectives emphasize that the organism is not a product of blind or random design, nor is it a passive byproduct of our genes, but is an active, self-organizing agent. This conception has been central to cyberneticians and philosophers who work with the notion of 'autopoiesis,' which refers to a 'self-creating' metabolic network that maintains itself in the face of environmental disturbance (Bich & Damiano, 2008; Bitbol & Luisi, 2004; Di Paolo, 2005; Luisi, 2003; Maturana & Varela, 1980; Weber & Varela, 2002).

The idea that organisms are purposive and self-organizing is articulated by Stuart Kauffman (1993), Ilya Prigogine (1977), and Howard H. Pattee (1973), who use the thermodynamic concept of ‘constraint’ to explain ‘teleological’ biological organization. Constraints act as boundaries that channel the transfer of energy throughout a system. In ‘work cycles,’ these constraints channel energy to reinforce other constraints, which reciprocally channel energy to reinforce the original constraints. These cycles are teleological insofar as they maintains a self-organizing causal regime where “the organisation of constraints can be said to achieve self-determination as self-constraint, since the conditions of existence of the constitutive constraints are, because of closure, mutually determined within and by the organisation itself” (Mossio & Bich, 2017). Current challenges to this approach seek to differentiate the uniqueness of living entities from other self-organizing phenomena (Cusimano & Sterner, 2020; Deacon, 2011; García-Valdecasas, 2022; Montévil & Mossio, 2015; Toepfer, 2012). Alicia Juarrero has also developed an extensive framework to understand how final causes can be interpreted as local efficient causes undergoing local constraints (2023). Kant is again pertinent to these discussions regarding the epistemic status we accord to self-organizing beings.

Kant clearly did not foresee how ‘mechanistic’ explanations would evolve from reductive, linear models into nonlinear frameworks—ones that can now be rigorously expressed through differential equations and computational modeling. Whether this is reason enough to move beyond Kant remains controversial (See Chapter 3; Gambarotto & Nahas, 2022). Yet what is crucial to Kant’s analysis of self-organizing beings is that he is not merely concerned with self-organization in organisms; he is concerned with the way in which *reason* (our practices of giving and responding to reasons) organizes such that we have the *right* to make the claims about, e.g., self-organizing organisms. Kant thinks it is indispensable to couple our investigations into nature with an investigation into the normative presuppositions that guide our own scientific practices. A Kantian analysis is hence incomplete without studying our own capacities to make valid empirical judgments about the world.[9] That is, it is incomplete without studying the way *we* organize when we make cognitive judgments about the world. This makes Kant's perspective unique in the philosophy of biology literature.

The contemporary resurging interest in the norms and principles latent in our scientific practices mean that philosophy of biology is not the only sphere that would be aided by a resurgence of Kantian themes.[10] The 'practical turn' in contemporary philosophy, history and social studies of science comprise a shift towards what scientists actually *do* against the background of norms and practices of their own creation (Barad, 2006; Cartwright, 1983; Chang, 2004, 2022; Dupré, 1993; Graham, 1981; Hacking, 1983; Keller, 2003; Latour & Woolgar, 1979; Longino, 2002; Pickering, 1995; Rheinberger, 1997; Rouse, 2002; Soler et al., 2014). This implicitly mimics Kant's concern with unpacking the normative presuppositions of the then-current practices of Newtonian science and Linnaean biology. The findings of this contemporary movement are too diverse to summarize at this juncture. A core agreement, however, is the rejection

[9] Kant argues that science is not merely the discovery of an empirical world but also a human activity that self-organizes around fundamental norms—what he calls 'Ideas'—and cognitive capacities—what he calls the categories. See Chapter 2

[10] In this sense, I interpret Kant as a 'pragmatist' along the lines of Pierre Keller (1998), Ernst Cassirer (1981), and Karl Jaspers (1962). See chapter 2 for a defense of this view.

of the 'propositional' image of science, where scientific understanding is identified with a body of knowledge or set of justifiable positions. People who are attuned to scientific practice do not reject this body of knowledge but situate it within an active process of understanding our world from the backdrop of our various epistemic interests, institutional and material settings, practical skills, and political circumstances. As Joseph Rouse puts the point,

> The sciences continually revise the terms and inferential relations through which we understand the world, which aspects of the world are salient and significant within that understanding, and how those aspects of the world matter to our overall understanding (2015, 206-7).

Rouse echoes a Kantian theme by analyzing the structure of science's constantly revised inferential relations—the very relations that constitute our understanding of nature. The history of science reveals the norms that unite and guide these relations, a perspective decisively advanced by Thomas Kuhn, whose work relocated Kantian questions to the arena of historical scientific change. To Kuhn, science alternates between periods of relative stability and order—what he calls 'normal' science, where most scientists agree upon paradigmatic explanatory frameworks and fundamental principles—and 'revolutionary' periods, where explanatory frameworks and fundamental principles are rapidly revised due to accumulated anomalies (Kuhn, 1962). This perspective in fact calls on us to question whether our scientific practices are *themselves* self-organizing adaptive systems. For instance, as we saw with the study of criticality and panarchy in II.B, as a system becomes increasingly ordered, it rigidifies, becoming less resilient to novel disturbances and thereby vulnerable to collapse. When the system collapses, it is followed by a complete re-organization to adapt to the shocks that initiated the collapse.[11] Kuhn's scientific 'revolutions' seem to follow a similar trajectory. Are scientific practices just an evolution of natural self-organizing processes?

Such a position seems to be accepted by contemporary philosophers influenced by the Stanford school. Composed of Nancy Cartwright (1983, 1999), John Dupré (1993), Ian Hacking (1996), and Peter Galison (1997), the Stanford school is famous for influencing a generation of philosophers of science to reject the 'unity' of science and instead advocate for scientific pluralism. On this self-organizing conception of science, knowledge is no longer seen as a unified, coherent whole dictated from above, but rather as a decentralized collection of multiple theories and practices that generate their own order from the bottom up. There is no single, 'fundamental' set of natural laws to which all other explanations can be reduced. Instead of a 'theory of everything,' these philosophers advocate for insights into local, context-dependent regularities that emerge adaptively in specific situations. This perspective has led members of the Stanford school to focus on the actual practices of scientists, highlighting the diverse methods and models they use to understand specific phenomena—practices that are often guided by specific theoretical interests rather than an abstract search for a unified truth. In all these respects, science as depicted by the Stanford school behaves like a paradigmatic self-organizing system: adaptive, decentralized, and generative of its own order.

It is no surprise that Robert C. Bishop and Roman Frigg—in a collection edited by Nancy Cartwright— have argued that the naturalist basis for the 'messiness' in science

[11] See (Nickles, 2008) for an interpretation of Kuhn in terms of complex systems theory.

is to be found in complex systems science. They 'naturalize' this epistemic situation by grounding the complexity of science within a conception of nature that is self-organizing. They argue that the traditional approach, which sees nature as governed by exceptionless laws, "crumbles in the way of chaos and complexity studies." (Bishop & Frigg, 2016, 55). For them, the alternative is that nature is composed of "self-organized order" that "depends on the context of our universe and the localized contexts in particular regions of the universe" (Ibid., 74). Here, we see that the ultimate explanation of the "dappled world"—as Nancy Cartwright calls it to describe the referent of the plurality and patchiness of our explanations—lies not merely within a feature of our epistemic practices, but in a reality that is dynamically self-organizing into patchy domains of order independently of those human practices.

Whether we accept this 'naturalist' explanation of our scientific practices, it seems that our scientific practices are like natural self-organizing systems. In fact, they seem to be one of the most paradigmatic examples of self-organizing behavior. Recall the properties of self-organizing systems above: they are open to their environment, resilient, generate their own properties for themselves, and are nonlinear. Scientific practices have all these features. With their experimental component, scientific practices are open to their environment; they are resilient in the face of recalcitrant evidence and other counterexamples, and adapt to these 'disturbances' once they become overwhelming; they generate their own norms/rules themselves, thereby being autonomous in the sense that they have to come up with their own laws or regularities (nature does not hand them to us on a silver platter); and they behave nonlinearly, relying on large theoretical edifices to prove even the simplest of facts, and not ‘linearly’ establishing these facts by reading them off of nature as if we were copying God’s notes.

In summary, the rise of complexity in the sciences, the renewed interest in philosophical questions concerning final causality and self-organization, and the ‘practical turn’ in philosophy of science are all ample motivation to revitalize Kantian philosophy. Not only does Kant introduce the term self-organization and give a sophisticated account of when we can judge things to be self-organizing, he does this against the backdrop of the inherent organization of our scientific practices. The primary motivation to bring back Kant is his sensitivity to the conjunction of the self-organization of nature and the way we self-organize through norms and reason. This conjunction brings about an underappreciated tension. If our scientific practices organize nature through their concepts and paradigms, does this mean nature’s order is merely our projection? Conversely, if nature’s self-organizing complexity is what explains our practices, does this mean our reason is not truly autonomous, but dictated by the external forces of nature? We are faced with a dilemma: either we privilege the world’s self-organization as the source of our own, thereby rendering our scientific practices mere products of external determination rather than genuinely self-organizing, or we privilege our mind’s self-organization as the source of the world’s, thereby reducing nature’s apparent autonomy to a projection of our own cognitive dynamics. The solution to this problem seems to require a co-investigation into the self-organizing mind and the self-organizing world without reducing one to the other—something that Kant masterfully achieved. Before we go into that, however, we need to investigate the dilemma in further detail.

IV. The Insufficiency of Complexity Realism and Idealism

The dilemma between privileging either our own self-organization or the self-organization of the cosmos crystallizes into two opposing philosophical positions, which I will call complexity realism and complexity idealism. While few may defend these positions in the form I attribute to them here, they represent the logical endpoints of the explanatory strategies available to contemporary discussions of self-organization. My aim is not to critique specific contemporary scholars. My contribution is rather to make explicit a tension that most remain largely unaware of. I aim first to unpack the logical structure of each position, and then show how each when made foundational, remains insufficient. Yet this insufficiency is instructive: it reveals that the self-organization of the mind and the self-organization of nature cannot be coherently thought in isolation, but demand an alternative framework where both are constituted simultaneously. Thus, by making the implicit tension explicit, I point us towards revitalizing Kant, who faced and overcame a similar tension in his time.

Let us define each position. Imagine we are trying to explain why our scientific practices are so complex and self-organizing. Complexity realism argues that the explanation lies in the world itself. Our practices can be trusted to make claims about a cosmos that is independently self-organizing. It is then straightforward to explain the self-organization of our scientific practices: the complexity of things 'out there' explains the complexity of our practices. This view can make two interrelated arguments: 1) since we can trust our scientific practices and these practices point to the self-organization of entities within the cosmos, these practices can in turn be fully explained by the real processes bringing about the emergence of life, human social organization, and finally scientific thought (a sufficiency argument); 2) since we must use multiple points of view, levels of explanation, abstractions, simplifications, and idealizations to make sense of the cosmos, the self-organization we see must exist in nature apart from us—nothing else can, with necessity, explain the complexity of our methods (a necessity argument). Both arguments start from a fundamental assumption: scientific practices 'look at' nature from a standpoint outside of nature, and nature is hence independent in the sense that our practices are not 'projecting' their complexity onto it.

Although nobody explicitly defends 'complexity realism' in the form I have described it, it has a basis in real currents of philosophical thought. According to William Wimsatt, for instance language is complex because it is a response to the complexity of the world (2007, 211). Wimsatt similarly attributes the complexity of science in its reliance on heuristics, points of view, biases, simplifications, generalizations, and idealizations (all from our endless epistemic aims) to the complexity of the world around us. Complexity realists could also be encouraged, we might think, by the alleged defeat of the unity of science program. In the wake of Nancy Cartwright's work dismantling the notion of universal scientific laws (1983), various philosophers of science have sought to ground such 'disunity' in a much messier, disunified notion of nature. For instance, John Dupré argues that his metaphysics is based on "skepticism" of the old unity approach. His goal is "to anchor that skepticism in a more appropriate metaphysics" (Dupré 1993, 7). The unity of science, he claims, failed to give us an accurate picture of the constitutive building blocks of the cosmos. To him, an alternative metaphysics is necessary: if reality is not simple, then it is complex, composed of an evolving field of self-organizing processes. Under these approaches, complexity realism states that the complexity of our epistemic practices is explained by a cosmos that is inherently complex. [12]

Complexity idealism, in contrast, locates the source of complexity in the knower. It attributes the apparent self-organization of the cosmos to the complexity of our own scientific practices, which are shaped by human interests, historical contingencies, and cognitive frameworks. The complexity of our knowledge reflects not a complex cosmos, but the complexity of our scientific practices, especially how they depend upon our human interests and the particular ways in which we have adapted to our environments. Social constructivists Shapin and Schaffer seem to echo this perspective when they state, "it is ourselves and not reality that is responsible for what we know" (1985, 344). While it may appear that the cosmos is complex, this just could reflect our biological and

[12] Complexity realists seem to represent the majority of the philosophical community concerned with such issues. Scientists such as Stuart Kauffman, John Holland, and Ilya Prigogine all seem to argue that self-organization is not only real, but the fundamental process within our cosmos. Aside from William Wimsatt (2007) and post-unity philosophers of science within the analytic tradition (see above), strong emergentist metaphysicians align themselves with complexity realism as so described. In the realm of philosophy of mind, Anthony Chemero—integrating James Gibson's (1979) ecological psychology with non-linear dynamics— argues that the mind-world interaction is a dynamical system that can be studied by 'real' relational properties that he calls 'affordances (Chemero, 2009). In continental philosophy, complexity realists mostly fall under the general umbrella of 'speculative realists' who, following Quentin Meillassoux, reject post-Kantian "correlationism" which Meillassoux defines as "the idea according to which we only ever have access to the correlation between thinking and being, and never to either term considered apart from the other" (2008, 5). Deleuzians (perhaps not Deleuze himself) such as Manuel DeLanda (2002) argue that intensive properties (such as temperature, pressure, chemical gradients) exist in nature that drive phase transitions and emergent structures. Reality is inherently shaped by these 'virtual' gradients and thresholds, which are in turn governed by singularities that also exist (virtually) in nature. Alicia Juarrero also aligns herself with speculative realism in arguing that self-organizing complexity is derivable from 'constraints' on efficient causes, constraints that actually exist in the real world (2023). Iain Hamilton Grant's 'transcendental materialism' is also a speculative realist argument derived from Schelling where matter is inherently self-organizing and productive (2006).. Such 'agency' of matter is also shared by Karen Barad (2006), who argues that quantum mechanics has proved that matter is dynamically self-organizing with us in an observer/observed system.

psychological faculties that have adapted to view the world in a particular way. The complexity idealist approach tends to focus on the psychology and sociology of science, including its political dynamics. [13]

While seemingly distinct, complexity idealism shares the same core assumption as complexity realism. They both conceive of our scientific practices as independent from the natural world. They both also think our scientific practices are self-organizing and complex. Where they differ is the explanatory priority they attribute to either our practices or the real world. Whereas a self-organizing cosmos explains the self-organization of our scientific practices under the realist view, the idealist view claims that the self-organizing complexity of the cosmos is attributable to the complexity of our human-interest-based scientific practices.

If we grant that both complexity realism and complexity idealism rely on a new conception of the 'natural' based on emergence and self-organization, both seem to fail as a naturalist project. Each position must relinquish one of two claims: 1) that there exist self-organizing phenomena in nature, or 2) that our scientific practices are self-organizing. Consider complexity realism first. Complexity realism attributes the complex self-organization of the cosmos to explain the complex self-organization of our scientific practices. We could not help but make our practices complex: the world just is that way.

---

[13] Contemporary complexity idealists are much rarer. Their most striking exemplars are Ernst von Glasersfeld (1995), whose 'radical constructivism' claims that cognitive construction is shaped by our survival needs, reducing all complexity we see in the world to adaptations we have. Donald Hoffman (2015) seems to echo this claim in his claim that what counts as complex for a given system (like human cognition) depends on a system's distinctive filters or 'interface' (2015). Thomas Metzinger (2009) also seems to be a complexity idealist, where the 'world'—even the world according to science— is just a brain constructed model optimized for survival, what Metzinger calls an 'ego tunnel.'

The complexity realist must assert, then, that the self-organization of the cosmos organizes our thought. In other words, an external feature to our practices (the self-organizing cosmos and the self-organizing entities and processes within it) dictates how our practices self-organize. However, if there was outside external control, our practices would cease to be self-organizing. They would be organized, certainly, but not *self*-organizing. The autonomy of scientific practices would be eliminated in favor of the cosmos dictating precisely how our practices are supposed to organize (as if God was forcing us to write down, e.g., Newton's laws). However, the argument that the cosmos is self-organizing was established on the very premise that our practices are self-organizing. The self-organizing cosmos *explains* the organization of our practices. Since our practices no longer seem to be self-organizing if an external world seems to organize them for us, we can no longer call our practices a naturalized extension of the self-organization of the cosmos.

Complexity idealism might seem to resolve this problem by claiming that our scientific practices are autonomous and independent from being dictated from any outside external control, such as nature. However, this comes at the expense of sacrificing the claim that the cosmos is in fact self-organizing. If the cosmos really is nothing other than what we project onto nature, then nature is organized according to our human interests, and therefore does not have any intrinsic self-organization. If the scientific consensus seems to claim that nature is self-organizing, then nature is no longer natural—it is organized by us. If the claim is, instead, that the self-organization of our practices leads to the mere *appearance* that the cosmos is self-organizing, whereas we remain agnostic whether the cosmos is *actually* self-organizing or not, this again fails to embed our scientific practices in a natural process. Rather, they are embedded merely in the *appearance* of a natural process.[14]

Complexity realism and complexity idealism hence both seem to fail to be naturalist projects. Yet they also seem insufficient for reasons entirely independent of naturalism. For complexity realism to be necessary, it would have to refute the skeptic represented by the complexity idealist, who insists that the complexity we see in the world could be a projection of our practices. To refute complexity idealism, we would have to *know* that our practices are *not* projecting complexity on to the universe as a side effect of achieving our various ever-so-complicated epistemic ends that are suited to our own biological needs. But the skeptic here would be impossible to refute, as we have no way of stepping 'outside' our own knowledge to know whether we are projecting complexity onto the world or if the world is 'truly' complex. Thus, the complexity realist faces a skepticism problem that the complexity idealist can easily resolve, as no such knowledge is required for the idealist.

---

[14] The same argument can be applied to a form of self-organization that does not 'touch' the world but seemingly is entirely self-contained. Here I am thinking about John McDowell's (1994) criticisms of Donald Davidson (1984). Davidson understood perception—our 'contact' with the world—as a causal prompting of discursive judgment in thought. If this is true, thought would be according to McDowell a "frictionless spinning in a void (1994, 66). A parallel criticism can be applied to contemporary understandings of scientific practice. Although he is not a traditional realist, Ian Hacking's "self-vindicating" conception of laboratory science seems to fall under a similar critique. Here, theories are not checked against a "passive world" as in traditional realism, but instead evolve together with our experimental "apparatus, and our observations" creating a "coherence theory of thought, action, materials, and marks" (Hacking, 1992). From the perspective of a self-organizing naturalism, Hacking's and Davidson's conceptions of the natural world would be quite unnatural. The world would not be organizing 'internally,' as it were, but organized 'externally' according to self-contained human practices.

Complexity realists might instead argue that although our multiple human aims, heuristics, and interests make our scientific practices complex, what *best* explains the complexity of our scientific practices is simply the complexity of the world. In other words, while they might concede the necessity of their views, they might double down on their sufficiency. The complexity idealist is simply insufficient in explaining the complexity of the cosmos. Their argument might work something like this: we make idealizations because the world is complex. These idealizations fail frequently when applied outside their limited domain of validity due, again, to the complexity of the cosmos. Hence, the insufficiency of our simplifications and generalizations point to a complex world beyond our practices. Only a complexity realist can sufficiently explain this.

Yet the explanation that we are projecting the complexity of our own epistemic idiosyncrasies onto the cosmos seems just as plausible, and equally as sufficient. The claim that the universe is complex may simply reflect the nature of beings like us, with various epistemic interests and goals, who possess limited epistemic tools and harbor lofty ideals of explaining everything through these constraints. In other words, the idea that our idealizations fall short is not due to the complexity of the cosmos, but the plethora of simplifications and complex practices that we are already engaging in. By contrast, consider a relatively simple self-organizing system. Since a ladybug has relatively few aims (eating small insects and reproducing, perhaps giving people luck), would it be justified in drawing the conclusion that the world is simple? The ladybug's simple practices seem to reveal more about the ladybug than the world.

Perhaps this is still too far-fetched. There is yet a deeper reason to doubt the complexity realist. To understand this, let us ignore the challenge presented by the complexity idealist. There is no projection of our complexity onto the world: we can rely on our scientific practices to tell us the truth about the complex universe, rather than truths about our own organization. We can assert with confidence the reality of the complex systems that we study in our practices, and thereby the complexity of the world they belong to. But what explains the complexity of our scientific practices? If the reason why we have so many aims and practices is because the world is complex, we have argued in a circle. For our scientific practices justified the existence of the complexity of the world in the first place. If these practices are to be explained by appealing to the complex world they aim to explain, we have established neither the complexity of the world nor the complexity of our scientific practice.[15] Thus complexity realism, in its most ideal form, seems to be circular at best.

We cannot ground the complexity of our scientific practices on the claims put forth by the complexity realist. The complexity of the cosmos could have been a projection from the complexity of our scientific practices; it seems to be insufficient as the best explanation of the complexity of our scientific practices; and it cannot be grounded on the empirical findings of science without risking circularity. That the universe is complex could be merely an ideal regulating the way finite intellects navigate the world, in place so we can respond to diverse and unexpected events. In science, we would apply

[15] Even if we were to secure a definition that lays out the necessary and sufficient conditions of a complex system (Estrada, 2023; Ladyman et al., 2013; Rathkopf, 2018; Strevens, 2005), we would still have to face this circularity challenge. What explains our ability to define complexity, if not some emergent property of a complex system?

this ideal so that we do not redundantly clump up our scientific theories into one empty, over-generalized theory.

The alternative of complexity idealism seems equally insufficient, and equally circular in its most ideal form. Instead of saying that reality is complex, they argue that it is really our psyche (or equivalently, our social/economic/political organization) that is complex, and that this structure is projecting its complexity onto the world. Our own complexity explains why we think the world is so complex. Yet this also seems circular. How do we know that our psychologies, etc. are complex if our self-conception of what our psyches are is a mere projection from our own psyches? Positing that our psyches are complex would thereby be a projection by our complex psychologies. Complexity idealism, taken to its natural conclusions, seems to be circular.

Complexity idealism and complexity realism both are insufficient. Neither can adequately explain the complexity of our scientific practices, let alone the complexity of the world. While they each begin with a profound and insightful intuition—that our practices and the natural world seem to be self-organizing—they run into difficulties when linking the two together in an overarching explanation. The presupposition that they both share is the unidirectionality of explanation: organization must originate either in the self-organization of the cosmos or in the self-organization of our minds, with one serving as the foundational cause of the other. This shared assumption is what generates the circularities and contradictions that lead to the insufficiencies of each project.

We are prompted to start a new investigation: How does the self-organization of our reason interrelate with the self-organization of nature such that each is possible only through its relation to the other? Are scientific practices really 'mental' as opposed to the 'physical' matter they study? These are not questions for empirical science—which presupposes the separation of knower and known—but a transcendental question concerning the conditions of possibility for successful practice itself. It is the very question that Immanuel Kant's critical philosophy is uniquely designed to address.[16]

V. Roadmap for Dissertation

I hope the preceding dialectic has done more than merely critique two extreme positions. Rather, I hope it points to a deeper problem: any philosophy that treats mind and world as independent substances struggles to explain how they become productively entangled in the very practice of giving and asking for reasons. Kant's revolutionary move is to begin not with the mind or the world as it exists independently of the mind, but with the structure of this normative practice. His critical philosophy therefore provides our essential framework, because it shows how the intelligibility of a self-organizing world and the self-organization of our own agency are bi-constituted within the practices of giving and responding to reasons. Now that we have established the necessity of Kant, let us see how this changes how we understand the proceeding roadmap of this work.

Kant's framework is itself contested ground: not all Kantians would agree with the normative and pragmatist reading I attribute to him. I therefore turn, in Chapters 2 and

[16] Certain Kantians in contemporary pragmatism are allies in this endeavor—see Chang (2022) for one striking example.

3, to developing a pragmatist reading of Kant that distinguishes my approach from more traditional 'impositionist' readings, on which the mind imposes form onto the world. Only once this interpretation is secured can we trace how the Kantian logic of bi-constitution extends beyond human practices to encompass nature itself—and ultimately, what ethical implications follow from doing so (Chapter 6).

In Chapter 2, I outline the central Kantian concept in my reading: the Idea. For Kant, Ideas are the normative contexts—like truth, justice, or the good—that guide our practices of reasoning. I crucially promote the thesis of 'bi-constitution' where our mental apparatuses and the external world do not pre-exist our practices but are simultaneously constituted as we participate in the practices guided by the Idea. The necessity and success of acting from these Ideas presupposes that nature is amenable to them—that it (at least according to the Idea of autonomy that guides moral practice) exhibits a purposive unity we can pick out as a whole. Kant thus shows that autonomous reason and a self-organizing nature are not separate; they are the mutually necessary conditions for our reason-giving practices to exist.

Chapter 3 extends this logic. If we understand ourselves and the world through Ideas, how do we understand natural entities like organisms or ecosystems? I argue that Kant's analysis of "natural ends" provides the answer: we must judge such entities 'as if' they were agents analogous to our own Idea-mediated agency. This analogy operates through a four-part schema of bi-constitution: the entity's form orients its parts into a self-sustaining whole; its parts reciprocally produce and maintain that whole; the entity selectively draws from its environment, which its form renders purposive as a niche; and success or failure in this process tests coherence between activity and norms, pressuring adaptation or collapse. This 'as if' judgment is not mere projection but a necessary regulative principle of reflective judgment, objectively validated when the practices it guides—biology, ecology, and complex systems science—lead to successful discovery. Against contemporary "ecological agency" theories that reduce agency to a matching relation between pre-given capacities and environments, I show that such accounts tacitly presuppose the Kantian logic of purposive self-organization they seek to replace. Kant thus populates the self-organizing world with non-human agents—from organisms to convection cells—by conceiving them, in analogy to ourselves, as responsive to an immanent Idea.

Chapter 4 turns to Ernst Cassirer to show how teleological judgment is successfully affirmed in both our understanding of human culture and the natural world. Cassirer's philosophy of symbolic forms reveals science, myth, language, and art as distinct meaning-making systems, each guided by its own Form or Idea. I argue that Cassirer extends Kant's bi-constitution to all symbolic activity: the "formal imperative" demands that nature ought to conform to forms, a transcendental condition of reason itself. When we interrogate nature with specific theories, nature answers back in ways not under our control, and the success of our inquiries confirms the harmony between our cognitive frameworks and nature's own self-organizing structure. Cassirer thus navigates between intellectualist readings (nature as projection) and naturalist readings (mind as copy), showing instead that thought and nature co-arise: the self-organization of culture—its dynamic of limitation, creative act, and adaptive self-preservation—is the very process through which nature's self-organization is disclosed. Science succeeds not because our concepts copy reality nor because we impose order, but because the symbolic wholes we create prove adequate to a nature that is itself dynamic and structured. Cassirer thus shows how our understanding of ourselves and the natural world successfully affirms the agential framework of the cosmos, without collapsing into subjective idealism or uncritical realism.

Chapter 5 will unpack the core assumptions of contemporary complex systems science. I will examine whether the core theoretical edifices of complexity science—graph theory, dynamical systems, criticality, and panarchy—presuppose or validate the Kantian judgment that sees nature as comprised of self-organizing agents, or whether they instead render teleological judgment unnecessary. I will argue that the successful application of these formal tools to natural phenomena proves neither a 'deflationary' reading of teleological causes, nor do they prove a realist metaphysics of self-organization. Rather, they validate the ongoing success of the Kantian hypothesis that nature can be judged *as if* it were composed of interacting agents. Furthermore, the success of this hypothesis seems to grant it objective validity, albeit we can never know whether it is true.

Chapter 6 will articulate the ethical culmination of this synthesis. If our autonomy presupposes nature's self-organization, then the ultimate end of our morality—the highest good—includes both the kingdom of ends and the kingdom of nature. Our moral vocation cannot be isolated from the integrity of the natural world. I construct a Kantian environmental ethics that moves beyond both the indirect-duty tradition (which grounds obligations to nature in moral self-perfection) and direct-duty revisionism (which attributes moral standing to animals based on shared traits). Instead, I argue that the highest good—the harmony of virtue and happiness—constitutively requires a purposively unified natural world. To will the highest good is to will a self-organizing cosmos, and a maxim that degrades nature's integrity generates a contradiction in will, yielding a direct but imperfect duty to preserve and promote nature's self-organizing form. What we are obligated to preserve is not any static state of nature but the form of self-organization itself—a dynamic, multi-scalar architecture of interlocking agents. Consequently, we incur two imperfect duties: to refrain from disintegrative practices and to promote integrative ones. The highest good is therefore the harmony between human freedom and the autonomous flourishing of nature.

In sum, this dissertation will argue that Kant's philosophy, properly understood as a pragmatic-transcendental inquiry into the norms of practice, offers a coherent path through the impasse that our historical situation seems to warrant. It provides a framework that navigates the dual resurgence of the study of scientific practice and the discoveries of complex systems science. The outcome is encompassing environmental ethical vision grounded in the very conditions of our existence as rational, natural beings.

Bibliography

Anderson, P. W. (1972). More is different. *Science,* 177(4047), 393–396.

Bak, P., Tang, C., & Wiesenfeld, K. (1987). Self organized criticality: An explanation of 1/f noise. *Physical Review Letters*, 59 (4), 381-384.

Barabási, A. (2016). *Network Science*. Cambridge University Press.

Barad, K. (2006). *Meeting the Universe Halfway*. Duke University Press.

Bateson, P. (2005). The return of the whole organism. *Journal of Biosciences*, 30(1), 31-39. https://doi.org/10.1007/BF02705148

Bechtel, W. & Richardson, R. (1993). Emergent phenomena and complex systems. In J. Kim, A. Beckermann, & H. Flohr (Eds.), *Emergence or Reduction? Essays on the prospects of nonreductive physicalism* (257-288). De Gruyter

Belousov, B. P. (1959). A periodic reaction and its mechanism. *Collection of Abstracts on Radiation Medicine*, 147, 145.

Berkes, F., & Folke, C. (Eds.). (1998). *Linking Social and Ecological Systems: Management Practices and Social Mechanisms for Building Resilience*. Cambridge University Press.

Bich, L., & Damiano, L. (2008). Order in the Nothing: Autopoiesis and the Organizational Characterization of the Living. In *Physics of Emergence and Organization*.

Bishop, R. C., & Frigg, R. (2016). From Order to Chaos and Back Again. In N. Cartwright & K. Ward (Eds.), *Rethinking Order after the Laws of Nature* (55–74). Bloomsbury Academic.

Bitbol, M., & Luisi, P. L. (2004). Autopoiesis with or without cognition: Defining life at its edge. *Journal of the Royal Society Interface*, 1(1), 99-107. https://doi.org/10.1098/rsif.2004.0012

Camazine, S., Deneubourg, J. L., Franks, N. R., Sneyd, J., Theraulaz, G., & Bonabeau, E. (2003). *Self-Organization in Biological Systems*. Princeton University Press.

Cartwright, N. (1983). *How the Laws of Physics Lie*. Oxford University Press.

Cartwright, N. (1999). *The Dappled World: A Study of the Boundaries of Science*. Cambridge University Press.

Cassirer, E. (1981). *Kant's Life and Thought*. Yale University Press.

Chalmers, D. (2006). Strong and Weak Emergence. In. P. Clayton & P. Davies (Eds.), *The re-emergence of emergence: the emergentist hypothesis from science to religion* (244-256). Oxford University Press

Chang, H. (2004). *Inventing Temperature: Measurement and Scientific Progress*. Oxford University Press.

Chang, H. (2022). *Realism for Realistic People: A New Pragmatist Philosophy.* Cambridge University Press.

Chialvo, D. R. (2010). Emergent complex neural dynamics. *Nature Physics,* 6, 744-750.

Chemero, A. (2009). *Radical embodied cognitive science*. MIT Press.

Cross, M. C., & Hohenberg, P. C. (1993). Pattern formation outside of equilibrium. *Reviews of Modern Physics*, 65(3), 851–1112.

Cummins, R. (1975). Functional Analysis. *Journal of Philosophy*, 72.

Cusimano, S., & Sterner, B. (2020). The Objectivity of Organizational Functions. *Acta Biotheoretica,* 68(2), 253-269. https://doi.org/10.1007/s10441-019-09365-9

Deacon, T. (2011). *Incomplete nature: How mind emerged from matter.* Norton.

DeLanda, M. (2002). *Intensive Science and Virtual Philosophy*. Continuum.

Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences,* 4(4), 429-452. https://doi.org/10.1007/s11097-005-9002-y

Dupré, J. (1993). *The Disorder of Things: Metaphysical Foundations of the Disunity of Science.* Harvard University Press.

Eigen, M. (1971). Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften,* 58(10), 465–523.

Estrada, E. (2023). *What is a Complex System, After All?* Foundations of Science.

Galison, P. (1997). *Image and Logic: a Material Culture of Microphysics*. University of Chicago Press.

Gambarotto, A. & Nahas, A. Teleology and the organism: Kant's controversial legacy for contemporary biology. *Stud Hist Philos Sci,* 93 (C), 47-56.

García-Valdecasas, M. (2022). On the naturalisation of teleology: self-organisation, autopoiesis and teleodynamics. *Adaptive Behavior*, 30(2), 103-117.

Gibson, J.J. (1979). *The ecological approach to visual perception*. Houghton, Mifflin, and Company.

Glasersfeld, E. (1995). *Radical Constructivism: A Way of Knowing and Learning*. The Falmer Press.

Goodwin, B. C. (1982). Development and evolution. *Journal of Theoretical Biology*, 97(1), 43–55. https://doi.org/10.1016/0022-5193(82)90275-2

Graham, L. R. (1981). *Between Science and Values*. Columbia University Press.

Grant, I. H. (2006). *Philosophies of nature after Schelling*. Continuum.

Gunderson, L. H. & Holling, C.S. (2002). *Panarchy: Understanding Transformations in Human and Natural Systems*. Island Press.

Hacking, I. (1983). *Representing and Intervening.* Cambridge University Press. https://doi.org/10.1017/CBO9780511814563

Hacking, I. (1996). The Disunities of the Sciences. In P. Galison & D. J. Stump (Eds.), *The Disunity of Science* (37–74). Stanford University Press.

Haken, H. (2006). *Information and Self-Organization: A Macroscopic Approach to Complex Systems*. Springer Science & Business Media.

Hawking, S. (2000). "Unified Theory" is Getting Closer, Hawking Predicts. *San Jose Mercury News,* 29A.

Hoffman, D. (2015). The interface theory of perception. *Psychonomic Bulletin & Review,* 22(6), 1480-1506.

Holling, C. S. (2001). Understanding the Complexity of Economic, Ecological, and Social Systems. *Ecosystems*, 4, 390-405.

Holling, C. S. (2010). Resilience and Stability of Ecological Systems. In L. Gunderson, C. Allen, & C. S. Holling (Eds.), *Foundations of Ecological Resilience* (19–50). Island Press.

Holling, C. S., & Gunderson, L. H. (2002). Resilience and Adaptive Cycles. In C. S. Holling & L. H. Gunderson (Eds.), *Panarchy: Understanding Transformations in Human and Natural Systems* (25–102). Island Press.

Hooker, C. (2011). Introduction to Philosophy of Complex Systems. In C. Hooker (Ed.), *Philosophy of Complex Systems* (Vol. 10, 3–90). Elsevier.

Huneman, P. (2010). Assessing the prospects for a return of organisms in evolutionary biology. *History and Philosophy of the Life Sciences*, 32(2–3), 341-371.

Jablonka, E., & Lamb, M. (1995). *Epigenetic Inheritance and Evolution: The Lamarckian Dimension.* Oxford University Press.

Jaspers, K. (1962). *Kant: From* The Great Philosophers*, Volume 1*. (H. Arendt, Ed.). Harcourt, Brace & World.

Juarrero, A. (1999). *Dynamics in Action: Intentional Behavior as a Complex System*. MIT Press.

Juarrero, A. (2023). *Context Changes Everything: How Constraints Create Coherence*. MIT Press.

Kant, I. (2000). *Critique of the Power of Judgment* (P. Guyer, Ed.). Cambridge University Press.

Kauffman, S. A. (1969). Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology*, 22(3), 437–467.

Kauffman, S. (1993). *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press.

Kauffman, S. (1995). *At Home in the Universe*. Oxford University Press.

Keller, E. F. (2003). *Making sense of life: explaining biological development with models, metaphors, and machines.* Harvard University Press.

Keller, P. (1998). *Kant and the demands of self-consciousness.* Cambridge University Press.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press.

Ladyman, J., Lambert, J., & Wiesner, K. (2013). What is a complex system? *European Journal for Philosophy of Science,* 3(1), 33–67. https://doi.org/10.1007/s13194-012-0056-8

Langton, C.G. (1990). Computation at the edge of chaos: Phase transitions and emergent computation. *Physica D: Nonlinear Phenomena,* 42 (1-3), 12-37.

Latour, B., & Woolgar, S. (1979). *Laboratory Life: The Social Construction of Scientific Facts*. Sage Publications.

Laughlin, R. B. (2005). *A Different Universe: Reinventing Physics from the Bottom Down*. Basic Books.

Lewens, T. (2004). *Organisms and Artifacts. Design in Nature and Elsewhere*. MIT Press.

Longino, H. E. (2002). *The Fate of Knowledge*. Princeton University Press.

Luisi, P. L. (2003). Autopoiesis: A review and a reappraisal. *Naturwissenschaften,* 90 (2), 49-59. https://doi.org/10.1007/s00114-002-0389-9

Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and Cognition: The Realization of the Living*. D. Reidel Publishing Company.

Mazzocchi, F. (2008). Complexity in biology. Exceeding the limits of reductionism and determinism using complexity theory. *EMBO Reports*, 9 (1), 10-14.

Meillassoux, Q. (2008). *After Finitude*. Continuum.

Mersini-Houghton, L. *Before the big band: the origin of the universe and what lies beyond*. Mariner Books.

Metzinger, T. (2009). *The Ego Tunnel: The science of the mind and the myth of the self*. Basic Books.

Millikin, R. (1989). In Defense of Proper Functions. *Philosophy of Science*, 56 (2), 288-302.

Mitchell, M. (2009). *Complexity: A Guided Tour*. Oxford University Press.

Montévil, M., & Mossio, M. (2015). Biological organisation as closure of constraints. *Journal of Theoretical Biology*, 372, 179-191. https://doi.org/10.1016/j.jtbi.2015.02.029

Moreno, A. & Mossio, M. (2015). *Biological Autonomy: A Philosophical and Theoretical Enquiry*. Springer.

Morin, E. (1990). *Introduction á la Pensée Complexe*. ESF Éditeur.

Mossio, M., & Bich, L. (2017). What makes biological organisation teleological? *Synthese*, 194(4), 1089–1114.

Neander, K. (1991). Functions as Selected Effects: The Conceptual Analyst's Defense. *Philosophy of Science*, 58 (2), 168-184.

Newman, M.E.J. (2010). *Networks: An Introduction*. Oxford University Press.

Nicholson, D. J. (2014). The Return of the Organism as a Fundamental Explanatory Concept in Biology. *Philosophy Compass*, 9(5), 347–359. https://doi.org/10.1111/phc3.12128

Nickles, T. (2008). Disruptive Scientific Change. In L. Soler, H. Sankey, & P. Hoyningen-Heune (Eds.), *Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities?* (349–377). Springer.

Odling-Smee, J., Erwin, D. H., Palkovacs, E. P., Feldman, M. W., & Laland, K. N. (2013). Niche Construction Theory: A Practical Guide for Ecologists. *The Quarterly Review of Biology*, 88(1), 3–28. https://doi.org/10.1086/669266

Packard, N. (1988). *Adaptation Toward the Edge of Chaos.* University of Illinois at Urbana-Champaign, Center for Complex Systems Research.

Pickering, A. (1995). *The Mangle of Practice: Time, Agency, and Science*. University of Chicago Press.

Pickett, S. T. A., Cadenasso, M. L., & Grove, J. M. (2004). Resilient cities: meaning, models, and metaphor for integrating the ecological, socio-economic, and planning realms. *Landscape and Urban Planning*, 69(4), 369–384. https://doi.org/10.1016/j.landurbplan.2003.10.035

Prigogine, I., & Nicolis, G. (1967). On symmetry-breaking instabilities in dissipative systems. *The Journal of Chemical Physics*, 46(9), 3542–3550.

Prigogine, I., & Stengers, I. (1984). *Order Out of Chaos*. Bantam New Age Books.

Rathkopf, C. (2018). Network representation and complex systems. *Synthese*, 195, 55–78.

Reynolds, C. W. (1987). Flocks, Herds, and Schools: A Distributed Behavioral Model. *Computer Graphics*, 21(4), 25-34.

Rheinberger, H.-J. (1997). *Toward a history of epistemic things: synthesizing proteins in the test tube.* Stanford University Press.

Rouse, J. (2002). *How Scientific Practices Matter: Reclaiming Philosophical Naturalism.* Chicago University Press.

Shapin, S., & Schaffer, S. (1985). *Leviathan and the Air Pump*. Princeton University Press.

Soler, L., Zwart, S., Lynch, M., & Israel-Jost, V. (Eds.). (2014). Science *after the Practice Turn in the Philosophy, History, and Social Studies of Science.* Routledge.

Strevens, M. (2005). How are the Science of Complex Systems Possible? *Philosophy of Science*, 72 (4), 531-556,

Strogatz, S. (2015). *Nonlinear Dynamics and Chaos*. Taylor & Francis.

Thompson, E. (2007). *Mind in Life: Biology, Phenomenology and Sciences of Mind*. Harvard University Press.

Toepfer, G. (2012). Teleology and its constitutive role for biology as the science of organized systems in nature. *Studies in History and Philosophy of Science Part C :Studies in History and Philosophy of Biological and Biomedical Sciences,* 43(1), 113-119. https://doi.org/10.1016/j.shpsc.2011.05.010

Verlinde, E. Emergent Gravity and the Dark Universe. *SciPost Phys,* 2(16), 1-41.

Waldrop, M. M. (1992). *Complexity: The Emerging Science at the Edge of Order and Chaos*. Simon & Schuster.

Weber, A., & Varela, F. J. V. (2002). Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences*, 1 (2), 97-125.

West-Eberhard, M. J. (2003). *Developmental Plasticity and Evolution*. Oxford University Press.

Wimsatt, W. (2007). *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality.* Harvard University Press.

Wright, L. (1973). Functions. *Philosophical Review,* 82 (2), 139-168.

Wu, X., Zhang, J., Geng, X., Wang, T., Wang, K., & Liu, S. (2020). Increasing green infrastructure-based ecological resilience in urban systems: A perspective from locating ecological and disturbance sources in a resource-based city. *Sustainable Cities and Society*, 61, 102354. https://doi.org/10.1016/j.scs.2020.102354

Zhabotinsky, A. M. (1964). Periodic processes of the oxidation of malonic acid in solution (study of the kinetics of Belousov's reaction). *Biofizika*, 9, 306–311.

Chapter 2: Kant, Ideas and Bi-constitution

I. Introduction

Our brief survey of the contemporary positions surrounding the complexity 'revolution' has led us to two untenable options: complexity realism, which attributes the self-organization we observe to the cosmos independently of our own self-organization; and complexity idealism, which claims that the organization we see in the cosmos is entirely a product of our own cognitive or practical activity. Both positions share a common but hidden flaw: they assume that order must originate on one side or the other—either the world's order explains the mind's, or the mind's order explains the world's. Neither sees that the self-organization of our agency and the self-organization of the cosmos might be united in a single activity. Kant's uniqueness lies in exactly this insight: it is moral reasoning, and not theoretical reasoning about the world's structure, that binds them together.

For Kant, our own agency is self-organizing through a process I call bi-constitution. When we act under the guidance of what Kant calls Ideas—comprehensive normative frameworks like the Idea of a pot, the Idea of the understanding, or the Idea of autonomy—we do not simply apply pre-existing rules to a pre-existing world. Rather, participation in Idea-guided practices simultaneously constitutes two things at once. At the horizonal level, the Idea constitutes a normative identity for the agent (what it means to be a good potter, a rigorous scientist, or a virtuous person) and a purposive world (a domain of significance within which tools, evidence, or moral actions show up as meaningful). At the actual level, concrete acts guided by the Idea constitute a determinate practitioner (this potter shaping this clay) and a determinate product (this pot). The agent and its world, the practitioner and the product, are bi-constituted together through the same normative practice. This is what it means for human agency to be self-organizing: we give ourselves our own laws, set our own ends, and structure our own activities, all within the normative space opened by Ideas.

But what makes Kant truly unique is how he grounds the self-organization of the cosmos: not in theoretical observation (contra realism), nor in mental projection (contra idealism), but in the necessary structure of moral bi-constitution itself. Moral bi-constitution is guided by the Idea of Autonomy, where we simultaneously constitute ourselves[17] as autonomous moral agents and the world as a moral world. Practical reason sets as our ultimate end the highest good—a state where virtue and happiness are harmonized. This command would be empty if nature were ultimately indifferent or hostile to moral ends, for moral action requires that the world be such that our ends can be actualized. Since the moral law is authoritative, its necessary condition—that nature is a purposive, self-organizing system amenable to ends and ultimately the realization of the highest good—must hold. Thus, the self-organization of our moral autonomy and the self-organization of the cosmos are therefore not two separate facts to be reconciled. They are united in the single activity of moral bi-constitution. What theoretical reason can only hypothesize, practical reason affirms as a necessary postulate—not as theoretical

[17] At least, our phenomenal selves, since noumenally we already are free (see section II and IV below)

knowledge of nature-in-itself, but as an objective commitment for practical purposes that makes our moral striving coherent.

This chapter proceeds in four stages. Section II establishes the interpretive framework of bi-constitution, arguing against the standard 'impositionist' reading of Kant in favor of a practical-normative interpretation centered on what Kant calls the 'Idea.' Ideas, I argue, are neither mental constructs nor metaphysical entities but normative standards that guide practices, simultaneously constituting both purposive subjects (normative identities capable of pursuing ends) and purposive worlds (domains of significance within which objects show up as meaningful). Section III applies this logic to theoretical and practical reason, demonstrating their isomorphic, bi-constituting structure while establishing the primacy of the practical. Theoretical reason presupposes nature's systematic unity—never establishing it as known. Practical reason, grounded in the fact of reason, must constitutively postulate this unity as the necessary condition for the highest good, thereby providing the grounds that theoretical reason depends on but cannot secure on its own. Section IV unpacks what this means for our conception of nature. If practical reason demands nature's purposive unity for moral action, then nature must be understood not as mere mechanism but as a unified system amenable to moral ends. I then show how the third Critique adds a complementary mode of access. This is the disinterested pleasure felt when inquiry succeeds, confirming that the self-organizing world we are morally compelled to believe in is the same world that proves intelligible through science. Section V concludes by showing how this Kantian position offers a compelling alternative to both complexity realism and complexity idealism, grounding nature's self-organization in the autonomous self-organization of practical reason itself.

II. Kantian Ideas

Interpreters often frame Kant's 'Copernican Turn' as a decisive reversal of the traditional epistemological model. Against empiricist views of the mind as a passive receptor that 'copies' regularities from the external realm of nature into our theoretical models, Kant supposedly argues that the mind imposes innate categories and forms of intuition upon the undifferentiated "manifold of sensibility," thereby constituting the coherent, objective world of our experience. On this standard reading of Kant, the mind imposes a 'form' onto the disorganized 'matter' of the phenomenal world.[18]

While extremely influential, this "impositionist" interpretation appears to reinstate a form of the very dogmatic metaphysics Kant sought to dismantle. Here, the mind's innate structures function as a pre-given framework—whether cast as a replacement for

[18] Take Paul Guyer, for instance: "[Kant's] assumption is that it is precisely mind-independent objects which come before the mind by means of these ideal and mind-dependent forms" (Guyer 1987, 21). Peter Strawson has also been influential in popularizing the impositionist view: "The doctrine of transcendental idealism, and the associated picture of the receiving and ordering apparatus of the mind producing nature as we know it out of the unknowable reality of things as they are in themselves, are undoubtedly the chief obstacles to a sympathetic understanding of the *Critique*." (1966, 22). Norman Kemp Smith also seems to be very influential for this view: "Kant arrives at the conclusion that the *a priori*, the distinguishing characteristics of which are universality and necessity, is not given in sense but is imposed by the mind" (1992, xxxiii). See Fiona Hughes (2017) for a summary of the 'impositionist' school of Kantian interpreters. Hughes argues that Dieter Henrich—although not an impositionist himself—sets the grounds for impositionism, since he "does not sufficiently develop the object pole of the subject–object relation," thereby granting the subject exclusive authority as the primary source of order (2017, 35).

the rationalist's innate ideas[19], or reduced to a mere inventory of psychological facts.[20] This view not only struggles to explain the origin and authority of these forms,[21] but, as noted in the previous chapter's critique of complexity idealism, it renders the present inquiry impossible from the outset: if the order of the world is solely our own foisting of patterns onto nature, then nature cannot coherently be thought of as self-organizing. Instead, it would be organized by us.

Therefore, to avoid a dead end and remain faithful to the critical spirit of Kant's project—which seeks the grounds of possibility for our scientific, moral, and aesthetic practices in our self-understanding as human agents—we must turn to a less popular but more fundamental reading.[22] This interpretation locates the core of the Copernican Turn not in a cognitive imposition, but in a practical-normative revolution. It argues that reason, in pursuing its own essential ends—such as knowledge of nature, bringing about a moral world, and creating art—generates the normative frameworks ('Ideas') that simultaneously constitute us as specific kinds of agents and structure our conception of a world fit for our action.

[19] See Vanzo (2018) for a defense of Kant's similarities to Liebnizian intellectual concepts. In a similar vein, Robert Brandom claims that the normative authority of our concepts for Kant was unsatisfactorily "punted…into the noumenal realm" (2000, 33-34).
[20] For an appraisal and defense of Kantian 'psychologism', see Waxman (2019).
[21] I agree with Patricia Kitcher (2011, ch. 13) here, who argues that both the Leibnizian 'noumenalist' and the psychologistic interpretations of Kant cannot ground the normative authority of Kantian concepts. Kitcher's solution, however, is to claim that that the I-rule—the requirement that representations must be capable of accompanying the 'I think'—is a norm that any cognizer must follow. As we will see, my own solution comes not from the Transcendental Deduction alone, but situating the Deduction within the deeper Idea-guided practices that are ultimately grounded in practical reason.
[22] Pierre Keller (Keller, 2013, 2015), Ernst Cassirer (Cassirer, 1981), and Karl Jaspers (1962) read Kant in the practice-based, action-oriented focus way that is guiding my inquiry.

In the practical-normative reading (which I will henceforth adopt), the axes that dominate Kantian philosophy are not innate 'forms' imposing themselves on a formless 'matter.' Rather, they are two standpoints that reason takes on itself and its world: the theoretical standpoint, through which we aim to comprehend nature as a system of efficient causes, and the practical standpoint, through which we aim to determine what we ought to do. [23] These two practices create a seemingly insurmountable gulf: theoretical reason shows us a world of cause and effect where everything, including our own actions, appears predetermined. Our inclinations—our desires for food, shelter, love, and community—belong entirely to this deterministic, sensible world. They are part of nature's causal order, not the expression of a free will. Practical reason, on the other hand, demands that we are free, autonomous agents, capable of making choices for ourselves that are not merely the product of prior causes. However, this freedom to act from duty can never be captured theoretically. We are thus caught between two perspectives: one sees ourselves as pre-determined objects in a scientific universe subject to inclinations outside of our control, while the other experiences ourselves as free subjects in a moral cosmos who have control over our own actions. Kant was driven to heal this rift between nature and freedom.

The key to understanding Kant's unification of the practices of theoretical and practical reason lies in analyzing a concept central to both: the Idea. Kant lays out his conception of the Idea in his reading of Plato. Kant begins his discussion by remarking

---

[23] As we will see, they are united in the Highest Good which prioritizes practical over theoretical reason (see section III). Many read Kant's *Critique of the Power of Judgment* as harmonizing freedom and nature, both in biology and aesthetics. See Chapter 3 for how such a unity is spelled out in biology.

that Ideas go far beyond anything we can encounter in experience. He highlights that Plato "found his ideas preeminently in everything that is practical, i.e., in what rests on freedom" (A314-5/B371). The Idea of virtue, for instance, is not to be drawn from experience, which can only give "an ambiguous non-entity, changeable with time and circumstances, useless for any sort of rule" (A315). It can never be refuted from adverse experience because it is not derived from experience. Rather, it guides experience: "it is only by means of this idea that any judgment of moral worth or unworth is possible" (B372). To elucidate this, Kant claims that Plato's republic is "at least a necessary idea, which one must make the ground not merely of the primary plan of a state's constitution but of all the laws too" (B373). Although it may never come to pass, it guides the "legislative constitution of human beings ever nearer to a possible greatest perfection" (B374). Ideas hence act as comprehensive norms that guide our practices towards a common goal. Without these Ideas, our practices would be nothing—just aimless, disorganized physical motions. The practice of pottery, for instance, would be nothing without norms delineating both what a good pot and what a good potter is. Ideas are what give practices (like pottery) intelligibility and orientation. Kant thinks we ought to follow Plato in unpacking the normative frameworks behind these practices by emulating his "spiritual flight" which "considers physical copies in the world order, and then ascends to their architectonic connection according to ends, i.e., ideas" (B375). Kant does not intend to emulate Plato in ascending to the metaphysical forms behind all sensible experience. Instead, in critical inquiry we ought to ascend from the mere physical, sensible world to cognize the comprehensive *normative contexts* that make our experience within that world possible.

But if Ideas are not derived from experience, where do they come from? Against the psychologistic reading of Kant, Kant clearly states that they are not innate mental structures (by no means are they "figments of the brain" [B371]). Against the noumenalist readings, Kant clearly claims that they are not located beyond mind and matter in a metaphysical 'third realm' (a point where Kant is quite critical of Plato's treatment of Ideas as metaphysical substances [A314/B371]). For Kant, giving a spatial location to an Idea is a category mistake. Ideas do not exist in the natural or some metaphysical world. Instead, they originate in the normative realm: what we "ought to do," not from what "is done" (A319). Grasping a comprehensive standard—like the Idea of virtue—does not require positing a psychological, physical or metaphysical entity in space or time. Even if no perfectly virtuous person has ever existed, this does not mean that the normative pull to become one is any weaker. Furthermore, these norms are not merely "in our heads" as psychological facts; rather, they actively organize our mental lives. When guided by the Idea of virtue, for instance, we shut down impulsive thoughts to steal, organizing our mental states in a virtuous way. Therefore, norms exist neither as entities in the external world nor as mere psychological states we happen to have; they exist in their capacity to guide our practices, how we conduct our lives towards meaningful ends. Kant names this guiding faculty 'reason,' the very source of Ideas.

An Idea is defined by its power to orient action. Without this, it remains an empty abstraction. Consider a shipbuilder. One only becomes a shipbuilder through the act of building, and this act is generated by the governing norms of what a ship ought to be—a vessel that floats, navigates, and carries cargo. The builder does not merely "follow" this Idea like a map. Instead, they participate in it, where every hammer-strike and joint-sealing finds its meaning and purpose from the normative context provided by this Idea. The finished ship, on the other hand, is not just a copy of a blueprint stored in the Idea, but a successful attempt at following the normative guidelines set by the Idea. Therefore, the Idea is an active, not a passive, organizing principle. It functions as a constitutive norm: it defines the practice (*what* is being done), which in turn defines both the practitioner (*who* is doing it), and the product (*the end* that organizes the materials).[24] Crucially, Ideas do not create practitioners and products as if from thin air. Rather, they give meaning—meaning that qualifies otherwise indeterminate individuals as practitioners and otherwise indifferent materials as products, by uniting them in a practice oriented by a common end. Without this unification, there is no determinate agency, no meaningful product, and no intelligible practice—only an indeterminate capacity, mere stuff, and blind motion.

What are these practitioners and products that are given meaning? For Kant, the subject (qua end-in-itself) pre-exists[25] any particular normative practice—but only as an indeterminate capacity for free agency, not as a fully formed agent.[26] Reason (through

---

[24] To use Kantian terms, the Idea does not constitute the 'noumenal' subject nor the thing-in-itself, but the empirical subject and object. It gives the empirical subject and the object a meaning, outside of which they are literally nothing.

[25] Here is where Hegel might develop Kant's Ideas in a way clearer than Kant himself did, where noumenon are not separate from the Idea, but are moments in the unfolding of the Idea. The noumenon qua indeterminate capacity is merely the Idea in its abstract moment, where the Idea itself is the process of its own actualization.

[26] The literature on Kant's conception of the noumenal subject is deeply divided, but the dominant positions share a commitment that my view rejects. Metaphysical realists like Julian Wuerth argue that the

Ideas) demands that the subject become an agent of a particular kind. The same holds for the object side. The indeterminate something that serves as the object of our activity pre-exists, but it has no positive characteristics of its own. The Idea demands something of the world: that it show up as a field for cognition, a domain for moral action, or a collection of meaningful materials for building. But the response is not guaranteed. The clay may crack, nature may refuse to yield its laws, or a moral action may remain unrealized. Whether or not[27] nature yields to our demands, it is determined by the frameworks demanded by the Idea.

This process, whereby participation[28] in an Idea-constituted practice simultaneously brings into being the meaning of both a certain subject and the object of

---

noumenal self is a simple, enduring substance with intrinsic powers—a "thinking thing" that grounds the unity of consciousness and exists independently of its mental states (Wuerth, 2014). Deflationary or epistemic interpreters such as Henry Allison and Patricia Kitcher contend that the transcendental subject is merely a logical or psychological condition of thought—a formal unity of apperception with no ontological significance beyond its function in unifying experience (Allison, 1996, 2004; Kitcher, 2011). Neither camp treats the noumenal subject as essentially practical. Allison and Kitcher locate it entirely within theoretical cognition; Wuerth locates it in a metaphysical realm indifferent to practical concerns. By contrast, my reading follows a different tradition, exemplified by Karl Schafer and Lucy Allais, which treats the noumenal subject as a practical posit—not a knowable substance nor a mere cognitive structure, but the indeterminate capacity for free agency that we must take ourselves to be insofar as we act under the idea of freedom (Schafer, 2018, 2023; Allais, 2015). For Schafer, our practical awareness of ourselves as free possesses the marks of cognition on distinctively practical grounds, revealing the self as a capacity for rational self-determination. For Allais, autonomy is the anchor point of Kant's account of freedom, and what we are as noumenal beings is fundamentally a matter of our capacity to respond to reason and act with others. This capacity pre-exists any determinate normative practice, but it is not a fully formed agent; it becomes such only through participation in Idea-guided practices.

[27] If nature does not yield to a particular scientific framework, it is still articulated in terms of the theory, only negatively

[28] That reason is not a property like sight or taste but something we 'participate in' through practices of using reasons to achieve our ends finds a parallel in Stephen Engstrom's (2013) reading of Kant's synthetic unity of apperception. Engstrom thinks that synthesis (combining our representations) is not a secondary operation performed on pre-existing materials; rather it is the primary act that makes representations possible. This synthesis presupposes the higher unity of the synthetic unity of apperception, since all knowledge requires holding diverse representations together in a single self-consciousness. The act of unifying them just *is* the understanding—therefore we do not have an understanding, we enact it every time we think.

Engstrom's language of participation shares much in common with the view advanced here which identifies the synthetic unity of apperception with the *Idea* of the understanding. I go one step forward by its activity, is Kant's implicit thesis of *bi-constitution*. To fully grasp bi-constitution, however, we must distinguish two levels at which it operates. The first is the level of 'actual' bi-constitution. This is the concrete moment of action. When a potter, guided by the Idea of a pot, shapes this lump of clay into this vessel, the potter becomes an actual potter, and the clay becomes an actual pot through this act. The bi-constitution here happens in and through the activity itself. But this actual level presupposes a second, deeper level: 'horizonal' or 'potential' bi-constitution. Before any specific act occurs, the Idea of a pot discloses an entire world of meaningful possibility where clay can be shaped, tools can be used, markets can sell and buy pots/equipment, and a traditions of sculpting can be learned. Similarly, the Idea of the pot forms a normative identity of the 'potter' who is accountable to the standards internal to the practice of pottery itself. At this horizonal level, what is bi-constituted is not a specific subject and object but rather the form of agency that can inhabit the practice (the purposive subject) and the domain of significance within which objects show up as relevant to the practice (the purposive world). From this perspective, we can judge whether the 'actual' potter and the 'actual' pot are good or not by comparing them to their ideal counterparts as established by the horizonal constitution.

One becomes a potter by adopting the normative identity demanded by the Idea. This normative identity is defined by standards that distinguish success from failure,

claiming that this Idea—as constitutive of the practice of cognition which renders intuitions intelligible—bi-constitutes us as knowers and nature as cognizable. Furthermore, whereas Engstrom focuses on theoretical reason, I argue that this practice is ultimately subordinate to our moral vocation, which gives purpose to all our rational activity and provides constitutive grounds for positing what theoretical reason cannot establish: namely, for our purposes, the self-organization and purposive unity of the natural world.

participation from mere imitation. The potter is not someone who happens to shape clay, but is someone who can succeed or fail *as* a potter. A good potter produces vessels that are functional, durable, and well-formed, while a bad potter produces cracked or misshapen vessels. This normative identity, in turn, determines what capacities matter for the purposive subject. Capacities are not natural endowments that pre-exist the practice, but are called into being by the norms of the practice. The potter requires the capacity to judge the consistency of clay, to coordinate hand and wheel, to anticipate how a form will hold under firing. One does not first possess manual dexterity and then apply it to pottery. Rather, dexterity becomes the capacity for potting only insofar as the subject is oriented by the norms of pottery. The same holds for other purposive subjects: a scientist is defined by norms of empirical adequacy, explanatory power, and the capacity to evaluate evidence; a moral agent is defined by norms of autonomy, respect for persons and the capacity for self-legislation. In each case, the Idea constitutes a normative identity that one can inhabit more or less successfully, equipped with the specific capacities that the practice itself generates.

Correlatively, horizonal bi-constitution also yields the purposive world. This is the field of affordances and meanings that the Idea makes salient. The Idea of a ship does not govern only shipbuilders and their vessels. It governs an entire practical world of tools, trade routes, weather patterns, and navigation techniques. This world is neither subjective nor objective—it is the normative space of potential action opened by the Idea, determining what counts as relevant and what counts as a problem or solution. A meaningless mass becomes timber by being a solution to the problem of shipbuilding; a stretch of water becomes a navigable route by virtue of the practice of seafaring. The purposive world is thus the horizon of significance within which objects can appear as meaningful for the agent's activity, just as the purposive subject is the normative identity the agent inhabits. Together, they are the two poles of horizonal bi-constitution.

The two levels of constitution—actual and horizonal—are not opposed but mutually reinforcing. The horizonal level makes actual acts intelligible as instances of a practice. That is, a particular shaping of clay counts as potting only because the action participates in the Idea of pottery. Conversely, actual acts continually renew and articulate the horizonal field of possibilities. Every pot successfully made reinforces the norms of the practice, and every failed pot forces a revision of what the practice demands. This mutual reinforcement, however, ensures that the pursuit of Ideas initiates an infinite, self-perpetuating process (see chapter 4). Ideas are unconditioned—they represent complete, absolute wholes—while our existence and our engagement with the world remain conditioned by time, matter, and circumstance. This gap ensures that we can never fully embody an Idea. We may participate in the Idea of a 'pot,' for example, but we will never become the perfect potter or craft the perfect pot. Every error or limitation we encounter forces a revision of our understanding and practice, which in turn drives a further articulation of the Idea. Thus, the Idea functions not as a static goal, but as an active, generative demand to perform an action a certain way. It perpetually motivates the very actions that aim to realize it, while guaranteeing that the coherence it promises can be approached but never fully attained.

Having seen the general, bi-constitutive characteristics of the Idea, we can now analyze the specific mechanisms of bi-constitution in Kant's account of theoretical and practical reason. This examination will accomplish two things: first, it will reveal how Kant's system bridges the gulf between the theoretical world of efficient causality and the practical world of freedom by grounding the former in the latter. Second, it will set the stage for Kant's argument (in section IV) that this very synthesis produces the objective reality of a self-organizing world—a concept that emerges as necessary from the fact of our practical reason. For our free, moral projects to be possible, the world cannot be a rigid, alien mechanism; it must be a system receptive to and organized by the ends we legislate.

III. Twofold Employment of Reason

How can Kant harmonize the self-organization of nature—the cognition of which is the ultimate end of theoretical reason—with the organization of ourselves as rational moral agents? The answer lies in a deep structural parallel: both cognition and moral action are guided by an Idea. Despite governing different practices (one aiming to know what is, the other aiming to determine what ought to be done), theoretical and practical reason share an isomorphic bi-constitutive logic. Theoretical reason is guided by the Idea of the understanding[29], while practical reason is guided by the Idea of autonomy. In both,

[29] I am speaking generally here: it is more accurate to say that intelligible cognition is guided by the Idea of understanding. Yet *comprehensive* cognition is guided by other Ideas (God, the Soul, the World). Practical reason too is supplemented by other Ideas.

the pursuit of an immediate object at the actual level (a scientific law, a moral action) simultaneously constitutes the normative identity of a subject capable of that pursuit at the horizonal level (a scientist, the demanded autonomous self for the phenomenal self), discloses a purposive world (for theoretical reason: nature as a system of shareable, law-governed phenomena; for practical reason: nature as amenable to moral ends), and points toward a total, comprehensive object (a unified theory of nature, the highest good) that serves as the ultimate target within that world. This structural identity makes their harmonization conceivable.

The following analysis traces this logic through Kant's system. First (A), it delineates the bi-constitutive machinery of theoretical reason to show why its comprehensive Idea remains merely regulative. It then (B) examines the transcendental errors that arise when this limit is ignored. Finally (C), it develops the bi-constitutive structure of practical reason before demonstrating how it grounds theoretical reason. Section IV will unpack the consequences of this grounding relation, showing how practical reason (and our power of judgment) founds the purposive, self-organizing unity of nature that theoretical reason depends on but fails to establish.

A. Theoretical Reason

The *Critique of Pure Reason* investigates the question, What can I know? As rational inquirers, we are driven toward investigating fundamental features of our lives—the nature of the soul, the structure of the cosmos, the existence of God. Yet to responsibly engage in these pursuits, Kant argues, we must first undertake a critique of our own cognitive faculty. We must unpack the very aptitudes and valid norms that make knowledge possible. This critical self-examination is, for Kant, the task of explicating what he calls the Idea of the understanding.[30]

As we saw, an "Idea" is not a vague notion but a precise technical term. An Idea of reason is a concept of an unconditional totality that transcends all possible experience. The Idea of the understanding acts as a normative, comprehensive guide for cognition, constituting both a form of agency (the normative identity of the inquirer) and a purposive world (the domain of significance within which inquiry becomes meaningful). The Idea of the understanding hence guides our *epistemic* practices in a systematic and holistic way. It is an "idea of the whole of the a priori cognition of the understanding" (A64/B89). For Kant, the goal of the understanding is to interpret the sensible world (A51/B75). However, the sensible world presents itself to us from a multiplicity of perspectives and standpoints, yielding a chaotic manifold. We would be unable to achieve coherent experience at all if these perspectives were not unified systematically in a single, comprehensive standpoint. This standpoint, which Kant calls the 'synthetic unity of apperception'—is the normative identity we must participate in if we are to acquire understanding. It ensures that the diversity of intuitions can be combined into unified objects of experience, that the past and future are connected in a single temporal consciousness, and that different spatial locations are all locations within the same unified

[30] Kant claims that "reason thus prepares the field for the understanding" (A657/B685), refuting the common view that understanding precedes reason, held by Eckart Förster (2000, 148-150).

space. If we participate within this standpoint with others, we can acquire knowledge not only of other's perspectives but of objects and laws that pertain to all perspectives. This makes possible scientific knowledge by disclosing a purposive world in which objects are not mere private sensations but shareable, law-governed entities.

Because the synthetic unity of apperception is a priori and can never be given in experience, it can only be grasped through an Idea. The whole of our cognition, says Kant,

> will constitute a system that is to be grasped and determined under one idea, the completeness and articulation of which system can at the same time yield a touchstone of the correctness and genuineness of all the pieces of cognition fitting into it (A65/B90)[31]

Here, the "one idea" is the Idea of the understanding itself—the blueprint or normative archetype of our cognitive system that guides our practices of theoretical knowledge. The "whole of this cognition" refers to the system of a priori concepts that are the "pieces" of the system. Their validity and interrelation are not discovered empirically but are derived from the necessity of their normative function in reaching the end of cognition. Kant compares the systematicity of theoretical reason itself to an organic body:

> [It] contains a truly articulated structure of members in which each thing is an organ, that is, in which everything is for the sake of each member, and each individual member is for the sake of all, so that even the least frailty, whether it be a mistake (an error) or a lack, must inevitably betray itself in its use (Bxxxvii-iii)

To articulate the Idea of the understanding is, therefore, to unpack the central functions of the organism that is our theoretical reason—to specify the normative identity of the knower and the structure of the purposive world of objects that such knowing discloses.

[31] Kant further emphasizes that this unity is a "unity of the theme in a play, speech, or fable" (B114)

Kant's 'Transcendental Analytic' identifies judgment as the fundamental act of thought—the active process through which the raw material of sensible intuition is interpreted or 'synthesized' into a unified representation that can be an object of knowledge. Judgment does not interpret sensory experience passively using contingent, habitual modes of thinking; it is the spontaneous activity of subsuming particular intuitions under universal rules. These universal rules are the pure concepts of the understanding: the categories.

Crucially, categories like substance and causality are not free-floating concepts we innately possess and then apply to the world (as the impositionist might suggest). They are, Kant argues, the fundamental and invariant norms of judgment itself, and thereby the necessary forms of any possible objective experience. They are the rules we must use whenever we make a determinate claim about an objective world, rules that we must follow to be intelligible. For instance, in judging that "the sun warms the stone," we do not observe two independent events—sunshine and a stone heating up—and then infer a causal link between them. Rather, by participating in the rule of causality (and, by extension, taking on the normative identity of the synthetic unity of apperception), we actively constitute them as a causally related event. The category does not describe a connection we find; it prescribes the normative rule through which a connection first becomes objectively intelligible. The category is the rule of synthesis. As Kant explains the objective validity of the categories:

> Concepts of objects in general lie [categories] at the ground of all experiential cognition as a priori conditions; consequently the objective validity of the categories, as a priori concepts, rests on the fact that through them alone is experience possible (A93/B126)

But how do we discover *which* categories these are? Kant's method is a "transcendental deduction," which aims to establish the "entitlement" (*Befugnis*) of the categories (A84/B116). He argues we cannot find them empirically or haphazardly, as Aristotle attempted. While praising Aristotle's effort as "worthy of an acute man," Kant criticizes his list as rhapsodic, assembled without a principle (B107).

The guiding principle for discovering the categories is the Idea of the understanding itself—the synthetic unity of apperception. This Idea demands a logically exhaustive table of the fundamental forms of judgment.[32] Kant therefore starts from the accepted logical forms of judgment (universal, singular, etc.) and asks: what corresponding synthetic function must the understanding possess to employ each form of judgment to determine an object? The transition from the logical form of judgment (the way we combine representations in thought) to the transcendental category (the way we think an object as determined by such combination) yields the complete, systematically derived table of categories. This derivation shows the categories are not a random list but the revealed cognitive norms necessary for valid judgment. They are the indispensable organs that function within the larger organic system of the understanding, each playing a necessary role in achieving the end of that system: a valid cognition.

[32] This derivation has been the subject of interpretive debate. Henry Allison (1996) argues that the principle of apperception is analytic, a logical precondition rather than an active, normative Idea. In my interpretation, by contrast, the Idea of the understanding acts as a generative norm—not merely a condition that representations must satisfy, but sets an end (interpreting the manifold of intuition) that orients the activity of synthesis itself.

The systematic derivation of the categories culminates in Kant's insight that the understanding operates through a single, unified process of bi-constitution. This logic hinges on unpacking the transcendental conditions for constitution, as well as between the horizonal and actual levels of constitution. At the transcendental level, Kant identifies the transcendental subject, the "I think," as the pure, logical unity of consciousness that must be able to accompany all representations (B131-132). This is not a knowable substance or a person; it is merely the formal condition for the possibility of any unified experience. Correlatively, the transcendental object, the "=X" (A109), is the mere concept of an object in general to which our cognition is directed—an indeterminate placeholder without positive content. Crucially, these transcendental conditions are never themselves constituted, but act as poles towards which bi-constitution is directed.

At the horizonal level, the Idea of the understanding—which is the synthetic unity of apperception—constitutes the normative identity of the knower as such. The Idea discloses the form of agency proper to the practice of knowing: the normative identity defined by the end of knowledge, equipped with the capacities for judgment, and subject to standards of coherence. Correlatively, the Idea discloses a purposive world—a domain of significance within which objects appear as interpretable through the categories, and thus as law-governed, shareable, and open to scientific inquiry. This is the significance of the often-misunderstood claim that "the conditions of the possibility of experience in general are at the same time conditions of the possibility of the objects of experience" (A158/B197). The conditions—the categories and the unity of apperception—are not tools applied by a finished subject to a finished world. They are the rules of a constitutive activity whose performance generates the horizons of the relata themselves.

Finally, at the actual level, when the manifold of sensory intuition is synthesized using a valid form of connection—for instance, the category of causality—an indeterminate sensory given is judged to be a determinate, law-governed object of experience, such as a causal event. The category provides the rule that actively constitutes the empirical object as having a specific objective form and relation. What this means is that this object abides by the universal forms of the understanding, such that others who also participate in the norms of the understanding would also understand it in the same way. Simultaneously, this act of synthesis unifies the manifold of inner sense—the temporal flow of conscious states—in a singular consciousness. The indeterminate stream of perceptions thereby coalesces into the coherent experience of a single, knowing subject undergoing a particular experience. This means that others who partake in the norms of the understanding comprehend not just the same object, but the mutual ways in which we—as knowers— are judging the object.

While the understanding's bi-constitutive activity successfully establishes the framework for universal, shareable, and objective experience, this does not exhaust the demands of theoretical reason. The understanding provides the rules for determining specific causal relations and substantial unities—it tells us that every event has a cause, that every substance persists, and ultimately yields a certain universal perspective through which all interpretation of the sensible world becomes intelligible. This still leaves open the possibility that the particular laws governing phenomena could be so diverse and heterogeneous that they resist unification into a singular theory of nature. Yet science requires not just lawfulness but systematic unity among laws. To achieve this, reason, in its regulative employment gives "to the understanding... unity a priori," a unity called "the unity of reason," and which is "of an altogether different kind than any unity that can be achieved by the understanding" (A302/B359). In a pivotal metaphor, Kant describes reason as the faculty that "prepares the field for the understanding" by guiding it toward comprehensive systematic goals (A657/B685). Reason does this not only be generating the Idea of the understanding, but drives us to seek the unconditioned—to push beyond any given condition toward the totality of conditions, and thereby to unify the understanding's cognitions into the most coherent possible system.

To execute this unifying function, theoretical reason utilizes Ideas that direct our understanding towards satisfying our reason's deepest curiosities. These Ideas—most fundamentally, the psychological Idea of the Soul, the cosmological Idea of the World, and the theological Idea of God—are not concepts of possible objects of experience. They are concepts of absolute totalities that transcend all possible empirical knowledge. Yet since they transcend empirical knowledge, they do not serve as knowable objects but as indispensable regulative principles for our scientific inquiry. These Ideas—like all Ideas—guide practices. Instead of the practice of valid cognition in general, they guide psychology, cosmology, and theology. Kant argues that these Ideas open up inquiry: we "question nature according to these ideas" (A645-46/B673-74).[33] This relationship is

[33] This active stance has led some commentators to attribute a stronger role to the ideas than mere heuristic guidance. Robert König (2022) argues that the Transcendental Dialectic has a "positive and founding character for cognition" and that ideas have "indirect objective reality" as a procedure—"i.e., an action"—rather than as properties of objects. For König, the ideas manifest themselves as "an action full of

captured in his famous analogy that reason must not approach nature as a pupil who listens to whatever the teacher chooses to say, but "like an appointed judge who compels witnesses to answer the questions he puts to them" (Bxiii). The Idea of the Soul, for instance, guides psychology by demanding we seek the ultimate, unconditioned unity of mental phenomena —thereby constituting the normative identity of the psychologist as a specific form of inquirer and the psyche as a purposive domain of investigation; the Idea of the World directs cosmology toward the totality of conditions in space, time, and causation —constituting the cosmologist and the cosmos as a unified field of inquiry; the Idea of God posits the ultimate ground of all contingency, providing a unifying principle for the entirety of scientific investigation —constituting the theologian and the divine as a regulative horizon for all inquiry.

Kant adds that the regulative function of the theoretical Idea is supplemented by the fundamental a priori interests of reason, which structure all inquiry: the principles to seek unity (homogeneity), to honor manifoldness (specification), and to presuppose affinity (continuity) between concepts (A662/B690). In contemporary terms, any scientific paradigm—be it in physics, ecology, or neuroscience—is guided by a specific "Idea," conceived from the "vantage point of a certain general interest" (A834/B862). Yet within physics, for instance, one can seek the unified laws of nature (homogeneity), particular laws that govern, e.g., the flow of water (specification) and how the, e.g., laws of water fit

purposiveness, maxims, and imperatives within cognition". This emphasis on Ideas as action resonates with my practice-based reading. However, König focuses primarily on how this "action" bridges theoretical and practical reason; my contribution is to specify the mechanism of that action—namely, bi-constitution, whereby participation in an Idea simultaneously constitutes the subject and the object of the practice through acting within such practice.

into an overarching theory of nature (continuity). Together, these Ideas and principles systematically guide the understanding, ensuring that even the varied sciences of today cohere within reason's overarching drive for a unified whole.

These Ideas possess a fundamentally different status from the categories of the understanding. While the categories are constitutive of experience itself—that is, they make empirical objects and our own identities as knowers intelligible—these Ideas are merely regulative for knowledge. From a strictly theoretical standpoint, the content of Ideas can never be fully known or demonstrated. They are what Kant calls "problematic concepts," and their use is "hypothetical" (A647/B675). He explains that the "hypothetical use of reason... is not properly constitutive," because we can never derive all possible consequences from an assumed Idea to prove its objective reality." "Rather," he continues, "this use of reason is only regulative, bringing unity into particular cognitions as far as possible and thereby approximating the rule to universality" (A647/B675). The Ideas suggest certain patterns to care about, projecting the form of a systematic whole, but they can never guarantee that such a whole actually exists. Consequently, from the vantage point of theoretical inquiry alone, the ultimate self-organizing unity of nature, qua Idea, remains a necessary presupposition of our questioning, a "problem without any solution" (B384).

The fact that theoretical reason's highest end is forever relegated to the status of a regulative ideal creates the need for, as we will see, practical reason to intervene, which treats this same unity not as a mere heuristic guide but as a constitutive postulate of the moral will. As he details in the "Architectonic of Pure Reason," our cognitive interests are ultimately subordinated to other human ends. Our ultimate interest of reason, he argues, is not theoretical but practical. The speculative use of reason finds its true value in "preparing the field for the practical" and is intentionally subordinated to our fundamental interest in the highest good—the realization of "general happiness" in a moral world (A851/B879). This subordination is necessitated by the very limitation we have traced: theoretical reason can only regulate the search for nature's ultimate unity but can never know such unity. The highest good, however, as the necessary object of practical reason, demands that we postulate the very systematic and self-organizing unity of nature that theoretical reason could never secure on its own.

Thus, the limit of theoretical knowledge is not a mere negation but has "made room" for the "extension" of pure reason through "practical data of reason" (Bxxi-Bxxii). It transforms from a barrier into an opening, making way for a necessity of a different order. To understand how this practical constitution is possible, and how it rescues the unity of reason from its theoretical dead end, we must first examine the precise nature of the errors that arise when this theoretical limit is ignored.

B. Theoretical Shortcomings

Kant's Transcendental Dialectic exposes a fundamental illusion of pure reason. When the Ideas of the Soul, the World, and God are hypostasized—that is, mistakenly treated as actual, knowable objects that constitute the unconditioned bedrock of reality—they lead reason into inevitable and systematic error. These errors all stem from the misapplication of the categories of the understanding beyond their only valid function, which is interpreting sense-experience. The Dialectic demonstrates the systematic illusions that arise when this valid context is discarded. When Reason mistakes the logical "I think" for a substantial soul, projects the demand for totality into a knowable cosmos-in-itself, or hypostatizes the ideal of systematic unity into a demonstrable God, it transgresses the bounds of possible experience. This transgression throws reason into the inescapable contradictions of the paralogisms, antinomies, and speculative theology, generating false knowledge instead of absolute truth.

Pertinent for our purposes is the systematic unity of nature that the Idea of God captures. The section entitled 'Ideal of Pure Reason,' which concludes the Transcendental Dialectic, diagnoses the fundamental error of hypostasizing the Idea of God. In this theoretical context, "God" represents the ultimate terminus of all explanation of the natural world: the necessary, uncaused cause of the cosmos and the ultimate unity underlying all contingent existence. This is Kant's name for the implicit, final target of any scientific quest for a complete "theory of everything," such as that proposed by complex systems scientists claiming that nature can be described entirely as a unified, self-organizing whole. As such, this Idea represents the inevitable horizon of all theoretical inquiry.

Kant's analysis reveals that the traditional proofs for God's existence share the same flaw: they all depend on the ontological argument's error, which is to treat existence as a real predicate. Human reason, in its drive for complete explanation, first constructs the logical idea of an *ens realissimum* (a most real being) as the ground of all possibility. The transcendental illusion occurs when reason hypostatizes this logical ideal, erroneously counting "existence" among its perfections, thereby generating the concept of a necessarily existing God. This illicit move is key because existence is not a predicate that adds a new determination to a concept (like "omniscient" or "good"). Rather, it is merely the illegitimate positing of the concept itself in relation to experience. Consequently, the cosmological argument, which seeks a first cause, and the physico-theological argument, which infers a designer, both ultimately rely on this flawed ontological leap to convert a contingent cause or a cosmic architect into a necessarily existent being. By exposing this shared foundation, Kant demonstrates that theoretical reason is incapable of attaining knowledge of God, as it inevitably transgresses the bounds of possible experience.

This critique of the theological proofs reinforces a crucial distinction: while Ideas like that of God cannot serve a 'constitutive' function for knowledge, they remain indispensable as 'regulative' principles. The Idea of God, for instance, regulates scientific inquiry toward a unified theory of nature, encouraging different fields to cohere into a single, systematic whole. However, a scientist who claims that "the cosmos is a self-organizing system" based on generalizing the results of their science commits the same category of error as the rational theologian. They are illicitly promoting a regulative Idea—the goal of cosmic unity—to the status of a constitutive truth about the world-in-itself. In doing so, they use reason's necessary lens (the command to 'research as if all natural laws were unified in a system') to make a dogmatic claim about what the world is. This is a classic transcendental error, where they mistake a necessary condition of thought for a discovered feature of reality itself.

C. The Primacy of the Practical

Kant argues that practical reason has primacy over theoretical reason.[34] I interpret this as claiming that practical reason *grounds* theoretical reason. To make this clearer, we can distinguish three distinct notions of ground. In a footnote in the *Critique of Practical Reason*, Kant distinguishes two types of grounds: *ratio essendi*—the grounds that make something what it is (the ground of its being or possibility)—and *ratio cognoscendi*—the grounds that make something known to us (the ground of our cognition of it) (5: 4n).[35] The *ratio essendi* of theoretical reason's Ideas would be the actual conditions that make them true—the real grounds in virtue of which the world is intelligible, coherent, and amenable to our cognitive efforts. On its own, theoretical reason can only presuppose that the world is such that our concepts can grasp it. It can never grasp what underwrites this match. The *ratio cognoscendi* for its regulative Ideas would be genuine cognition—the assurance that our cognitive endeavors track something real, not mere heuristic fictions. To Kant's two types of grounds, we may add a third grounding role—*ratio finalis*—which establishes why we should pursue theory at all, why the end of knowledge matters to begin with.

[34] Thus, in the union of pure speculative with pure practical reason in one cognition, the latter has primacy, assuming that this union is not contingent and discretionary but based a priori on reason itself and therefore necessary" (5: 121). For the primacy of practical reason see Neiman (1994) and O'Neill (1990).
[35] "Whereas freedom is indeed the ratio essendi of the moral law, the moral law is the ratio cognoscendi of freedom. For, had not the moral law already been distinctly thought in our reason, we should never consider ourselves justified in assuming such a thing as freedom (even though it is not self-contradictory). But were there no freedom, the moral law would not be encountered at all in ourselves" (5: 4n)

Practical reason plays the role of *ratio finalis*, *ratio credendi* (not fully *ratio cognoscendi*), and *ratio essendi* for theoretical reason. That is, the fact of reason and its ultimate object, the highest good, provide a threefold grounding function. First, as the final end that gives all rational activity its point, they provide the *ratio finalis* that makes theoretical inquiry meaningful: because realizing the highest good requires knowledge of nature, theoretical cognition acquires an ultimate significance beyond its own internal end. Second, the highest good demands conditions for its possibility—the postulates of God and immortality—which serve as the *ratio credendi* that warrants our belief that nature is systematically unified and amenable to our moral ends. Unlike a *ratio cognoscendi*, which would imply theoretical cognition of nature's unity and our own immortality (and is therefore impossible, given the opacity of Ideas of reason), the *ratio credendi* provides practical, non-cognitive warrant—rational faith (*Vernunftglaube*)—that is subjectively sufficient for moral action but does not amount to objective theoretical knowledge. Third, these same postulates function as the *ratio essendi*—the real ground that nature's purposive unity objectively obtains, making it the kind of world where the highest good can be realized. In what follows, we will trace how the fact of reason and the highest good exercise each of these grounding functions, while spelling out the general structure of practical reason's bi-constitution.

1. Ratio finalis

Kant establishes the primacy of the practical via the "fact of reason"—our immediate consciousness of an unconditional "ought" that directs us to act in moral ways (5:31; 5:47). This fact demands an explanation: how can the moral law possess such authority over us? The Idea of autonomy[36] (5:33) provides the answer by revealing the bi-constitutive structure of practical reason.[37] At the horizonal level, the Idea of autonomy constitutes the normative framework of morality by defining the form of agency in which the will gives the moral law to itself—acting from duty and subject to the standard of universalizability. Correlatively, it constitutes the purposive world of morality where events call on us to intervene and the highest good appears as the necessary object of moral striving.[38] At the actual level, through concrete moral actions guided by this Idea,

[36] Kant calls autonomy the "formal supreme principle of pure practical reason" (5:39). This principle is the formal expression of what the Idea of autonomy means: a will that gives itself the law, independent of any material incentives, or as Kant puts it in the *Groundwork*, "the idea of the will of every rational being as a will giving universal law" (4:431).

[37] Kant seems to identify autonomy with freedom: "the moral law expresses nothing other than the autonomy of pure practical reason, that is, freedom, and this is itself the formal condition of all maxims, under which alone they can accord with the supreme practical law" (5:33). We can therefore call it the Idea of Freedom, with the caveat that we are using it in a practical sense and not in its theoretical function. He writes that the idea of freedom "is, however, absolutely impossible to give anywhere in experience an example of it" (5:49); we cannot theoretically cognize it. Yet pure practical reason "fills this vacant place with a determinate law of causality in an intelligible world (with freedom), namely the moral law" (5:49). Thus, the Idea of Freedom receives objective reality not through theoretical insight but through the moral law itself, which gives it "undoubted reality" though "only practical" (5:49). The Idea of Freedom is not a cognitive claim about a noumenal object but the practical condition under which we must act as if we were free, with the moral law providing the determinate content that speculative reason could never secure on its own.

[38] "For, the moral law in fact transfers us, in idea, into a nature in which pure reason, if it were accompanied with suitable physical power, would produce the highest good, and it determines our will to confer on the sensible world the form of a whole of rational beings. The most ordinary attention to oneself confirms that this idea is really, as it were, the pattern for the determinations of our will" (5:43).

we constitute ourselves as determinate moral agents and contribute to the realization of the moral world.

Crucially, the Idea of autonomy does not *create* our status as ends in ourselves. The noumenal self—the indeterminate capacity for freedom—already is an end in itself and already possesses dignity (rather than a 'price' [4:434]). This status belongs to our intelligible character and does not depend on any Idea. What the Idea does is make this status intelligible to us as the ground of obligation. It reveals the gap between what we already are (ends in ourselves, at the noumenal level) and what we actually do (our empirically conditioned actions, at the actual level). The moral law's authority is the authority of our own noumenal status over our empirical self. We do not aspire to become ends in ourselves; we aspire to act in a manner worthy of the fact that we already are ends in ourselves, capable of freedom. The Idea of autonomy does not constitute our noumenal status, but constitutes the normative framework through which we can recognize, honor, and act in accordance with that status. By acting guided by this Idea, we close the gap between our noumenal status and our actual performance, participating more fully in what we already are as autonomous ends-in-themselves. This noumenal status is not our 'true self' that we return to, but the capacity for free self-legislation.

This brings us to a crucial asymmetry between the subject and object sides of practical bi-constitution. On the subject side, the noumenal self—the indeterminate capacity for freedom—already is free, already an end in itself, and already possesses dignity. This status pre-exists any moral practice. Through the moral law, we practically recognize this status and make it intelligible as the ground of obligation. The moral law reveals this status and calls us to act in accordance with this normative identity. On the object side, by contrast, the moral world does not pre-exist. It is not yet realized. Through the Idea of autonomy, we are obligated to bring about the moral world.

Kant calls the moral world the 'highest good.'[39] This is state of the world where "happiness [is] distributed in exact proportion to morality" (5:110). The moral law thus commands us to strive to realize this highest good, making it "a priori (morally) necessary to produce the highest good through the freedom of the will" (5: 113). Yet a fundamental problem arises from our finite nature. We cannot achieve perfect virtue ("holiness") in a finite lifespan, nor can we ensure that our actions are either efficacious or lead to happiness, since acting effectively depends not merely on our virtuous intentions but "knowledge of the laws of nature and the physical ability to use them for one's purposes" (5: 113). Given nature's apparent indifference, the moral project commanded by the very law that constitutes us seems, paradoxically, impossible to fulfill.

We can now start to see the relationship between theoretical and practical reason. Practical reason's task of promoting the highest good requires knowledge of nature. To preserve the conditions under which happiness is possible and to act effectively on our moral commitments, we must understand causal relations, predict outcomes, and grasp the systematic unity of the natural world. This requires precisely the kind of cognition

[39] Eric Watkins (2010) analyzes the antinomy of practical reason that gives rise to the highest good, showing how reason's demand for the unconditioned manifests as a conflict between virtue and happiness that only the postulates can resolve. Where Watkins emphasizes the logical resolution of the antinomy, I emphasize the ongoing practical, bi-constitutive function of the highest good as the horizon of an endless process. It is not a concept we have to believe in to avoid irrationality in practical deliberation, but rather a normative Idea that generates the ongoing practice of attempting to make the world a world it ought to be—a moral world—simultaneously constituting us as moral agents.

that theoretical reason pursues. Yet theoretical reason must remain autonomous in its own sphere. It cannot posit the highest good as its internal end, for doing so would corrupt scientific inquiry by subordinating the pursuit of what is to what ought to be. Practical reason does not replace theoretical reason's internal purpose. Rather, practical reason has primacy, integrating theoretical inquiry into reason's complete vocation, giving it ultimate significance. The scientist's pursuit of systematic cognition, pursued for its own sake, becomes meaningful within the larger project of realizing the highest good. Here the highest good functions as the *ratio finalis*—not as a premise within theoretical inquiry, but as the final end that orients the whole of reason's activity.

Because Kant thinks that "all interest is ultimately practical" (5:121), the ultimate end of practical reason—the highest good—gives direction to all our purposes, including our theoretical pursuits. Without this common purpose, we could not even trust one another's epistemic claims, making science (let alone basic communication) impossible. Hence, it is practical reason that makes theoretical reason's goals coherent. This is precisely what Kant has in view when he distinguishes, in the first Critique's Architectonic of Pure Reason, between philosophy in the scholastic concept (*Schulbegriff*) and philosophy in the world-concept (*Weltbegriff*). The former concerns "a system of cognition that is sought only as a science without having as its end anything more than the systematic unity of this knowledge" (A838/B866). The latter, by contrast, is "the science of the relations of all cognition to the essential ends of human reason" (A839/B867)—ends that are ultimately practical and concern "what necessarily interests everyone" (A839n/B867n). Philosophy in the world-concept thus aims at nothing less than the final end that unites all rational beings, the very same end that Kant elsewhere identifies with the highest good. Crucially, the world-concept itself delineates a domain of significance within which all cognitive pursuits acquire meaning by being oriented toward the essential end of humanity. Kant is very clear that the world-concept "has always grounded" the scholastic concept, and hence theoretical reason (A837/B866). This grounding is a *ratio finalis*: theoretical reason's autonomous pursuit of systematic cognition finds its ultimate significance only within this cosmopolitan conception of philosophy, where knowledge serves the essential end of humanity—the highest good.

2. *Ratio essendi* and *Ratio credenda*

The Idea of autonomy does not constitute our noumenal status. Rather, it constitutes the self we demand our phenomenal self to be—the normative identity our empirical self must adopt and approximate. This demand is mediated by the faculty Kant calls *Willkür* (the power of choice), which stands at the crossroads between the moral law (given by *Wille*, the legislative will of the noumenal self) and the inclinations of the phenomenal self. The moral law does not create our noumenal status but reveals it and, through the exercise of *Willkür*, demands that our phenomenal self approximate the self we demand it to be. This Idea can be approximated, but can never fully bring our phenomenal self in alignment to it in practice.

Action will always be tested against what Kant calls the categorical imperative, which reveals a procedure that we can use to see whether our maxims approximate the moral law. Every maxim we adopt is a provisional attempt to legislate ourselves and a piece of the world. [40] The categorical imperative subjects these attempts to a test of universalizability, forcing us to ask: "Would the world constituted by this maxim be a coherent, non-self-contradictory moral order?" When a maxim fails this test—when its universalization leads to a logical or volitional contradiction—it is rejected, compelling a revision of our proposed legislation. Often this contradiction is only realized after the fact, as we can rarely ever cognize the complete implications of our often-vague maxims.

Yet even when a maxim passes the categorical imperative test, a further question remains: are we acting from duty or merely in accordance with duty, driven by mere inclination? The moral law demands not just lawful conduct but conduct done for the sake of the law. The certainty that we act for the sake of the moral law is notoriously opaque to self-knowledge. We can never be certain that our motives are purely moral. Thus, (at least) two distinct obstacles block our phenomenal self from approximating the self we demand it to be: the difficulty of knowing whether our maxims are universalizable, and the difficulty of knowing whether our intentions are pure. In this way, the moral law provokes a continuous process of self-correction, driving us to endlessly refine our actions and intentions toward the ever-receding horizon of the autonomous self we demand ourselves to be.

---

[40] This emphasis on the ongoing function of autonomy distinguishes my reading from interpretations that focus primarily on the justificatory role of the postulates. Marcus Willaschek (2010), in his influential analysis of the primacy doctrine, examines how practical reason's needs justify theoretical propositions (the postulates) that reason cannot demonstrate theoretically. He argues that postulates are "theoretical propositions" that are "not as such demonstrable" but are "inseparable corollaries" of the moral law. While Willaschek's analysis is invaluable for understanding the epistemic status of the postulates, my reading focuses on a different dimension: the postulates do not merely justify beliefs but generate practices. The commitment to the highest good does not just license the belief in God and immortality; it structures the entire activity of moral striving, calling us to endless self-correction through the universalizability test.

For this endless striving be a rational enterprise, we must believe that we can achieve our moral self-perfection, as well as achieve the highest good. This requires two practical 'postulates.' First, because "complete conformity" with the moral law is only possible in an "endless progress," we must postulate the immortality of the soul to provide the necessary duration for this asymptotic approach to holiness (5: 122). Second, because we cannot establish the necessary harmony between the moral law and the laws of nature, we must postulate the existence of God as a supreme intelligence who can orchestrate this union, ensuring that happiness flows from worthiness to be happy. We must, in other words, presuppose that the natural world embraces our moral actions.

These postulates perform two distinct but inseparable grounding functions for theoretical reason. First, they serve as *ratio credendi* (ground of believing)—they provide a "pure rational belief" (5: 126), a warrant for believing that nature is systematically unified and amenable to our ends. They cannot serve as a *ratio cognoscendi* (ground of knowing) because such a ground would require a schema through which we could theoretically cognize nature's systematic unity. Ideas of reason, however, have no schema. We can never point to an intuition and say, "this is the unity that God guarantees." To claim otherwise would be to fall into moral enthusiasm, the delusion that we can directly perceive the moral order in the world. The opacity of the ground of moral action forbids any such cognition. Theoretical reason, left to its own devices, can only ever regulate its inquiries by the Idea that nature and its laws comprise a systematic unity. It can never *know* that this unity obtains. The postulates change this situation, but not by providing theoretical cognition. Instead, they secure the conditions under which the purposive world of scientific investigation can be taken as more than a mere heuristic projection. Because they are necessary conditions of the highest good—and because the highest good is the necessary object of a will bound by the moral law—we are *practically justified* in treating nature as unified and conducive to our ends. This is not a theoretical justification. We gain no speculative insight into how nature achieves this unity. We hence have no *cognitio*. But we acquire a warranted conviction (rational faith) that the world we investigate is in fact coherent, intelligible, and fit for our cognitive effort, something theoretical reason alone could never secure.

Second, these same postulates function as *ratio essendi*—they are the real ground that nature's purposive unity objectively obtains. This is a stronger claim than mere warrant. The postulate of God does not merely justify our belief that nature is systematic; it posits God as that in virtue of which nature is systematic. Just as freedom is the *ratio essendi* of the moral law despite being theoretically unknowable, so too God is the *ratio essendi* of nature's amenability to moral ends. From the practical standpoint, the highest good requires that nature be the kind of world where virtue and happiness exist in proportion to each other. The postulate of God is the condition that makes this possible—the real ground that underwrites the fit between our moral ends and the natural world in which we must realize them.

This leads to a shift in the status of these rational ideas. For theoretical reason, concepts like God and the soul were "transcendent and merely regulative principles," empty of cognitive content (5: 135). However, through their necessary role in the moral project, "they become immanent and constitutive inasmuch as they are grounds of the possibility of making real the necessary object of pure practical reason (the highest good)" (5: 135). This is a practical, not a theoretical, extension of reason.

What becomes constituted is not the noumenal object (which remains a purely negative limit concept), nor the noumenal subject (the indeterminate capacity for freedom). Instead, what is constituted are the horizonal conditions of moral agency itself: the normative identity that demands the phenomenal self become what it ought to be (an autonomous agent who can reasonably and infinitely strive for moral perfection and the highest good), and the purposive world of nature as a domain where the striving for the highest good. The postulates of God and immortality do not create the noumenal subject or its capacity for freedom. Rather, they secure the possibility of the highest good, thereby giving the moral agent a rational basis for acting as if nature is amenable to moral ends. Kant is clear that "this extension of theoretical reason is no extension of speculation" (5: 134); we are "not able to show how their concept refers to an object, and this is not yet cognition of these objects" (5: 135). We gain no speculative knowledge of God's nature, but practical reason gives these ideas "objective reality" by making them indispensable presuppositions for moral action (5: 135).

Attaining the highest good is what Kant calls the "doctrine of wisdom," a philosophy which "objectively is represented completely only in reason alone, whereas subjectively, for a person, it is only the goal of his unceasing endeavors" (5:108-9). We are not commanded to succeed in creating the highest good, a task that surpasses our power, but to endlessly chase it. This chase is what constitutes the normative identity of the philosopher—the form of agency defined by the end of wisdom, equipped with the capacity for self-correction, and subject to the standard of asymptotic approximation to a fully autonomous being. This chase is what justifies our rational faith and the never-ending bi-constitutive task of becoming virtuous while making the world a virtuous place. The moral law alone remains the sole motivating ground, as “the moral law alone must be viewed as the ground for making the highest good and its realization or promotion the object” (5: 109). Our obligation is thus to the unceasing endeavor itself, a practical commitment that, in turn, requires us to view the world as if it were constituted by a moral author, making our duty a rational and coherent enterprise.

By establishing what theoretical reason cannot, practical reason gives purpose to theoretical reason. It establishes why it is necessary to care about the pursuit of a systematic and complete conception of knowledge at all. Theoretical reason’s autonomous domain is preserved—it alone can tell us how nature works—but its ultimate point comes from outside itself. The world we cognize through the understanding is the same world in which we must act morally, and it is practical reason’s need to act that finally makes theoretical reason’s project meaningful. At the horizonal level, the Idea of autonomy unifies the normative identity of the knower and the moral agent into a single form of agency—the philosopher in the world-concept—and discloses a purposive world in which nature is not a rigid, alien mechanism but a domain amenable to both cognition and moral action. In these ways, practical reason provides what theoretical reason cannot: the *ratio finalis* that gives its inquiries ultimate significance, the *ratio credendi* that warrants belief in its deepest presuppositions, and the *ratio essendi*—the real ground of nature’s purposive unity itself—which we will now examine in more detail.

IV. The self-organization of nature

We have seen that practical reason provides the *ratio essendi* of nature's purposive unity—the real ground that nature is the kind of world where moral ends can be realized. But this requires careful parsing. The purposiveness at issue here is the claim that nature is structured such that the highest good—the harmony of virtue and happiness—can be realized in it. This is not a judgment about nature that theoretical reason could verify. It is a postulate of practical reason, a necessary condition for the moral law to be binding on beings like us who act in the natural world. This postulate carries a demand that nature must be more than a mere mechanism of efficient causes. It must be the sort of whole that can harmonize with the ends of free beings. In short, nature must be *self-organizing*: a purposive, unified system that is not merely receptive to our moral projects but intrinsically oriented toward the realization of ends. Kant argues that this structure is demanded by practical reason, transforming how we *must think* about the natural world in which we act—even if we can never *know* it to be such theoretically.

Kant also identifies a second, distinct mode of accessing the harmony between nature and reason—one that operates not through practical reason's postulation but through the power of judgment. Here, the purposiveness at issue is the way nature appears to our cognitive faculties as if it were designed for our understanding. This is a reflective principle that guides empirical inquiry. Reflective judgment does not transcend sensibility and reason to grasp a supersensible ground. Rather, it reflects on reason's own activity within the bounds of possible experience. When reflective judgment succeeds in discovering a principle that unifies heterogeneous laws of nature, we feel a disinterested pleasure. This pleasure is not the satisfaction of a moral interest but the felt confirmation that nature answers to our cognitive needs. Crucially, this feeling does not prove that nature is purposive. It is a feeling that accompanies successful inquiry, signaling that our reflective presupposition—that nature is purposive for our cognition—has been answered by nature itself. What practical reason postulates as necessary for moral action, the power of judgment feels as confirmed in the success of its inquiries. The two are distinct in kind—one is a postulate of reason, the other is a feeling of reflective judgment—but they converge on the same conclusion: that nature is unified and a place where our rational ends can be fulfilled.

Kant articulates the first of these arguments in a pivotal passage in the first *Critique* that the purposive structure of nature (what he will later call a "self-organizing" structure) is grounded not in theoretical reasoning but in our morality:

> What sort of use can we make of our understanding, even in regard to experience, if we do not set ends before ourselves? The highest ends, however, are those of morality, and only pure reason can grant us cognition of these. But though equipped and guided with these, we still cannot even make any purposive use of our acquaintance with nature for cognition unless nature itself has introduced purposive unity; without this we would not even have any reason, since we would have no school for it and no culture through objects that would offer the material for such concepts. That purposive unity is necessary, however, and grounded in the essence of the faculty of choice itself, and therefore this one, which contains the condition of the application of that unity in concreto, must also be necessary, and thus the transcendental improvement of our rational cognition is not the cause but rather merely the effect of the practical purposiveness which pure reason imposes on us. (A816-7/B845-6)

Here Kant is arguing that the purposive unity of nature is grounded in the faculty of choice—that is, in our capacity to set practical ends (*Willkür*). Kant claims that nature must be unified if moral ends are to be realizable. This is not a claim about how nature appears to our cognitive faculties but a demand on nature's structure that follows from the bindingness of the moral law.

We can break this argument down into four fundamental components: (1) The first two sentences argue that the understanding and our morality are grounded in the structure of our end-setting. We saw that the structure of end-setting is isomorphic in both theoretical and practical reason: both are guided by the Idea (either of the understanding or autonomy, respectively), and both involve normative principles conducive to achieving their desired end. (2) The second sentence establishes the primacy of practical reasoning (see II.C). Our autonomy and its necessary object—the highest good—postulates what theoretical reason must depend upon but cannot establish on its own.

(3) The third sentence asserts that we must postulate a purposive unity in nature, claiming we cannot make any "purposive use of our acquaintance with nature for cognition unless nature itself has introduced purposive unity." In other words, we cannot achieve any ends—theoretical or practical—unless nature is a self-regulating unity amenable to ends. We cannot participate in the Idea—that is, we cannot constitute nature by achieving our moral ends, and thereby cannot constitute ourselves as capable of achieving those ends— unless nature is a self-constituting, self-organizing unity. If we planned to become vegan, help our friends move or speak up against authoritarianism, nature must be stable enough such that these actions could be actualized: "the action must be possible under natural conditions if the ought is directed to it" (A548/B576). Nature must be more than a realm of mere mechanisms, since mechanistic laws can be

infinitely disparate from one another. Nature must unite these mechanistic laws into a unified, coherent whole—a cosmos—and not some disorderly chaos that contains a unique law for every individual piece of matter. If nature was chaotic in this way, we would have no 'school' for our moral reasoning and no 'culture through objects' through which we can learn from the consequences of our actions. Furthermore, the consequences of the same action would be drastically different every time it was performed, making moral reasoning impossible. Historical examples of moral behavior would cease to act as models for us, since such behaviors could not be applicable to current incongruent circumstances.

(4) In the last sentence, Kant argues that since moral reasoning *is* possible, nature must necessarily be a purposive unity. This is what Kant means when he says this unity is "grounded in the faculty of choice." This is not something we can theoretically know, for this would supersede the capabilities of the understanding, which only operates in the context of sensible intuitions. But we can and must posit it as a necessary Idea of reason to act morally. The 'transcendental improvement' that Kant refers to is the organization of our Idea-mediated purposes and aims (our rational cognition) such that we are systematically oriented towards the highest good. We can only actualize the highest good if we bring about a moral world that harmonizes our moral ends with the cosmos. This is grounded in practical purposiveness, since the highest good is the necessary end of the autonomous will.

The ultimate justification for Kant's argument is the "fact of reason," which guarantees the possibility of moral reasoning. We saw earlier that this fact of reason is an acknowledgment of the fact that we are conscious of the normative pull of an "ought" to do the right thing. Yet this "ought" would be empty if it could not lead to a possible action. The fact of reason, hence, is the fact that we *can* use reason to successfully achieve our moral ends. The implication of this is that although we can never *know* that nature is a purposive unity, practical reason compels us to *postulate* this unity as the necessary condition for the realizability of its ends, including its ultimate end, the highest good. This is the argument Kant makes in the first *Critique*: the purposive unity of nature is "grounded in the essence of the faculty of choice itself" (A817/B845), a necessity demanded not by theoretical cognition but by the structure of moral agency trying to actualize itself in the phenomenal world.

We must ask Kant: do not all of our ends assume that nature is a regulated unity? What is so special about our moral ends in particular? I could not achieve my everyday ends of walking around the block and making breakfast if nature had not introduced a unified, regulated structure. The difference between these everyday ends and moral ends lies not in the need for regularity, but in the scope of the presupposition. Everyday ends—drinking coffee, building a bridge—presuppose only local, specific regularities (the laws of physics, chemistry). I can plant a tree without believing the entire universe is purposively aligned with my horticultural ends. Moral action, in contrast, demands a belief about the *total* system of nature. This is because it's necessary object—the highest good—is not one end among others within the world, but the final end of reason as a whole: a state where the entire cosmos is so constituted that virtue and happiness are in harmony. If nature were ultimately a chaotic and indifferent machine that systematically crushed justice and rewarded vice, the moral project itself would be futile. The "fact of reason"—our experience of moral obligation—proves that the moral project is coherent, and that nature can in fact become susceptible to our moral ends. Therefore, we are compelled to conclude that the world possesses the purposive unity necessary for the pursuit of morality.

What, then, is the status of scientific practice? Does it not *also* presuppose that nature is a systematic unity to perform its activities? Theoretical reason does indeed presuppose nature's systematic unity, but its commitment is merely regulative. Science must operate as if nature is a unified, intelligible whole to conduct its inquiries, but it can never cross the gap to establishing this constitutively. For theoretical reason, the failure of this presupposition would mean the collapse of a unified scientific endeavor. For practical reason, however, the failure of nature's unity would be a catastrophic failure of human rationality as such, making the foundational moral law (and the theoretical reason it grounds—see II.C above) a command to achieve an impossibility. This disparity in stakes justifies the primacy of the practical. Theoretical reason can only hypothesize unity; practical reason, grounded in the fact of reason, postulates this unity as a necessary, constitutive reality (*ratio essendi*) of the cosmos.

In the third *Critique*, Kant adds a second, distinct mode of accessing the harmony between nature and reason. This mode operates not through practical reason but through the power of judgment. Kant writes that "the attainment of every aim is combined with the feeling of pleasure" (5:187). For certain aims—specifically, the aim of reflective judgment to find unity in nature's empirical laws—we can experience a pleasure that is not personal but is determined by an "a priori ground valid for everyone" (5:187). This feeling occurs "merely through the relation of the object to the faculty of cognition, without the concept of purposiveness... having the least regard to the faculty of desire" (5:187). It is thus entirely distinct from the pleasures of the agreeable (like eating ice cream) or the good (like moral satisfaction). Grounded in the universal structure of cognition rather than in personal preference, this feeling indicates a congruity between nature and our cognitive faculties.

This kind of pleasure arises whenever reflective judgment succeeds in discovering a principle that unifies heterogeneous laws of nature, or more generally when reason's regulative Ideas successfully guide empirical inquiry (5:187–188). The harmony at work here is not the harmony *within* the subject—the free play of imagination and understanding that characterizes aesthetic judgment—but the harmony *between* reason and nature: the felt accord between our cognitive demand for systematic unity and nature's amenability to that demand. In such moments, we do not merely postulate that nature is purposive; we *feel* it. The success of our theoretical endeavors provides a felt confirmation that the presupposition guiding our inquiry—that nature is systematically unified—has been answered by nature itself. This confirmation is not reducible to wish-fulfillment or instrumental convenience. When I desire chocolate and eat it, my pleasure tells me only that my desire was satisfied—nothing about the objective quality of the chocolate or the reasonableness of my desire. The disinterested pleasure Kant describes is fundamentally different. It does not arise from getting what we wanted, but from the *accord* between nature and our a priori cognitive faculties that successful discovery reveals.

Since it is grounded in the a priori structure of cognition itself, not in the contingent interests of the investigator, this pleasure can claim universal validity.

The a priori commands of reason (as disclosed in the Idea) and the regulative principles of judgment are thus disinterested in that we must employ them regardless of whether we desire to, for they are necessary conditions of empirical inquiry itself (5:188; A647/B675). When we presuppose such a framework, the pleasure we feel upon success is not the satisfaction of a prior desire but the felt confirmation that this necessary presupposition has been answered by nature. This lends the pleasure epistemic weight, signaling that the presupposition was appropriate to the way nature is structured.

Within this framework, the spectacular, unifying success of science—exemplified by fields like complex systems theory—assumes its proper role. It does not *prove* the practical postulate, but it provides a powerful *corroboration* from its success to both model various complex entities and to unite various strands of science into one comprehensive framework. As Kant himself suggests, the very possibility of empirical investigation rests on presupposing that nature is purposive for our cognition (5:186–187), and the pleasure we feel in discovery—the admiration that accompanies the unification of heterogeneous laws (5:187–188)—is the *felt confirmation* that this presupposition is not merely subjective but is instead a condition of cognition that we experience as an objective truth. It demonstrates that the purposive, self-organizing world we are morally compelled to believe in is the same world that proves itself to be felt through our successful scientific investigations. Successful instances where science unifies many disparate laws into overarching theories is thus an *echo* of the unity that practical reason is compelled to postulate.

What emerges, then, is a dual account of how the harmony between nature and reason is accessed. Practical reason, operating through the faculty of desire, postulates nature's purposiveness as the necessary condition for the realizability of the highest good. This is a demand on nature's being, grounded in the bindingness of the moral law. Reflective judgment, operating through a universally accessible feeling, experiences nature's purposiveness as a disinterested pleasure that accompanies successful empirical inquiry. This is not a judgment about nature but a felt confirmation that our regulative presuppositions have been answered by nature itself. The two converge on the same conclusion—that nature is not alien to our rational ends and has, by virtue of its self-organization, the capacity to achieve the highest good—but they do so through different modes: one through postulation (practical reason), the other through feeling (reflective judgment).

V. Conclusion: Neither Complexity Realism nor Complexity Idealism

In our previous investigation (chapter 1), we saw how complexity realism—the idea that the cosmos really *is* self-organizing, independently of our cognition—and complexity idealism—that our consciousness projects its own complexity out onto an independent nature—are both problematic. The sciences of complex systems, however much they claim to reveal 'reality' apart from us, is not sufficient to make sweeping claims about the cosmos as a whole. Neither does reflection about our own complex utilization of idealizations, simplifications, and generalizations in our ever-more complex sciences lead us to believe that our own self-organization imposes its structure onto a passive cosmos. We saw that both positions lead us to forego the idea that either we or nature are self-organizing. Either we 'copy' the self-organization of nature into our minds, making us no longer self-organizing (and no longer 'natural', if what is natural is to self-organize), or we impose our own self-organization on to nature, making nature no longer self-organizing (and hence organized by us). We were required, then, to investigate the nature of how we, as autonomous agents, self-organize in relation to a self-organizing nature.

By studying how Kant navigates this problem, we have discovered a compelling alternative to both positions. He disagrees with complexity realism by claiming that we can have no metaphysical knowledge that nature is self-organizing. The certainty we have that nature is self-organizing is grounded not in theory, but in our own aspirations to be autonomous. Since to be autonomous is to follow the moral law, the self-organizing, purposive nature is a condition that we must postulate if achieving our moral ends is to be possible at all. Since it is a fact that we *can* pursue these ends, it is necessary that nature be self-organizing.

In this way, Kant skirts all the possible objections to complexity realism. However, is that because he is a complexity idealist, where the complexities of the mind are imposed upon nature? As I have argued, Kant does not think that the mind imposes order onto nature, as the standard reading of him suggests. It is only through the Ideas that our minds and the world simultaneously gain structure. Ideas originate in neither the mind nor the world, but constitute them. The origin of the Ideas lies within our practices, which constitute both what it means to be a practitioner and what counts as the object of that practice's ends. Neither does reason 'impose' its structure onto both the self and nature. Rather, reason demands something of both—without being able to control the response. The Idea demands that the indeterminate subject become a determinate agent of a particular kind, and that nature show up as a domain amenable to that agent's ends. But the demand is not a guarantee. The subject may falter; inclinations may pull elsewhere; the world may resist; nature may refuse to yield its laws, the moral world may remain unrealized. The Idea calls forth a demand, but the response is answered, or not, in the unpredictable course of practice. This fallibility—the ever-present possibility that the demand will go unanswered—is not a failure of Kant's system but a feature of it. It is what generates the endless process of self-correction that drives both our cognitive and moral striving.

It is purely within the purview of our own practices of autonomous action that we postulate that nature is self-organizing. Yet although Kant's insights seem to not depend on contemporary science's findings, these sciences nevertheless seem to confirm that nature is self-organizing at every turn: the pervasiveness of self-organization in the cosmos at both the smallest and largest scales is felt as it is confirmed again and again in scientific practice, leading to our disinterested pleasure that nature conforms to our cognitive structures. Theoretical reason fleshes out the purposive unity that practical reason demands and in doing so, it can actively inform and refine our conception of it. It is because of this that scientific advances can alter our articulation of the highest good.

Since complex systems science and environmental science have shown the fragility of the self-organizing ecosystem we live in, this suggests that a strong environmental ethics is required to achieve the highest good. We will consider a Kantian environmental ethics in this manner in chapter 6. In the next chapter, we will see how Kant understands non-human systems to be self-organizing, spelling out in further detail his general theory of (human and non-human) agency.

Bibliography

Allais, L. (2015). *Manifest Reality: Kant's Idealism and His Realism*. Oxford University Press.

Allison, H. (1990). *Kant's Theory of Freedom*. Cambridge University Press.

Allison, H. (1996). *Idealism and Freedom: Essays on Kant's Theoretical and Practical Philosophy*. Cambridge University Press.

Allison, H. (2004). *Kant's Transcendental Idealism: An Interpretation and Defense* (Revised Edition). Yale University Press.

Brandom, R. (2000). *Articulating Reasons: An Introduction to Inferentialism*. Harvard University Press.

Buchdahl, G. (1992). *Kant and the Dynamics of Reason: Essays on the Structure of Kant's Philosophy*. Blackwell.

Cassirer, E. (1981). *Kant's Life and Thought*. Yale University Press.

Chulanon, P. (2025). Kant on the Theoretical Use of the Ideas of Reason: A Transcendental Interpretation. *Canadian Journal of Philosophy,* 54 (3), 230-250.

Engstrom, S. (2013). Unity of Apperception. *Studi Kantiani*, 26, 37-54.

Fisher, N. (2023). Life, Lawfulness, and Contingency: Kant and Schelling on Organic Nature. *Archiv für Geschichte der Philosophie*, 105 (1), 163-188.

Förster, E. (2000). *Kant's Final Synthesis: An Essay on the Opus Postumum*. Cambridge University Press.

Ginsborg, H. (2001). Kant on Understanding Organisms as Natural Purposes. In E. Watkins (Ed.), *Kant and the Sciences* (231-254). Oxford University Press.

Guyer, P. (1987). *Kant and the Claims of Knowledge.* Cambridge University Press.

Grier, M. (2001). *Kant's Doctrine of Transcendental Illusion*. Cambridge University Press.

Hughes, F. (2017). *Kant's Aesthetic Epistemology*. Edinburgh University Press.

Jaspers, K. (1962). *Kant: From* The Great Philosophers*, Volume 1*. (H. Arendt, Ed.). Harcourt, Brace & World.

Kant, I. (1993). *Opus Postumum*. (E. Förster, Ed.). Cambridge University Press.

Kant, I. (1998). *Critique of Pure Reason* (P. Guyer & A. Wood, Eds.) Cambridge University Press.

Kant, I. (2000). *Critique of the Power of Judgment* (P. Guyer, Ed.). Cambridge University Press.

Kant, I. (2015). *Critique of Practical Reason.* Cambridge University Press.

Kant, I. (2019). *Groundwork of the Metaphysics of Morals* (R. Stern, Ed.) Oxford University Press.

Keller, P. (2013). Ideas, Freedom, and the Ends of Architectonic. In J. Stolzenberg & F. Rush (Eds.), *Freiheit/Freedom* (51-78). De Gruyter.

Keller, P. (2015). Cassirer's Retrieval of Kant's Copernican Revolution in Semiotics. In J. Friedman & S. Luft (Eds.), *The Philosophy of Ernst Cassirer* (259-288). De Gruyter.

Kemp Smith, N. (1992). *Commentary to Kant's Critique of Pure Reason*. Humanities Press.

Kitcher, P. (2011). *Kant's Thinker*. Oxford University Press.

Kleingeld, P. (2010). Moral consciousness and the 'fact of reason' In J. Timmerman & A. Reath (Eds.), *A Critical Guide to Kant's Critique of Practical Reason* (55-72). Cambridge University Press.

König, R. (2022). Towards a Unity of Theoretical and Practical Reason: On the Constitutive Significance of the Transcendental Dialectic. *Open Philosophy,* 5 (1), 622-635.

McLaughlin, P. (1990). *Kant's Critique of Teleology in Biological Explanation*. Edwin Mellen Press.

Neiman, S. (1994). *The Unity of Reason: Rereading Kant*. Oxford University Press.

O'Neill, O. (1990). *Constructions of Reason: Explorations of Kant's Practical Philosophy.* Cambridge University Press.

Schafer, K. (2018). Practical Cognition and Knowledge of Things-in-Themselves. In D. Heide & E. Tiffany (Eds.), *Kant's Transcendental Deduction and the Theory of Apperception* (83-109). Oxford University Press.

Schafer, K. (2023). *Kant's Reason: The Unity of Reason and the Limits of Cognition*. Oxford University Press.

Strawson, P. (1966). *The Bounds of Sense*. Routledge.

Vanzo, A. (2018). Leibniz on innate ideas and Kant on the origin of the categories. *Archiv für Geschichte der Philosophie*, 100 (1), 19-45.

Watkins, E. (2010). The antinomy of practical reason: reason, the unconditioned and the highest good. In J. Timmerman & A. Reath (Eds.), *A Critical Guide to Kant's Critique of Practical Reason* (55-72). Cambridge University Press.

Waxman, W. (2019) *A Guide to Kant's Psychologism*. Routledge.

Willaschek, M. (2010). The primacy of practical reason and the idea of a practical postulate. In J. Timmerman & A. Reath (Eds.), *A Critical Guide to Kant's Critique of Practical Reason* (55-72). Cambridge University Press.

Wuerth, J. (2014). *Kant on Mind, Action, and Ethics*. Oxford University Press.

Ypi, L. (2021). *The Architectonic of Reason*. Oxford University Press.

Zammito, J. H. (1992) *The Genesis of Kant's Critique of Judgment*. University of Chicago Press.

Chapter 3: Bi-constituting Agents in the Natural World

I. Introduction

In chapter 1 I argued that due to a conflux of historical circumstances—contemporary complex systems science's understandings of self-organization, the rising philosophical interest in scientific practice and the return of the whole organism in biology—it was imperative to return to Kant to understand these interlinking phenomena. Only Kant can harmonize the self-organization of our own agency with the self-organization in the cosmos under reason or the Idea. I argued that contemporary understandings of self-organization are either complexity 'realist', prioritizing the self-organization latent in nature to explain the self-organization of our human practices, or complexity 'idealists' prioritizing instead the self-organization of our minds to explain the apparent self-organization in nature. These two accounts fall short, failing to properly explain both our own apparent self-organization as well as the self-organization we see in nature. I suggested that Kant's account offers a way out of this dilemma by rejecting the inherent dualism in either of these accounts.

In chapter 2, I investigated how Kant thinks our own organization as rational agents relates to the self-organization of the cosmos. For Kant, we are self-organizing beings because our practical (moral) agency is guided by comprehensive normative structures called Ideas— such as the Idea of autonomy—that constitute and guide our practices. These practices operate through a dynamic of bi-constitution: in pursuing an

Idea, we simultaneously constitute ourselves as specific kinds of rational agents and constitute the world as a meaningful domain for achieving our ends. Crucially, the moral project commanded by reason—the realization of the highest good—would be incoherent if the cosmos were a blind, mechanical chaos. Therefore—since morality is a 'fact of reason'— practical reason compels us to postulate that nature itself is a purposive, self-organizing unity receptive to our moral aims. While theoretical science can only hypothesize such unity, moral action proves it as a necessary condition of its own possibility. Furthermore, Kant argues that we 'feel' the harmony between our own cognitive faculties and the self-organization in the cosmos when our Idea-mediated practices are successful. Both practical reason and this universal feeling acquired from the power of judgment ground the concept of a self-organizing cosmos not in metaphysical speculation but in our own Idea-guided practices, thereby avoiding the fallacies of complexity realism and complexity idealism.

Although the preceding inquiry characterizes nature as a self-organizing unity, how are we to think of entities within nature to be self-organizing? What is Kant's theory of these non-human agents? For Kant, non-human, natural agents are to be understood 'as if' they were ends. That is, we must treat them analogously to works of art or intentional productions. As such, they are to be understood as if their activities, too, are mediated by an Idea. Although the Kant literature amply acknowledges that Kant views organic agency as analogous to human agency, it does not yet account for my novel interpretation of Kant's bi-constitutive, Idea-mediated account of agency, which I develop in chapter 2. As we have seen for Kant, Ideas orient practices which constitute an internal constitution of the subject paired with an external constitution of the environment. Can natural agents be understood in this way? Are they 'ecological' agents, that dynamically constitute themselves and their environments?

This chapter will unfold in three parts. First, I will unpack how Kant's core teleological understanding of natural agents—what Kant calls "natural ends"—relies on an analogy with our own Idea-mediated agency. Second, I will propose that Kant intended this kind of agency to be understood on the model of bi-constitution established in chapter 2. I will first outline a formal, four-part schema of self-organizing agency, where an immanent "Idea" orients a dynamic process that constitutes internal parts and an external environment in a ceaseless striving to maintain its own structure. I will then textually substantiate this claim as Kant's intended understanding of natural agents. Finally, I will contrast this Kantian model of agency with contemporary theories of "ecological agency." I contend that while such theories aim to explain agency through dynamic matches between given capacities and environments, they either fail to capture the essential, self-organizing character of agency or must tacitly presuppose the very bi-constitutive logic Kant articulates. The chapter thus shows how Kant populates the self-organizing world that morality necessitates with pockets of self-organizing, non-human agents by conceiving them—in analogy to us—as responsive to the Idea.

II. Kant and Teleology

According to Kant, a natural end—such as an organism—is defined as "cause and effect of itself" (5: 371). This definition posits a teleological structure of circular causality, where the whole (e.g., the organism) organizes its parts (e.g., the organism's organs), and the parts, in turn, reciprocally sustain and produce the whole. Kant explains that such self-organizing wholes must satisfy two criteria. The first is that its parts "are possible only through their relation to the whole" (5: 373). To be an end, an organism, for example, must fall under an Idea "that must determine a priori everything that is to be contained in it" (ibid). For example, a tree's roots, leaves, stems, and seeds are all determined through its "idea of a whole" (i.e., the tree itself). The second criterion is "that its parts be combined into a whole by being reciprocally the cause and effect of their form" (5: 373). Each organ, for instance, must produce the other parts (the other organs) while also itself being a product of these other organs. In this way, Kant calls such "organized and self-organizing being" a "natural end" (5: 374).

Kant claims that such self-organizing beings cannot be understood according to mechanistic causality. While the "idea has to ground the possibility of the product of nature" within an "absolute unity," matter is a "multitude of things, which by itself can provide no determinate unity of composition" (5: 377). Physical-mechanical necessity, according to Kant, is where a "thing is possible in accordance with mere laws of efficient causes" (20: 241). Mechanistic causality refers to the 'linear' laws of matter, whereby a cause produces a determinate effect in a one-directional sequence. Kant calls the current state of such a system a "causal nexus" (i.e., of efficient causes), that is, a "connection that constitutes a series (of causes and effects)" where the "things themselves, which as effects presuppose others as their causes, cannot be the causes of these at the same time" (5: 372).

Whereas mechanical systems exhibit necessary and universal laws of linear processes, self-organizing systems exhibit non-linear, circular causality where a part of the system (what we might call a subsystem or an organ) can be both cause and effect at the same time. The causality of a self-organizing being obeys mechanistic causality, but the mechanisms involved in the maintenance of such a system would be entirely contingent if taken in abstraction from the overarching unity through which they occur: "nature, considered as a mere mechanism, could have formed itself in a thousand different ways without hitting precisely upon the unity in accordance with such a rule" (5: 360). What is it about nature that seems to constitute these extraordinarily unlikely pockets of order? How can we understand such order?

Our cognitive faculties have no constitutive concept from our understanding that can capture an entity that is both cause and effect of itself. Our cognition is prepared to establish universal laws for causes and effects that happen in a linear series, not for causal processes that are non-linear. Non-linear causal processes, as we have seen, are extremely contingent. No universal law can be prescribed to them. To resolve this, Kant introduces a specific faculty he calls reflective judgment. Unlike determinative judgment, which applies a given universal concept (like "cause") to a 'particular,' reflective judgment searches for a universal concept for particulars for which we have no ready-made concept (5: 179). As the name suggests, reflective judgment's task is to 'reflect' upon that particular to find a concept for it. A natural end (such as an organism) is such a particular in need of a concept. To guide this search, reflective judgment needs a principle. It cannot get one from the understanding, whose categories can only establish linear cause/effect relations, so it borrows an "Idea" from reason—the Idea of a purpose or end (20: 235; 5: 361; 5:372).[41]

By 'borrowing' the Idea of an end from reason, Kant means to do so using an *analogy*. By using an analogy with our own capacity for purposive action, we acquire the necessary model for investigating self-organizing beings. According to Kant in *The Critique of the Power of Judgment*, the agency we ascribe to organic entities is in "analogy" with "the kind we encounter with ourselves" (5: 360). That is, for Kant, we look at organisms and other self-organizing phenomena 'as-if' they were organized according to an Idea. The analogy is not to our practical, moral reasoning of which we saw is mediated by Ideas. Rather it is to our "technical" reasoning: "the art of bringing about that which one wishes should exist" (21: 200). The analogy is closer to that of a craftsman, whose Idea of a pot, for instance, guides their construction of a pot. Similarly, the activities of, for instance, an organism is guided as if by some coordinated plan—an 'Idea'—that constitutes the organization of its parts and how it interacts with its external environment.[42] This Idea

[41] In these passages, Kant refers to the concept of an end to be a 'concept of reason,' which Kant in the first Critique identifies with an Idea: "a concept…which goes beyond the possibility of experience, is an idea or a concept of reason" (A320/B377).

[42] A popular view in the literature is to claim that Kant is a 'naturalist' by claiming that reason seems to be intelligible in light of the teleology of the organism (Ypi 2001, ch.3) or that reason is a form of life (McDowell, 1994). This is to put the cart before the horse for Kant. Rather, it is only on the basis on an analogy with reason (particularly, technical reason) that we attribute to life its self-organizing, as-if-Idea-mediated character. This does not preclude, however, that we can give a naturalist account of the origins of reason in nature, similar to the "daring adventure of reason" (5: 419) that Kant describes as the process where we trace the origins of life from inorganic matter. This natural history of reason, however, must always happen *within* reason itself, meaning reason is the inescapable ground for any such naturalist

establishes, as it were, a normative identity for the agent, which constitutes the basis through which we can say things are good or bad for such an agent. That is, a woodpecker that is bad at pecking is not good at being a woodpecker, because it fails to enact the capacities established by the 'Idea' of the woodpecker.

For Kant, an analogy is not a direct resemblance, but a shared "rule of reflection" between a sensible intuition (a symbol) and an Idea (what is symbolized). He claims,

> the power of judgment performs a double task, first applying the concept to the object of a sensible intuition, and then, second, applying the mere rule of reflection on that intuition to an entirely different object, of which the first is only the symbol. (5: 352).

This double task of judgment can be explained as follows. For Kant, the first 'task' of judgment is to take a normal sensible thing and apply a concept to it. For example, take the sun as an object that we apply an empirical concept to: "sun." We understand what the concept of "sun" entails: a supreme, unique entity in the solar system providing the necessary conditions for both the existence of life and the visibility of everything on earth. In this context, the 'rule of reflection' dictates certain patterns that organize our thoughts and feelings around the concept "sun." This pattern might revolve around, for instance, the fact that the sun provides the ultimate conditions that make possible all life and knowledge. We can then apply this rule of reflection to a completely different object, something that is not sensible, like an abstract Idea of reason. For example, Plato often talks about the Idea of the Good in analogy with the sun. Here, the sun (the sensible object in task 1) now serves as a symbol for the Good (the abstract Idea from task 2). The

account. We cannot escape reason to find something more fundamental than reason; however, we can use reason to investigate its natural origins.

sun is not the Good, but it is a stand-in that gives us an analogical model to think about it. Just like the sun produces all life and knowledge, we can then think of the Good as the source of the moral law that nourishes and sustains our rational nature.

The 'double task' of the analogy between technical reason and organized beings goes as follows. The first task establishes the symbolic basis by reflecting on our own capacity for technical reason. Consider the craftsman guided by the Idea of a pot. The 'rule of reflection' we derive from this case is not the craftsman's subjective intention, but the objective, bi-constitutive logic of the activity itself. The Idea of the pot functions as a constitutive norm that simultaneously determines the practice (pot-making), defines the practitioner (the craftsman qua potter), and governs the product (the pot as a functional unity). The craftsman and the pot are thus distinct yet mutually constituted termini of a single, norm-guided activity. This structure—where a technical practice oriented by an Idea bi-constitutes an agent and an external artifact—provides the formal rule we will apply to the natural (non-human) agent.

The second task of judgment applies this rule to the natural agent. The comparison between the natural agent and a craftsman reveals a unique application of the analogy. For the craftsman, the practitioner (potter) and product (pot) are separate entities. In the natural agent, however, the primary product of its constitutive practice is the practitioner itself. At the actual level of bi-constitution, practitioner and product coincide: the organized being produces itself in a way that organizes all its internal parts for the end of maintaining its own structure. Additionally, at what I called, in chapter 2, the 'horizonal' level (the level of *potential* action), the Idea of the organized being also seems to constitute a series of capacities for the agent to act, as well as a world that is purposive to its own actions. This is also in direct analogy to the craftsman, who—when guided by the Idea of the pot—orients themselves into a series of capacities capable of making pots, while constituting the world in terms of materials and tools conducive to the task of pot-making. There is, however, one crucial difference: the organism is its *own* artifact; it is "cause and effect of itself." But this is not merely the same relation holding accidentally, as if the organism just happens to produce itself. Rather, the organism moves itself qua itself — the sameness of relation between producer and product holds non-accidentally, essentially, and of necessity, because the organism's identity just is this self-organizing activity. This self-referential closure is what sets natural ends apart from artworks and artifacts—the craftsman analogy helps to illuminate organic self-organization precisely by extending the analogy to a case where the end and the agent coincide.

What status could this analogy have for our knowledge? Do organized beings really exist? Or are they a mere heuristic? One might think that such a teleological principle is at odds with the universal dominion of mechanistic causality. However, Kant recognizes that judging nature teleologically is a regulative principle of our judgment. But this does not mean the analogy is merely subjective or that the rule of reflection applies only to the judger. On the contrary, for Kant, the rule of reflection is applied to the object itself: we reflect on the object under the guidance of the Idea, and the object (the organized being as it appears) objectively exhibits the structure of self-organization that the analogy articulates. The regulative status means only that we cannot have constitutive theoretical knowledge of this structure as a thing-in-itself; it does not mean the structure is a mere projection or heuristic without objective purchase on the phenomenon. As we saw in chapter 2, a regulative theoretical principle of reason (e.g., generated by the Idea of the Soul) guides inquiry (e.g., psychology) by orienting it to abide by certain norms. These principles can guide inquiry without actually positing that the object of the Idea (e.g., the Soul) actually exists or can be known as a noumenon. They exist to guide inquiry, not to assert metaphysical existence. A constitutive theoretical principle, on the other hand, such as the principle of causality, directly constitutes our phenomenal experience and can be known. By claiming that teleological judgment is regulative, Kant similarly intends that we cannot know that organized beings are really mediated by Ideas as things in themselves, yet they nevertheless guide inquiry into nature —and they do so because the objects themselves, as they appear, demand this rule of reflection for their intelligibility.

Kant's regulative principles are more than arbitrary heuristics. The regulative principle that we should investigate certain beings *as if* they were self-organizing and acting according to ends is certainly not arbitrary; it is a necessary guide for our inquiry into nature (5: 404). Without it, we would not look for unifying, systematic relationships between phenomena. Science would hence be impossible.

Furthermore, it is not as if mechanistic judgment holds a more privileged, constitutive status for Kant. To be sure, Kant does think we can have constitutive judgments of nature. These are determinative judgments, where the pure concepts of the understanding (e.g., causality) are directly constitutive of experience, guaranteeing that our experience is "lawlike." [43] Yet the specific content of these laws— "the empirical laws of motion" (20: 237)—remains contingent. To discover and systematize these contingent laws, we need regulative principles. Even in the first *Critique*, Kant describes the "idea of a mechanism" as an Idea that unites phenomena in "the chemical effects of materials" (A645-6/B673-4). He states that such an Idea serves for the "thoroughgoing unity of these concepts, insofar as the idea serves the understanding as a rule" (A645/B673). Thus, mechanism does not constitutively determine nature's material laws; it regulatively guides our search for linear causal explanations, just as teleology guides our inquiry into self-organizing beings.[44] They are not conflicting truths about nature but complementary rules for our judgment. [45]

I mentioned that for Kant, teleological judgment is necessary insofar as it makes empirical cognition possible. We cannot even judge things to exist in mechanistic relationships if we have not identified holistic contexts within which these mechanistic relationships are intelligible. However, Kant seems to imply a further justification for the regulative principle of teleological judgment. In the Critique of Pure reason, Kant mentions that Ideas in the theoretical sphere are like hypotheses we use to question

---

[43] We must keep in mind—as we saw in Chapter 2— that even the constitutive nature of the understanding is justified through recourse to the "Idea of the understanding" which as an Idea, regulates our investigations in how we critique our own cognitive faculties.

[44] This is a controversial claim that is debated extensively in the literature, as the *Critique of Pure Reason* and the subsequent *Metaphysical Foundations of Natural Science* seem to imply that Kant intended his categories to be constitutive of our experience of mechanical causality. See Breitenbach (2009), Ginsborg (2004), and McLaughlin (1990) for varying views.

[45] When an agent dies or begins to dissolve, one might assume that it returns to mere mechanism. This might suggest that mechanism is more fundamental for Kant. While this might be true from the perspective of theoretical reason, recall that for Kant that nature is inherently teleological, as derived from the necessary postulates of practical reason (see chapter 2). Hence, even if agents die here and there, it does not disprove the holistically self-organizing description of nature.

nature, subjecting nature to inquiry as if we were an "appointed judge." Theoretical Ideas also seem to be justified by their ability to guide inquiry. Similarly, we use teleological judgment as a regulative principle for the end of cognition in general, "guiding research into objects" (5: 375). This guideline serves to "supplement the inadequacy" of "research in accordance with mechanical laws" for the purpose of "the empirical search for all the particular laws of nature" (5: 383). Judging nature in terms of purposes therefore seems to be justified by its ability not to *prove* that nature is intrinsically purposive, but by its *fruitfulness* in explanation. When the guiding analogy of purposiveness productively intersects with the regulative Idea of mechanism, scientists discover lawful relations that would otherwise remain obscure. The "end" of this purposive framework is thus the successful completion of theoretical reason's task: the systematic unification of empirical knowledge.

In the last chapter, we saw how the successful instantiation of theoretical Ideas leads to a certain kind of universal pleasure. Kant intends his regulative principle of teleological judgment to have a similar sense of validation. Here, the felt confirmation of purposiveness concerns not, e.g., the systematic unity of our scientific laws in general, but the specific judgment that certain natural products must be regarded as if they were self-organizing ends.

Kant claims that "the attainment of every aim is combined with the feeling of pleasure" (5:187). In the context of teleological judgment, this pleasure arises when our investigation into organized beings brings about successful articulation according to mechanistic laws. When we approach an organism *as if* it were a self-organizing end—

asking not merely how its parts interact but what they *do for* the whole—and when this approach yields genuine insight, we feel a distinctive pleasure. For Kant, this pleasure is determined by an "a priori ground valid for everyone" and arises "merely through the relation of the object to the faculty of cognition, without the concept of purposiveness... having the least regard to the faculty of desire" (5:187).

Because this pleasure is detached from the faculty of desire, it is not reducible to wish-fulfillment or pragmatic convenience. When a medical researcher hopes her drug trial succeeds and it does, her satisfaction tells her only that her desire was fulfilled. But when a biologist discovers that a previously mysterious organelle serves a specific function within the cell, the pleasure she feels is different. It does not merely confirm that she got what she wanted; it confirms that her way of approaching nature—*as if* it were organized according to ends—was appropriate to the object itself. This is why the pleasure can claim universal validity: it is grounded in the universal structure of cognition itself, not in the contingent interests of the investigator. [46]

---

[46] Here I agree largely with Karl Jaspers's reading of Kant, where he claims: "In this existence of knowledge, which in itself cannot be concluded but remains open, the supersensible makes itself felt by acts, not through any known content. In the phenomenality of existence, in this area of universal objectivity produced by the understanding, the Ideas are the breach through which the supersensible enters into knowledge" (1962, 57). For Kant, the supersensible (das Übersinnliche) refers to the intelligible ground that underlies both nature (as appearance) and freedom (as the moral law). It is the "supersensible substratum" of nature (5:196) and the "supersensible in the subject" (5:195)—a single, unified ground that we can never know through concepts but can think as the hidden source of both domains. The Ideas of reason—the soul, the world, and God—are one "breach" through which this supersensible ground becomes felt rather than known. When a regulative Idea successfully guides empirical inquiry, the pleasure we feel is not merely cognitive satisfaction; it is the felt recognition that the supersensible ground is one, that nature and freedom (i.e., the freedom to generate epistemic practices from the Idea) are not ultimately opposed but are unified in a way that our discursive intellect can never fully grasp.

The regulative principle of teleological judgment provides another such breach. When we judge organized beings as if they were products of natural ends, we are not asserting the objective reality of such ends but are opening ourselves to a mode of judging that reveals the supersensible substratum in a different way. The success of teleological judgment—our ability to discover and unify biological phenomena under

As I argued in the last chapter, this pleasure lends the principle epistemic weight: it signals that the presupposition was appropriate to the way nature is structured. As we have furthermore seen in chapter 1 and will discuss in more detail in chapter 5, the regulative principle of teleological judgment has been wildly successful in contemporary science, culminating in fields like complexity sciences, systems biology, and artificial life. These fields presuppose that nature contains systems that produce, maintain, and adapt themselves—from ant colonies whose individual members generate collective intelligence, to gene regulatory networks that sustain cellular stability. The success of this approach, the discoveries it yields, produces the same disinterested pleasure Kant describes. This success does not prove that we can *know* nature be comprised of teleological causes: instead, it validates the framework we use to see nature in this way. Our judgment reveals a deep harmony between our cognitive faculties and nature, giving objective weight to the existence of natural ends, without knowing this to be the case.

To summarize so far, Kant's analysis of the non-human, natural agent as a natural end culminates in a sophisticated analogical framework. It establishes that we can only grasp self-organizing beings by viewing them *as if* they were guided by an Idea, analogous to the way a craftsman is guided by the Idea of the artifact they wish to create. This framework resolves the apparent conflict with mechanism by showing both principles to be regulative, complementary guides for our reason. However, while we have established *that* such an analogy is necessary and identified its general form—a "rule

---

this regulative principle—produces the same disinterested pleasure as the success of reason's Ideas, and for the same reason: it enacts a harmony between nature and freedom.

of reflection" concerning bi-constitution—we have not yet fully unpacked its precise logical structure. The critical question of *how exactly* the bi-constitutive logic of the craftsman's Idea maps onto the internal and external relations of an organized being remains to be systematically detailed. It is to this task—the derivation of a formal schema of self-organizing agency from the analogical rule—that we now turn.[47]

[47] Angela Breitenbach agrees with my view, so far, with Kant's view that we must draw an analogy between the organism and our end-setting reason. She concludes that "organized beings" must "be judged in analogy with the systematic unity and purposeful activity of our own reason." (Breitenbach 2009, 109—my translation). Much as we saw how reason is unified in a dynamic whole in Chapter 2, so are natural purposes. She finds the 'rule of reflection' to be that which supports a total system (ibid, Ch. 4). She gathers support for this in Kant's claim that for organized beings, their "internal possibility presupposes throughout the idea of a whole on which even the constitution and mode of action of the parts depends, which is just how we must represent an organized body" (5:408). Organized beings, like reason, must also be a systematic totality. However, Breitenbach's agent is not the dynamic, pragmatic, self- and world-constituting agent that we have discussed in chapter 2 and will see below. Breitenbach also does not seem to acknowledge the pragmatic success conditions of regulative Ideas, which in my reading give more ontological weight to the regulative Idea of purposiveness given its overarching success (a success we can feel but cannot know). This gives us a drastically different way of understanding the regulative analogy than Breitenbach supposes.

III. Self-organizing Agency

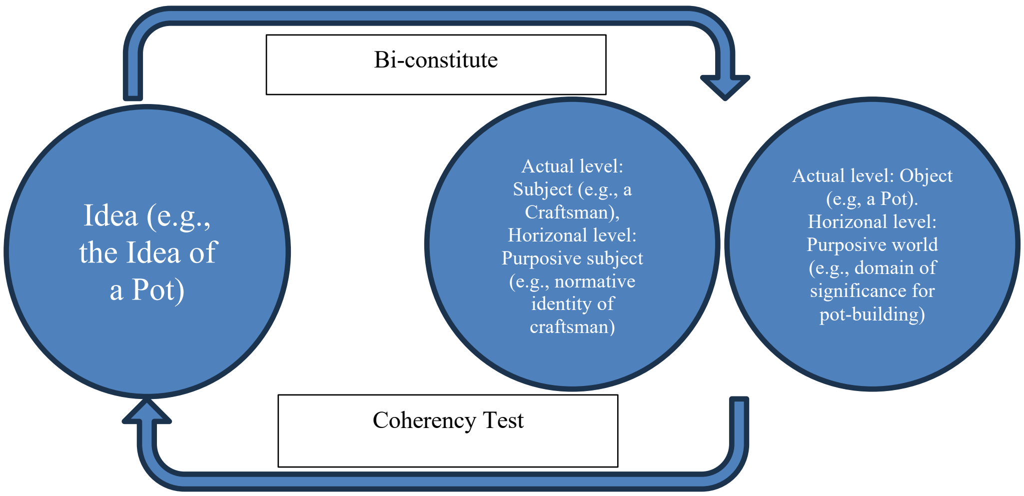

Figure 1: Schema of technical reason

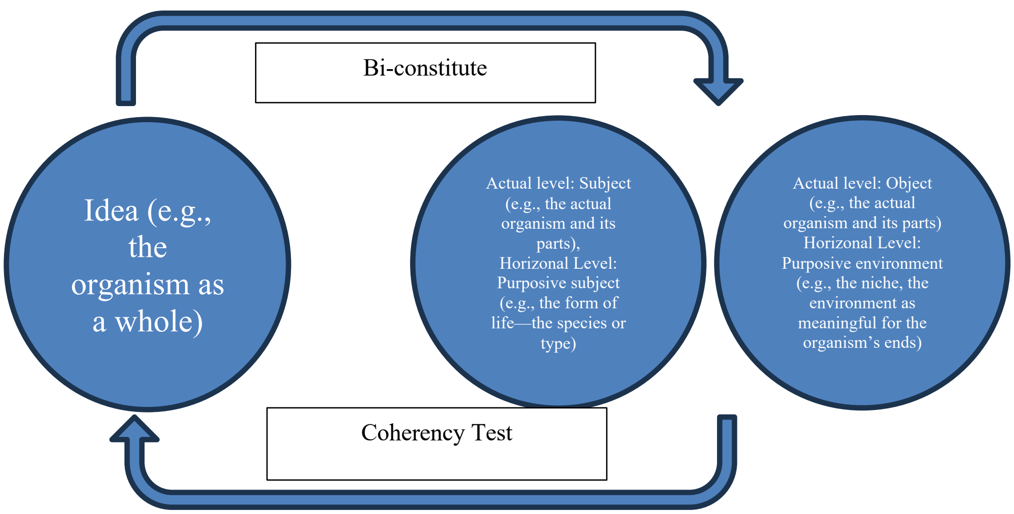

Figure 2: Hypothetical interpretation of Kant's schema of natural agency

Kant claimed that organisms are to be understood on the model of Idea-mediated technical agency. In chapter 2, we developed a robust account of Idea-mediated agency. I will now extract the essential elements of that account (figure 1), hypothesize the principles a natural agent must obey to be analogous to such an agent (figure 2), and validate these principles as Kant's by grounding them in the text.

To briefly recap Chapter 2, Ideas are comprehensive normative structures that generate practices. This generation occurs across two distinct but interdependent levels: at the horizonal level, the Idea constitutes a normative practical identity (the purposive subject) and a purposive world (the domain of significance disclosed by the practice); at the actual level, the Idea constitutes a concrete agent (the practitioner) and a concrete product (the object or end that organizes the materials). In the example of pot-making, the Idea of a pot guides the practice: at the horizonal level, it constitutes the normative identity of the potter (the standards of skill, the telos of craftsmanship) and the purposive world of pottery (clay as meaningful material, tools as extensions of the craft, the market as the context of exchange); at the actual level, it constitutes the particular potter (this person with these hands) and the finished pot (this vessel on the wheel). The two levels are not separate domains but mutually constituting: the horizonal gives meaning to the actual activities, while those actual activities continually test and articulate the horizonal norms.

In the case of the natural agent, however, practitioner and product seem to form a closed loop: the organism is both the agent that acts and the product of its own self-production. Yet the agent must draw materials and energy from outside itself to fuel this very act of self-production, meaning the self-constituting practice logically presupposes an external source. This is where we can hypothesize the full scope of the analogy to emerge at the horizonal level. Just as the Idea of a pot bi-constitutes, for the craftsman, both the purposive subject (the normative identity of the potter) and the purposive world (the domain of significance of pottery—clay, tools, market, tradition), so too does the Idea of the natural agent bi-constitute both a purposive subject and a purposive environment. This bi-constitution unfolds through four interlocking principles: [48]

1) First, the organizing being's form—its normative identity as described by the Idea—orients its parts (or subsystems) into a functional, self-sustaining whole. This is the horizonal constitution of the purposive subject, establishing what the organized being is and what counts as its proper functioning.
2) Second, the organized being's parts exist in reciprocal causality: each is both cause and effect of the others, and together they produce and maintain the whole. This is the actual constitution of the organized being, the ongoing activity through which the organized being sustains itself. [49]
3) To persist, the organized being requires selective input from its environment, which its form (Idea) thereby renders purposive, constituting a niche of

[48] Jakob von Uexküll's theory of the organism shares many similarities to Kant's schema. He shares Kant the idea of how an organized being constitutes its environment by selectively picking out certain input to both perceive and respond to. Von Uexküll was heavily inspired by Kant to arrive at these conclusions. Von Uexküll, however, does not seem to acknowledge the dynamic means in which environments and organisms change. Organisms are windowless monads, existing in pre-established harmony with each other. Here von Uexküll follows Leibniz more than Kant. Kant, on the other hand, by incorporating pragmatic criteria of success for organisms, can accommodate these dynamics by showing how failure implies a (relatively) independent dynamically self-organizing nature. This puts pressure on the organism to change its behaviors or die.
[49] This may include the actual environments it constructs (e.g., nests, burrows, hives) that, for our intents and purposes, count not as 'external' to the being, but its constitutive parts.

potential significance. This environment is itself composed of other self-organizing systems, forming an ecology of reciprocal dependence. This is the horizonal constitution of the purposive environment, the domain of significance in which nutrients, threats, and meanings show up for the organized being's self-maintenance.

4) Successful acquisition of environmental input leads to persistence of the whole (whether through growth, reproduction, etc.), while failure pressures the system toward collapse or adaptive reorganization. This is the coherency test between actual and horizonal constitution: the actual acts of the, e.g., organism—its feeding, growth, repair—continually realize and articulate the horizonal field, either reinforcing the norms of its form of organization (e.g., Idea) or, in cases of failure, forcing a revision of how it attains its self-preservation.

Let us turn to see if the text validates these four principles. That the Idea guides the organization of parts into a whole (1) derives from the first criterion Kant uses to define self-organizing beings as laid out above. At the horizonal level, the Idea of, e.g., the whole organism constitutes an immanent normative identity: the form of life defined by the end of self-preservation, equipped with capacities, and subject to standards of success and failure to be the form it ought to be. The Idea of the whole organized being structures the role of each part: the parts "are possible only through their relation to the whole" (5:373). To be a natural end, an organized being must fall under an Idea "that must determine a priori everything that is to be contained in it" (ibid). The form of, e.g., the organism thus determines the functional arrangement of its organs, giving each part its significance in contributing to the maintenance of the entire being.

The self-producing dynamic (2) is captured by Kant's second criterion: the parts are combined "by being reciprocally the cause and effect of their form" (5:373). At the actual level, this reciprocal causality is the mechanism of self-maintenance: each part is both cause and effect of the others, and together they produce and maintain the whole. They act as "an organ that produces the other parts (consequently each produces the others reciprocally)" (5:374). The organism is thus its own artifact—practitioner and product coincide in the circular dynamic of self-production. Furthermore, each subsystem "must be considered as organized" (5:377), suggesting a hierarchy of self-organizing forms.

That this self-organizing activity constitutes the purposiveness of its external environment (3) is evident in how the agent relates to its environment. At the horizonal level, the Idea of the agent constitutes practices that require a purposive environment to satisfy. This is the agent's 'ecological niche,' a domain of significance in which features of the agent's environment become salient to the agent's practices of self-maintenance. For example, a bird sees certain fruits as conducive to satisfying its hunger; a frog sees certain lily pads as suitable for sitting; a hawk finds a mouse purposive for its persistence. This is a "hypothetical" purposiveness (5:378). When the organism acts—for instance, when a bird sets out to eat some fruit—it tests a hypothesis[50]: it sees the fruit as an extension of itself, as falling under its Idea. But this relation is problematic; the fruit could be poisonous or indigestible. Success in the hypothesis ('if I eat that, I will be nourished') means the organism has successfully integrated that part of the environment into its activity. This is what Kant means in stating that "the purpose of nature must be extended to everything that lies in its product" (5:377). Everything the organized being uses for its purposes thereby falls—problematically—under its Idea, constituting the environment as a purposive, yet contingent, correlate of its own form and activity.[51]

Each organism's purposive environment overlaps and interlaces with those of others, forming an ecology of reciprocal dependence. Kant claims, in the teleological, regulative mode of judging nature, that "everything in the world is good for something, that nothing in it is in vain; and by means of the example that nature gives in its organic products, one is justified, indeed called upon to expect nothing in nature and its laws but what is purposive in the whole" (5:379). Hence, each organized being is purposive for others; each being's activity helps to form the environment that sustains another's. Kant illustrates this with the example of the pine forest:

> no soil is more favorable to pine trees than a sandy soil. Now the ancient sea, before it withdrew from the land, left so many sandy tracts behind in our northern regions that on this soil, otherwise so useless for any cultivation, extensive pine forests grew up, for the irrational eradication of which we frequently blame our ancestors; and here one can ask whether this ancient deposit of sandy strata was an end of nature for the sake of the pine forests that were possible there (5: 367)

[50] I am talking a bit loosely here: it is important to be clear that organisms do not have hypotheses. Since they are themselves Ideas, they *are* hypotheses. They test themselves.
[51] In a similar way, we saw in chapter 2 that the "purposiveness of nature" is a necessary presupposition for the practical use of our reason (5:455), a thesis we saw Kant defend as early as the Critique of Pure Reason.

The sandy soil serves as a purposive condition for the forest; the forest, in turn, becomes a purposive condition for insects and other life. The way an agent organizes itself may become purposive for another agent, and vice versa. Using teleological judgment, we can even judge ourselves as a means to nature:

> the plant-eating animals exist in order to moderate the excessive growth of the plant kingdom, by which many of its species would be choked; the carnivores exist in order to set bounds to the voraciousness of the plant-eaters; finally, humankind exists in order to establish a certain balance among the productive and destructive powers of nature by hunting and reducing the number of the latter. And thus the human being, however much he might be valued as an end in a certain relation, would in another relation in turn have only the rank of a means (5: 427)

Such a perspective is possible precisely because the principle of teleology is a regulative rule for judging nature. But this rule applies to the object itself: the reciprocal interdependence is objective, not merely subjective. It reveals nature as a system of overlapping, bi-constitutive processes where agents reciprocally create 'niches' in which their environments are purposive for it, and they are purposive for their environments.

Success, failure, and adaptation are inherent to this model. The coherency test (4) between the actual and horizonal levels determines whether the organism's self-organization succeeds or fails. Nature may or may not be conducive to an organism. A fox's eyes function to detect prey and shelter. When the fox's activity encounters an environment conducive to itself—say, where it is able to find and capture prey—the organism succeeds in its self-organization. [52] If it does not encounter such environment, it faces pressure to adapt its behavior or even its capacities. The normative identity of the

[52] Such a success, for Kant, might be accompanied by pleasure (or an analog thereof): "pleasure is the representation of the concordance of the object or the action with the subjective conditions of life, that is, with the faculty of causality of a representation with a view to the actuality of its object" (5: 9 n.)

organism—what it means to be a fox—is not a static template but is continually tested and approximated through the organism's actual engagements with its purposive environment.

Kant identifies concrete markers of successful self-organization: reproduction of the species (it "generates itself as far as the species is concerned") and "growth" of the individual (5:371). More dynamically, he acknowledges a response to failure, discussing how organisms adapt to injury as the "self-help of nature" (5:372). This implies that the gap between an organism's ideal form and its finite, contingent existence generates a teleological pressure. The being is driven to alter its activity and reorganize its internal and external relations to better realize a holistic harmony with its surroundings. Such adaptive pressure (which also suggests an openness of Kant's account to evolution) is integral to Kant's analysis of organized agency, where striving is not toward a static ideal but a perpetual process of re-coordinating self and environment to better approximate the agent's Idea.

We have seen ample textual evidence to support these principles. Let us now consider their scope. It might seem that Kant is talking exclusively about living beings when he talks about natural ends. Yet he never says this explicitly. This suggests that Kant's framework applies to more than living organisms, but to all non-human agents. What might this look like?

Consider a convection cell, which is an organized, rolling motion of water that appears as water boils, say, on the kitchen stove. The 'whole' is the stable, circulating cell itself in reciprocal exchange with its environment, whereas the subsystems or parts include hot water at the bottom and cold water at the top (we could add the container holding the water as a part as well). Such a structure maps on precisely to Kant's theory of agency

1) As the water boils, a whole emerges that is structured to efficiently dissipate heat. At the horizonal level, the 'Idea' of the convection cell constitutes an immanent normative identity: the form of the cell defined by the end of heat dissipation, equipped with the capacities of buoyancy and thermal exchange, and subject to standards of stability versus chaotic breakdown. The subsystems (parcels of hot and cold water) are organized by the whole cell to efficiently transfer heat from the bottom to the top.
2) At the actual level, hot water is less dense, so it is buoyant and rises to the top. Once at the top, it transfers heat to the cooler air above, making it cool down. This cooler water parcel is denser, so it sinks back to the bottom. At the bottom it is heated again, becoming a hot water parcel. Subsystems of hot and cold-water parcels both cause and effect each other in this way, constituting and giving rise to the whole convection cell. Practitioner and product coincide: the cell produces itself through its own circulatory activity.
3) The purposive environment of the convection cell is constituted by its form: the selective input for the cell is a certain amount of heat from the stove. This input is selective in that the system cannot persist with just any input (stirring, light, or sound): it requires continuous flow of a particular quantity of heat

from below. In turn, the boiling pot outputs steam and heat into the kitchen environment, which can be used as selective input for the persistence of yet other systems (providing input for other self-organizing beings, constituting in its way an ecology)

4) The coherency test between actual and horizonal levels determines success or failure. The convection cell succeeds to the extent that the pot successfully draws heat from the stove. Yet if the input is too low, the cells weaken and vanish. If the heat input is too high, the cells weaken and break down into chaotic bubbling. Since the system is relatively simple, we might say, is not "complex" enough to warrant adaptation or "self-help" if the environment no longer provides the right amount of heat.

Such systems were, of course, not understood by Kant at his time. But a simple application of his theory of agency proves that his theory has lasting relevancy, especially when deprived of anthropomorphic references and considered at the level of these formal rules of reflection.

The application of this framework to a convection cell may seem to strain the Kantian notion of an 'Idea.' However, this strain is alleviated when we distinguish between the *subjective representation* of an end in a conscious being and the *objective, immanent principle* of organization in any self-sustaining system. In the human potter, the Idea is a reflective normative identity: a conscious set of plans and activities that the agent can represent to itself and deliberately enact. In the non-human system, the 'Idea' is an immanent (non-reflective) normative identity: the stable, unconscious functional form towards which the system dynamically tends. Under this description, the 'Idea of the convection cell' is nothing more than what guides the specific, heat-dissipating roll pattern that selectively maintains itself against environmental perturbations—constituting at the horizonal level the purposive subject (the form of the cell) and the purposive environment (the heat gradient as a domain of 'significance' so that the cell persists). By abstracting the 'Idea' to this formal, functional level, we Kant's analogy seems to hold while extending its scope beyond the realm of organisms.

To take a step back, we must acknowledge that the four rules of reflection I have derived are the formal schema we use whenever we judge an object to be a self-organizing system. However, even with these rules guiding judgment, we still cannot know whether natural beings are 'really' agential in themselves. Yet instead of demanding theoretical certainty, we can derive confidence in the validity of this principle from the success we achieve when we judge in accordance with it. As we have seen, success here is defined by the explanatory power that this framework grants us. It is *felt* when the regulative lens of judging entities as analogous to Ideas guides scientific inquiry—for instance, by prompting us to discover the chemical mechanism by which plankton produce the oxygen necessary for human life, thereby solidifying a structural link between two seemingly disparate organized beings. In fact, such systematicity is a necessary presupposition for science; without it, we would be left with a heap of disconnected, contingent laws. This is why the feeling of success is no ordinary feeling, but is the universal feeling of harmony between our own cognitive faculties and nature. The wild success of self-organizing models across thermodynamics, economics, biology, and psychology (see chapters 1 and 5) further enriches this feeling, being a testament to the fruitfulness of these rules of reflection.

IV. Ecological Agency

We saw earlier that the boundaries between mechanical and teleological principles for Kant are not rigid. Rather, each guide our investigations into material nature in a particular way: mechanism by viewing all matter through the eyes of linear laws of motion; teleology by viewing entities in nature as belonging to purposive wholes. Since Kant's times, thermodynamics, natural selection, ecology, and systems thinking have blurred the boundaries between mechanism and teleology. These advances give mathematical form to *non-linear* mechanistic processes. Kant, recall, understood mechanisms to be linear processes. It seems then that Kant imposed false limitations on the explanatory scope of mechanisms. This result has led to responses from several critics: some have called for the elimination of teleological causes as a valid description of things,[53] while others claim that self-organizing beings constitute the most fundamental level of reality, no longer merely 'regulating' our investigations.[54] In both cases, the entire

[53] See Millikan (1989), Neander (1991) and also Wright (1973) who seek to naturalize and replace 'purpose' talk with functions grounded in evolutionary selection history. Chemero (2009 and below). See also Desmond and Huneman (2020) for contemporary ways to understand agency in biological systems.
[54] See for instance the proponents of 'autopoeisis'—Di Paolo (2005), Varela (1979), Moreno and Mossio (2015), and Walsh (2012, 2015)—who seek a naturalized account of agency, where self-organizing agency is a unique causal principle within nature 'constitutive' feature of biological systems. See Gambaretto and Nahas (2022) for an overview of this interpretation and its relation to Kantian teleology.

Kantian critical perspective is seen as creating arbitrary restrictions on what we can know. Kant, however, also saw the guiding principles of science (the theoretical Ideas) as problematic and merely hypothetical: we can never fully know them or predict how they will guide inquiry (see chapter 2). As they get unpacked in our various scientific practices, we learn about how they guide our thought. Hence, they are not arbitrary restrictions, but the dynamically evolving grounds that make our explanations possible.

It is crucial to remember that Kant thought that mechanism was *also* a regulative notion. This means that new advances in mechanistic frameworks does not change the regulative status of these mechanisms. For Kant, the founder of a science only has a weak grasp of the principles of their science: it takes time for the guiding principles to become transparent to us.[55] The evolution of mechanism to encompass nonlinear processes, therefore, merely means that we have a new grasp on the Idea of mechanism that guide our investigations into nature. Our understanding of the Idea did not emerge in a vacuum, but from a dialogue between Newtonian mechanics, the crude biological knowledge of Kant's age, and the later advances in biology and physics that prompted new ways of investigating nature. Crucially, it seems as if Kant's model of agency as expressed in the regulative principle of purposiveness too has been refined and its schema further articulated, particularly considering contemporary investigations in systems thinking (see chapter 5).

[55] "It is too bad that it is first possible for us to glimpse the idea in a clearer light and to outline a whole architectonically, in accordance with the ends of reason, only after we have long collected relevant cognitions haphazardly like building materials and worked through them technically with only a hint from an idea lying hidden within us" (A834-5/B862-3)

My aim in what follows is to clarify our understanding of Kant's legacy for contemporary discussions of non-human agency by juxtaposing Kant's conception of agency with a prominent contemporary alternative: an account of "ecological" agency that links a realist theory of psychological "affordances" with cutting-edge developments in systems thinking. I will argue that this contemporary theory fails to offer a convincing replacement of teleological conceptions of agency. Instead, by necessarily yet unknowingly presupposing a Kantian conception of agency, its reconstruction of agency affirms rather than negates a teleological conception of agency. This lends credence to the idea that Kant's principle of purposiveness remains unconsciously operative within contemporary psychology and systems thinking. Furthermore, Kant's conception of agency—albeit ecological—offers a compelling alternative to the sense of agency in the ecological agency literature.

Recently, the notion of "ecological agency" has become quite popular. Central to this theory is a robust threefold ontology: an organism's capacities, objective facts about an external environment, and the affordances that connect the two are all construed as real entities. I will first discuss the theory, how it differs from Kant's, and why it is inadequate to its intended goal of providing a model for agency. At best, it only captures such agency insofar as it tacitly *presupposes* the Kantian framework.

James J. Gibson founded the notion of ecological psychology in the late 1970's. Crucial to his description of animals are what he calls affordances: 'The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill" (Gibson 1979, 127). Affordances are features of the environment that exist relative to the animal's capacities to potentially experience them. A flat, clear path affords walking to any organism capable of walking; a doorknob affords grasping to organisms that can grasp; a sandwich affords eating to those that have the capacity to eat sandwiches.

Gibson considers himself a realist about affordances since he claims that we can 'directly' grasp affordances without mediation by mental representations, inferences, or theoretical constructions. Moreover, we perceive the *meaning* of the external world directly, in that we perceive what the environment affords us. We do not abstract data that we then synthesize and interpret, nor do we apply theories to experience in a two-step process. We directly access reality by directly observing affordances.

Philosophers have debated what kind of ontology Gibson's theory amounts to. Because it incorporates many elements of dynamic systems theory, Anthony Chemero's theory is the most relevant for our purposes. Chemero's theory of 'radical embodied cognitive science' asserts—using models generated by dynamic systems theory—that cognition is fundamentally constituted by the body-brain-environment system. Affordances, abilities, and the nervous system are all dynamically coupled together. For example, walking emerges from a dynamic interaction between leg biomechanics (abilities), the nervous system, and properties of the ground relative to our abilities (the affordances).

To ground the reality of affordances, Chemero draws on Ian Hacking's concept of 'entity realism.' Hacking famously held that the reality of an entity is demonstrated by its causal role in experimental practice—that is, if we can reliably spray, measure, or manipulate it, it must be real. In essence, real things are those we can use. Chemero extends this reasoning to affordances: they are, almost by definition, embedded in the practical sphere. We continually engage with them; they are the very medium of our interaction with the world. Yet if, as Chemero states, "relations are real, with the same status as the things that stand in the relations" (Chemero, 141), then the reality of the relation (the affordance) itself seems to demand the reality of its relata: the objective environment[56] and the concrete capacities of the subject.

In entity realism, it seems that affordances must be more fundamental that "the things that stand in the relations" (ibid). Our everyday experience is practical, engaging with tools and equipment rather than an objective environment independent of us. Yet elsewhere, to justify the existence of affordances, Chemero seems to *presuppose* the objective environment: "among the things we experience are relationships between things; so relations are real, with the same status as the things that stand in the relations" (Chemero, 141). Elsewhere, he argues that we 'discover' affordances as if they already existed prior to our awareness, such as when "very young children actively explore their environments in order to discover affordances" (Chemero, 201). Since affordances are defined as relations between capacities and the objective environment, they seem to actually depend on the pre-existence of the relata. Yet in entity realism the opposite is the case: the existence of the relata is grounded in the relation (the affordance).

Regardless of the how Chemero escapes the circularity in grounding his ontology (i.e., what comes first, the affordances or the entities they relate?), he still requires a

[56] Confusingly, Chemero sometimes seems to employ the term 'environment' to refer to the 'subjective' environment of the organism, which is their niche composed of the affordances that define the organism's lived environment.

burdensome threefold realist ontology, positing the reality of 1) subjects, 2) a subject-independent environment, and 3) a subject-dependent environment. While traditional realism readily accepts (2) and often (1), the third component—a subject-dependent environmental niche constituted by affordances—marks a significant departure.

Chemero's triadic structure is not unique among ecological approaches. Such theories typically analyze agency by examining how a given capacity 'matches' an objective environmental condition to generate the conditions for action. Jennifer Morton's (2010) "ecological" theory of deliberative norms exemplifies this shared structure. While her focus shifts from direct perception (Gibson) or embodied cognition (Chemero) to the internal process of deliberation, the underlying logic is analogous. For Morton, the rationality of a deliberative style is not governed by a universal, internal rule, but by how well it fits a specific environment. Just as a perceptual capacity adaptively matches a perceptual environment to yield an affordance, a style of deliberation adaptively matches psychological capacities to a deliberative environment. In an emergency room, for instance, the adaptive norm is to use fast, heuristic reasoning; the slow, systematic deliberation of a scientist would be a poor fit. Consequently, norms—like affordances—function as conditions for action that are always indexed to objective environments and subjective capacities.

Morton, like Gibson and Chemero, treats affordances—in her case, deliberative norms—as matching relations between subjective capacities and objective environments. This realist, triadic model reveals a fundamental divergence from Kant's dynamic conception. For Kant, agency is not a matching operation between two given relata, but an Idea-mediated practice where the relata are not given but are bi-constituted. There is no pre-given ontology of objective environments or fixed capacities—such an ontology is unintelligible apart from our practices. We use norms to navigate various identities and environments, but those identities and ends are never given from the outset, requiring us to 'pick' which norm to match the environment and our psychological state. Rather, the norm we use to guide our action depends on what action we want to do (e.g., fast medical care). It then generates an environment purposive to this end (e.g., patients with injuries of varying seriousness acquire significance) and the capacities we must take on mentally to achieve that end (e.g., rapid-reasoning medical professionals). Hence, the practice is more fundamental, establishing the significance of the relata. We are much more familiar with our practices than abstract mental capacities and a mind-independent reality. It is strange to think we must have some robust metaphysical capacities to know what our psychological states and environmental states are independent of us, where the proper norm of deliberation requires matching these given sets of states. Kant might be right to think we can never grasp the high abstractions contained in the latter two. Even if we were Gods that could, Kant offers a much less abstract and more parsimonious ontology than what the ecological agency theorists require.

To clarify this contrast, consider a fox pursuing squirrels under a Kantian framework. The fox's capacities—to identify, chase, and consume squirrels—are unintelligible apart from fox-like practices: to eat squirrels, and ultimately, to maintain itself. Likewise, the fox's purposive environment—the squirrels, the forest, etc.—are unintelligible apart from fox-like practices. The fox does not live in an 'objective world' but interrogates its environment and thereby constitutes it in a way where everything encountered is interpreted towards the fox's practice of maintaining itself. The fox's survival remains a continual problem—one whose resolution depends on whether nature proves conducive to itself. A successful hunt means that the fox's horizonal constitution of nature as a place purposive to itself was successful. Failure, by contrast, means that the purposive constitution of the external environment did not go as planned, exerting pressure to reorganize—whether within the fox's own behavioral repertoire or, at the species level, through adaptive evolution. In both success and failure, the environment is not a stable reality to be matched to a stable set of capacities, but a correlative pole in a ceaseless, bi-constitutive practice of self-organization.

In the ecological agency framework, as articulated by Chemero, the standard view treats the fox's capacities as abstract, detached from practical context—mere capacities to move legs, detect odors, or eat meat—that are matched with an objective environment comprised of abstract physical laws. This collision between two abstractions generates affordances, describing how the fox uses its capacities to engage the environment, as when a squirrel appears and the hunt begins. Yet the fox's own end of self-preservation remains irrelevant to the description. Chemero's ecological approach instead reconstructs the scenario mechanistically, identifying determinate environmental conditions, the affordances those conditions offer, and the specific capacities of the fox to engage with them—capacities that are in turn determined by the organism's nervous system, creating a dynamic feedback loop between body, brain, and environment. He thus proposes a concrete triad of nervous system, body, and environmental affordances. Robust metaphysics aside, the question remains whether this reconstruction actually captures anything properly agential at all. Is agency nothing more than a match between two puzzle pieces—an objective environment and a set of capacities? Where is the action? One wonders if in this mechanical reconstruction, Chemero has substituted theoretical mechanisms for the agency itself.

Although Kant would disagree with the metaphysical realism latent in Chemero's description, he would welcome the possibility to explain agency in mechanistic terms. After all, the principle of judging things teleologically was meant to *open up* the possibility of mechanical explanation. The very role of this principle is to provide a context that guides mechanistic explanation, not to replace it. Crucially, however, this means that any such mechanical account *presupposes* that we have already judged the fox-environment system as purposively self-organizing, using the rules of reflection outlined earlier.

Does Chemero presuppose agency in his theoretical construction? Yes. This is clearest in the way in Chemero's reconstructs the body-brain-environment system. He seems to highlight only those capacities and affordances that sustain the agent. Yet in reality, there are an *overwhelming* number of affordances that could be lethal to the agent—a tree affords a bird a fatal collision, just as land affords a fish desiccation. Yet the very selection of which capacities and affordances are deemed "relevant" and which irrelevant (such as the tree affording crashing) embodies a teleological judgment. Put simply, Chemero can demonstrate how the elements in the body-brain-environment system are interconnected, but the rationale for identifying the salience and significance of all these terms in the beginning must be based on a teleological judgment. Affordances are chosen for no other reason than that they serve the organism's self-organization—the very phenomenon his theory seeks to explain. Thus, the very act of selection—guided by the rules of reflection—presupposes the agency that his mechanical description relies upon but cannot ground.

Chemero is no fan of a teleological understanding of agency. He praises Newton for sidestepping teleological speculation through the formulation of general laws of motion. Chemero extends this analogy to psychology, suggesting that "once one has mathematical covering laws for psychology, laws that predict the behavior of agents in their environments with great accuracy, there may be no need for teleological explanations in psychology" (Chemero 77). Kant would readily agree that teleological explanations are not valid causal explanations. Their function, for him, is not to explain, but to guide explanation by providing a context for explanation—for instance, by helping us discern structural forms and the processes that aid in their maintenance. Hence, while agency must be presupposed to explain certain processes mechanistically, the agency itself cannot be explained mechanistically—it can only be identified when we judge entities as if they were purposive.

The fundamental question pitting ecological agency theorists against Kant is whether agency is a mere mechanistic matching relation or not. If it is, we might as well rule out the possibility of human autonomy. My free actions are not possible if all my actions are merely pre-ordained matches between my capacities with certain environmental conditions. My actions would be dictated entirely by my contingent psychological and environmental states. My own agency is incapable of organizing these states into an integral unit. That is, I have no way to control my fleeting desires and environmental states. Yet it is only when my desires are under check so that I can make complex plans, and my environment is subject to rules of my own creation, then I am a full-fledged agent. Only if we already presuppose a sort of hierarchical self-organization (e.g., by presupposing 'higher order' desires) can we justifiably assert some kind of autonomy: but this, again, moves beyond mere matching relations between two givens—our psychology and the environment. Matching is a mere mechanistic reaction, not an action. Instead, in genuine agency, we act by *creating* the structure of our own internal states and the external environment.

In summary, Kant's theory is unique from ecological agency theorists in that the agency of the organized being does not depend on matching certain given capacities with certain environmental conditions. Agency is not a mutually adjusting fit between two pre-determined puzzle pieces. Agency requires self-organization or self-constitution, which in turn requires some analogy with purposive agency to be intelligible in non-humans. Additionally, Kant's framework does not require a robust metaphysics that takes the external environment as something given, but as answers to a problem posed by the organized being. Such 'questions' posed by the organized being frame nature selectively. If the framing is successful, the organized being is rewarded: more (e.g.) energy, more persistence. If the framing is unsuccessful, the organism perishes or is pressured to adapt. Nowhere in this scheme must capacities or environmental conditions be posited as objectively specified. To do so is to perform a mechanistic reconstruction of the agent and its environment, a perspective that Kant thinks is a valid form of explanation. But this reconstruction depends on taking the teleological framework for granted, the framework that identified a structural, agential whole worthy of explanation. Insofar as such mechanistic explanations are successful, they affirm, rather than replace, the teleological judgment that made such explanations in the first place. As we will see in chapter 5, the very science of dynamical systems that Chemero's body-brain-environment model depends upon presupposes the principle of purposiveness down to its core.

V. Conclusion

This chapter has argued that Kant's framework of self-organizing agency, derived from the regulative principle of purposiveness, offers a cogent alternative to contemporary ecological theories of agency. Against ecological accounts that reduce agency to a matching relation between pre-given capacities and environments, Kant presents a bi-constitutive model: the organized being is an Idea-mediated being that dynamically constitutes both its internal structure and its external environment in a ceaseless striving for wholeness. What emerges from this contrast is a critical perspective that reveals the implicit teleological commitments within ostensibly mechanistic accounts, even those that embrace nonlinear causality. The Kantian framework does not seek to supplant mechanism, but to reveal its necessary regulative complement—the regulative principle purposive unity that guides our investigation and defines the very phenomenon of agency.

This Kantian insight lays the groundwork for a deeper engagement with the sciences of the twentieth and twenty-first centuries. As the sciences evolved, the problematic of teleology transformed. In the next two chapters, we will trace this transformation. Chapter 4 will explore how the neo-Kantian philosopher Ernst Cassirer, writing in the early twentieth century, interpreted the revolutionary developments in physics and biology. For Cassirer, these developments did not render teleological judgment obsolete but rather demonstrated its inescapable centrality as a constitutive feature of scientific thought. The regulative principle of purposiveness, he argued, had become embedded in the very structure of scientific explanation, guiding our understanding of relational wholes in ways that purely mechanistic models could not.

Building on this foundation, Chapter 5 will examine how this transformed understanding of teleology finds a direct expression in contemporary complex systems science, including the dynamical systems theory that Chemero's theory of affordances depends upon. Far from being an eliminable heuristic, the logic of purposiveness—the judgment of systems as striving, adaptive wholes—has become the conceptual core of fields ranging from theoretical biology and ecology to cognitive science and economics. The formal schema of self-organization derived here—with its emphasis on reciprocal causality, environmental constitution, and dynamic adaptation—provides a philosophical key to understanding these modern sciences. By returning to Kant, we have recovered a conceptual framework for understanding how purposive judgment remains indispensable for grasping the organized, agential character of nature itself.

Bibliography

Breitenbach, A. (2009). *Die Analogie von Vernuft und Natur.* Walder de Gruyter.

Chemero, A. (2009). *Radical Embodied Cognitive Science*. Bradford.

Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences,* 4(4), 429-452.

Düsing, K. (1968). *Die Teleologie in Kants Weltbegriff.* Bouvier.

Gambarotto, A. & Nahas, A. Teleology and the organism: Kant's controversial legacy for contemporary biology. *Stud Hist Philos Sci,* 93 (C), 47-56.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Houghton, Mifflin and Company.

Ginsborg, H. (1990). *The Role of Taste in Kant's Theory of Cognition*. Routledge.

Ginsborg, H. (2014). *The Normativity of Nature: Essays on Kant's Critique of Judgment*. Oxford University Press.

Guyer, P. (2005). Purpose in Nature: What is Living and What is Dead in Kant's Teleology? In *Kant's System of Nature and Freedom: Selected Essays* (343-372). Oxford University Press.

Jaspers, K. (1962). *Kant: From* The Great Philosophers*, Volume 1*. (H. Arendt, Ed.). Harcourt, Brace & World.

Kant, I. (1998). *Critique of Pure Reason* (P. Guyer & A. Wood, Eds.) Cambridge University Press.

Kant, I. (2000). *Critique of the Power of Judgment* (P. Guyer, Ed.). Cambridge University Press.

Kant, I. (2015). *Critique of Practical Reason.* Cambridge University Press.

Makkreel, R. (1990). *Imagination and Interpretation in Kant*. The University of Chicago Press.

McDowell, J. (1994). *Mind and World*. Harvard University Press.

McLaughlin, P. (1990). *Kant's Critique of Teleology in Biological Explanation*. The Edwin Mellen Press.

Millikin, R. (1989). In Defense of Proper Functions. *Philosophy of Science*, 56 (2), 288-302.

Moreno, A. & Mossio M. (2015). *Biological autonomy: A philosophical and theoretical enquiry*. Springer.

Morton, J. (2010). Toward an Ecological Theory of the Norms of Practical Deliberation. *European Journal of Philosophy*, 19 (4), 561-584.

Neander, K. (1991). Functions as Selected Effects: The Conceptual Analyst's Defense. *Philosophy of Science*, 58 (2), 168-184.

Varela, F.J. (1979). *Principles of biological autonomy*. Elsevier.

Walsh, D. (2012). Mechanism and purpose: A case for natural teleology. *Studies in History and Philosophy Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 43(1), 173-181.

Walsh, D. (2015). *Organisms, agency, and evolution*. Cambridge University Press.

Wright, L. (1973). Functions. *Philosophical Review,* 82 (2), 139-168.

Ypi, L. (2021). *The Architectonic of Reason*. Oxford University Press.

Zammito, J. H. (1992) *The Genesis of Kant's Critique of Judgment*. University of Chicago Press.

Chapter 4: Ernst Cassirer and the Self-Organization of Thought and Nature

I. Introduction

Chapter 2 established Kant's thesis of bi-constitution: Ideas guide practices that simultaneously bring into being purposive subjects (identities capable of pursuing certain purposes) and purposive worlds (domains of significance within which objects show up as meaningful to pursuing our purposes). This process operates at two inseparable levels. At the 'actual' level, the concrete performance of an action that is guided by an Idea constitutes a particular subject and a particular object: a person becomes a surgeon and cures a patient when they perform a successful surgery. At the 'horizonal' level, the Idea discloses both an entire field of possible agency—what I called the 'normative identity' of the subject, e.g., a doctor with a whole range of surgical and treatment capacities capable of a wide range of medical interventions—and a correlative domain of significance (a 'purposive world') within which objects and tools appear as relevant, e.g., for the doctor: a world of diseases, germs, nutrition and potential medical equipment. Bi-constitution thus rejects both the impositionist reading (where the mind projects order onto a passive nature) and the realist reading (where order inheres in nature independently of any normative practice). Instead, subjects and objects are bi-constituted (both horizonally and actually) through participation in Idea-governed practices.

Ernst Cassirer, the most prominent of the Marburg neo-Kantians, extends this bi-constitutive logic far beyond theoretical and practical reason to encompass the full range of symbolic forms, including myth, language, and art. For Cassirer, these forms are the very media through which anything becomes intelligible at all. These 'forms' are Ideas in the normative, practice-guiding conception I have argued that Kant bases his critical enterprise upon. As we have seen, these forms require a purposive world to achieve our ends. For example, the landscape becomes imbued with magical energy in myth; with representations of places relative to me and you in language; as a source of sublime artistic inspiration for the artist. Cassirer, like Kant, thinks that this purposive world must be self-organizing; it must be a meaning-laden totality that organizes itself around certain central rules and patterns. Nature is not something we intellectually construct; rather, Cassirer claims, the "forces of nature," are on a par with "all truly intellectual energies" insofar as "they all exist by virtue of being activated and by giving shape to themselves in this activation" (Cassirer 1996, 15).

Cassirer's position on nature is puzzling because it must navigate two commitments. He simultaneously thinks that the symbolic forms constitute the meaning of nature, while also believing that nature is self-organizing, constituting *itself* through its own endogenous, self-organizing dynamics. This apparent paradox threatens to collapse into subjective idealism (or complexity idealism, as I called it earlier, where nature is merely our projection) or into an uncritical realism (or complexity realism, where our symbolic forms merely copy a ready-made nature). The existing literature on Cassirer has tended to emphasize one side of this tension at the expense of the other. Intellectualist interpretations—exemplified by the work of Samantha Matherne (2021)—rightly stress that for Cassirer all access to nature is mediated by symbolic forms, and that the meaning of nature is always already shaped by the cultural frameworks we bring to it. However, these readings risk veering into claiming that nature's own self-organizing activity becomes epiphenomenal or is absorbed entirely into the self-organization of our minds. Conversely, more naturalist[57] readings—advanced by Thora Ilin Bayer (2001), Arno Schubbach (2022), and Tobias Endres (2021)—emphasize Cassirer's commitment to the natural continuity between human culture and non-human life and thereby accept that, for Cassirer, nature is an autonomous domain of form-generation. Yet these interpretations often struggle to account for Cassirer's insistence that nature is never given apart from the symbolic functions that mediate our grasp of it. Neither camp, I argue, has fully appreciated that Cassirer's achievement is rather to navigate between these positions by articulating a two-level bi-constitutive framework analogous to Kant's.

The key to understanding Cassirer is to understand what he calls the "formal imperative" (Cassirer 1996, 188). To interrogate nature—to ask it questions, to design experiments, to seek laws—we must *demand* that nature ought to conform to forms, rules,

---

[57] My reading is fundamentally distinct from other readings of Cassirer's 'naturalism.' While authors like Luft (2019) and Howe (1961) are correct in defining 'nature' as a singular dimension of man's culture, they fail to explain nature's relative independence from us, in that it self-organizes whether we want it to or not. Endres (2021), on the other hand, gives an excellent overview of Cassirer's philosophy of anthropology, showing how the faculty of expression can arise naturalistically. I do not wish to combat this view but give it deeper theoretical underpinnings. Endres argues how the expressive function emerges from organic life. But how can we say that organic life spontaneously emerges from the natural world, if not for a self-organizing *telos* towards autonomy within the natural world? Finally, my reading is in direct conflict from Kreis' (2010, 2021) view, who seems to see Cassirer's "expressivism" (where language is a natural expression of thought) as a "naturalization of the mental" (2021, 188). I find this view puzzling, especially since for Cassirer the distinction between nature and the mental seems to play no methodological role in his whole corpus. Instead, they are abstract poles that are generated from the holistic process of symbolization. They are not some abstract parts— presupposed at the outset—that need to somehow be glued together. In this way, I claim that nature is only mediated by the process of symbolic thought, yet we nevertheless can claim that it is relatively independent of us in this process. Furthermore, in this independence it is self-organizing. Hence, my arguments are entirely distinct from Kreis' aims.
For a reading that also simultaneously integrates thought and nature and aligns with my own view, see Keller's (2021) dynamic structuralist reading of Cassirer.

and principles. This demand is not a psychological disposition but a transcendental condition of reason itself. Without presupposing that nature is a self-organizing, law-governed whole, the activity of science would be unintelligible. The formal imperative thus horizonally constitutes both the purposive world of nature as a domain amenable to investigation and the normative identity of the subject as an inquirer who can legitimately expect nature to answer. Contra intellectualist readings, this horizonal constitution does not reduce nature to a mere mirror of culture; rather, it establishes the space within which nature's own self-organization can be discovered. Contra naturalist readings, our scientific discoveries are never direct insight about nature-in-itself but always a product of our interrogation.

At the actual level, the horizonal demand does not dictate which forms nature will disclose. When we pose determinate questions to nature—framed by specific theories, axioms, and experimental setups—nature responds. Its answers are not under our control. The form we uncover is simultaneously ours (insofar as our questioning gives it determinacy) and nature's own (insofar as the answer is not chosen by us). This is when the 'horizonal' and 'actual' moments of bi-constitution collide. We horizonally frame the meaning of anything we could potentially see in nature, but we cannot determine whether nature will actually abide by these forms. We experiment to find this out. The intellectualist is correct that there is no unmediated access to nature; the naturalist is correct that nature nevertheless answers back. Bi-constitution shows these to be two moments of a single process, not opposing commitments.

The present chapter develops this interpretation in two sections. Section II examines Cassirer's account of nature as self-organizing, tracing its origins to Kant's distinction between formal and material purposiveness and showing how Cassirer synthesizes them through the formal imperative. Section III turns to Cassirer's metaphysics of symbolic forms, demonstrating that thought and culture are self-organizing through the tripartite dynamic of energeia—activity that simultaneously limits itself, creates new forms, and adaptively preserves wholeness. I will argue that this process is precisely akin to Kant's thesis of bi-constitution. It will become clear that the self-organizing dynamic of culture is analogous to the self-organizing dynamic we must see nature as undergoing according to the formal imperative as established in section II. By reading Cassirer through the lens of Kantian bi-constitution, we can see that the intellectualist and naturalist strands in his work are two moments of a single bi-constitutive dynamic: the horizonal constitution of nature as a purposive domain for inquiry, and the actual constitution of determinate forms through nature's own responsive self-organization.

II. The Purposiveness of Nature: Formal Imperative and Material Success

Before turning to a discussion of Cassirer's conception of nature, I would like to flesh out further what his conception of nature and the natural is *not*. Cassirer thinks that "naturalists" commonly want to argue two things that he emphatically rejects.[58] The first

[58] See Hartung (2018, 158ff.) for an excellent summary of Cassirer's criticisms of naturalism in this sense

is to reduce entities like human autonomy, consciousness, and culture to mechanistic causal physical processes governed by general laws. This reductive naturalism proves inadequate by failing to account for the genuine autonomy of symbolic cognition—our capacity to project possible realms of meaning that cannot be explained as determined by fixed physical laws (which would render such activities 'fated' and thus devoid of autonomy) (Cassirer 1939, 151).[59] The second tenet of naturalism is the claim that it is intelligible to think of natural causal processes as existing completely independently of human cognition. In this view we passively derive knowledge of nature by copying an external world (that exists independent of our awareness of it) into our minds. Cassirer rejects any such conception of causal processes that are not mediated by the symbolic forms that science constructs and hence rejects this claim of naturalism. Such processes would have to be thought, and such thought is always invested with the frameworks, interests, and questions we approach nature with. Causal processes are incomprehensible without this mediation.

Cassirer does not merely disagree with the reductionism and realism of naturalists. He holds an entirely different conception of nature, one we might call 'systems-based.' This conception does not define the natural in terms of whether it can be reduced to a series of lower-level mechanisms governed by general, deterministic causal laws. Rather, what is 'natural' is best understood in terms of patterns of holistic, dynamic organization, adaptation, and emergence—an approach to nature that is shared by present-day

[59] Cassirer also claims—in agreement with Rickert's theory of the cultural sciences— that this naturalist conception—by positing general ultimate causes at the root of all explanation— also fails and abolishes the concrete individuality that is essential to cultural phenomenon. See (Cassirer 1937, 156ff.)

ecologists, systems scientists, physicists, mathematicians, economists, and neuroscientists.[60] In this sense, culture does not seem so anomalous: he claims that the sudden emergence of culture[61]—although it can "no longer be explained causally" (Cassirer 2000, 102)—is not some supernatural phenomena, but a further exemplification of a spontaneously generating, self-organizing process that occurs throughout nature. Cassirer thinks it is evident that "the study of nature has already taught us" this lesson (Cassirer 2000, 102), so it is prudent to now look to his studies of nature to elucidate this point.

To understand Cassirer's studies of nature we must first examine Kant's conception of purposiveness and self-organization in the *Critique of the Power of Judgment*, which profoundly influenced Cassirer.[62] To do this, we must first distinguish two senses of purposiveness in Kant, for this distinction will prove central to understanding Cassirer's own position.

---

[60] Cassirer emerges as a crucial intermediary between the Kantian concept of self-organization and cutting-edge research in systems and complexity science, which relies heavily on the idea of self-organization. See chapter 5. For a general introduction to the scientific background, see (Haken 2006; Holling 2010; Kauffman 1993; Mitchell 2009; Packard 1988; Prigogine & Stengers 1984; Waldrop 1992). Philosophical analysis is less available, but see (Bishop & Frigg 2016; Deacon 2011; Di Paolo, 2005; García-Valdecasas, 2022; Hooker 2011; Juarrero 2023)

[61] In *Essay on Man* Cassirer echoes this view of sudden emergence, claiming that Jakob von Uexküll's conception of animal life is continuous with, although qualitatively distinct from, spiritual life in this manner. See (Cassirer 1944, 44).

[62] See Schubbach (2022) for a thorough summary of the third critique's influence on Cassirer

A. Kant’s Two Senses of Purposiveness: Formal and Material

In the *Critique of the Power of Judgment*, Kant distinguishes two kinds of purposiveness. *Formal* purposiveness (formale Zweckmäßigkeit) concerns the harmony between nature’s forms and our cognitive faculties, a harmony we must presuppose for empirical inquiry to be possible at all (Kant 2000, 5:193–194). Formal purposiveness does not claim that nature *is* organized for our cognition; it claims only that we must judge *as if* this were so. It is a necessary regulative principle of reflection, without which empirical concept formation would be impossible. In this sense, nature must be a holistic structure amenable to our cognitive frameworks, albeit in a merely formal sense: we do not yet know the particular mechanisms and patterns that it makes use of to organize. As Kant argues, “we still cannot even make any purposive use of our acquaintance with nature for cognition unless nature itself has introduced purposive unity; without this we would not even have any reason” (A816–817/B845–846). This is the general horizonal principle that nature must be purposive for our understanding.

Material purposiveness (materials Zweckmäßigkeit), by contrast, characterizes not the purposiveness of nature for our cognition, but the purposiveness of individual natural structures (such as an organism). These beings are to be conceived ‘as if’ they were ‘natural ends’ whose parts are possible only through their relation to the whole, and whose parts reciprocally cause and effect one another (5:373–374). In the section “Critique of the Teleological Power of Judgment,” Kant is interested in making sense of nature’s apparent material purposiveness (see Chapter 3). By describing these beings as 'natural ends,' Kant resists the idea that these ends have anthropomorphic intentions, as well as Spinoza's idea that these ends are mere illusions of our imaginations. Rather, these beings are ends merely to the extent that they maintain their holistic structure. Their end is themselves, their own maintenance. Kant uses the term 'self-organization' to articulate organic purposive structures that are not maintained by some outside force (like God or humanity) producing them but are *self*-producing. To spell this out, Kant claims that natural ends have a reciprocal dependency between their parts (e.g., the organs) and the whole (e.g., the organism itself): "its parts (as far as their existence and their form are concerned) are possible only through relation to the whole" (5: 373). These parts must be "combined into a whole by being reciprocally the cause and effect of their form" (5: 373). While the parts produce the whole and make it possible, the whole produces the parts and makes them possible. Both are reciprocally determining in a dynamic feedback loop.

Formal purposiveness is the transcendental presupposition that nature's empirical laws are systematically unified for our understanding (a necessary condition for scientific inquiry as such), whereas material purposiveness is what regulates our judgment of particular natural entities as if they were organized for a purpose. Formal purposiveness is more fundamental than material purposiveness: we can treat nature as if it were purposive for our cognition (formal purposiveness) without ever claiming that any particular being is a natural end (material purposiveness). Conversely, treating a natural entity as if it were an end does not by itself justify the claim that nature as a whole is formally purposive for our cognition. Yet formal purposiveness is a necessary presupposition of the pursuit of science as a whole, including those sciences that themselves posit purposive wholes.

B. Cassirer's Formal Imperative

Cassirer is aware that Kant's third *Critique* tackles the problem of integrating nature's diverse forms into a logical system, asking: "whence comes this harmony between the natural forms and logical forms, and upon what is it based?" (Cassirer 1950, 124). How can we say that the forms of our thought 'line up' with the forms of nature? The question is not whether there are organic beings in nature, or that nature is self-organizing. For Kant, it is self-evident that nature is in the process of organizing itself, asserting that "one must also conceive of a world-organization in a unified body" (21: 212).[63] The question that Cassirer raises is whether the self-organizing forms we capture in science are to really harmonize with those in the natural world: how do we know that our theories and concepts accurately capture nature's structure? As we shall see, Cassirer and Kant obliterate this distinction between logic and nature, epistemology and ontology, by a demand of reason.

In an unpublished manuscript titled "On Basis Phenomena," Cassirer appears to follow Kant's approach, contending that nature's formal organization stems from reason's demand for structural regularity—an imperative arising from a particular mode of

[63] Kant also says: "Eventually, our all-producing globe itself (as an organic body which has emerged from chaos), completed this purpose [i.e., the purpose of organizing matter and the organization of the world to suit the needs of different species] in the mechanism of nature" (Cassirer 1998, 66).

apprehending the world. This demand for structure, he suggests, enables what he calls the 'discovery' of nature:

> [The discovery of nature] could not come about in any other way except by 'objective' perception—the experience of 'nature' as embodying forms and following rules…This discovery of nature in objective 'perception' is then confirmed, furthered, and supplemented, in a sense 'transcended,' in the discovery of the realm of Ideas, as the realm of 'pure forms' (Cassirer 1996, 190)

To 'experience' nature requires a certain demand: nature ought to follow forms and rules. The forms we do discover in nature, hence, arise initially from the forms we think nature 'ought' to have. We can then trace the forms we 'discover' back to the 'pure' forms that we originally invested into nature. In this sense, nature is not something with a determinate structure outside of objective thought. Rather, Cassirer thinks that nature does not even *exist* for us unless we can free ourselves from the immediate chaotic flux of our experience to demand this experience be organized according to certain forms and rules. Cassirer calls this the "imperative of the pure form" (Cassirer 1996, 188), an imperative that Cassirer traces to Kant[64] and Plato. This demand need not be a conscious declaration but is instead implicit whenever we compel nature to answer the questions we pose—questions which are implicitly shaped by certain forms,namely, principles, basic rules, and axioms that "constitute" them and their "conditions of its possibility" (Cassirer 1996, 188]. In posing such questions, we presuppose nature's capacity to answer them—an expectation that already assumes that nature conforms to certain patterns and structures. 'Objective perception,' which sees nature as ordered according to forms, is

[64] Anticipating the principle of formal purposiveness, Kant claims in the first *Critique* that we must approach nature "not like a pupil, who has recited to him whatever the teacher wants to say, but like an appointed judge who compels witnesses to answer the questions he puts to them" (Bxiii).

objective because it demands that nature conform to regular patterns. This demand does not guarantee which forms that nature abides by. We have no control over how nature answers our questions. Yet however nature answers our questions, the answer is always framed according to the form (the concepts, principles, basic rules and axioms) we imply in the question.

The formal imperative functions at the *horizonal* level of bi-constitution. It constitutes the *purposive world* of nature as a domain amenable to investigation—a world that is not a chaotic flux but a law-governed, self-organizing whole. Correlatively, it constitutes the *purposive subject* of the inquirer as one who can legitimately expect nature to answer, who possesses the normative identity of a scientist or objective perceiver rather than a mere creature of inclination. The formal imperative is not a metaphysical claim about nature-in-itself (realism) nor a projection of private mental states (subjective idealism). It is a normative presupposition of the practice of inquiry itself.

C. The Discovery of Nature's Material Purposiveness

When we interrogate nature under the formal imperative, what do we find? Cassirer's answer is that we discover self-organizing wholes at every level of investigation. Nature is not a collection of mechanical interactions reducible to efficient causes. Rather, nature is primarily comprised of purposive forms. These forms are neither lifeless, 'formal' and static entities that limit mechanistic reactions, nor are they quasi-living beings with intentions and desires. Rather, they are the *material*, dynamic self-preserving processes through which nature organizes itself.[65] Local mechanisms lose their role as building blocks of nature, functioning now only as parts within self-preserving wholes. Just as an organ can be understood only in relation to the organism in which it belongs, any physical object or interaction only makes sense in relation to the self-preserving whole in which it exists. For Cassirer, the formal imperative implies that we invest nature with *materially* purposive forms.

For Cassirer, form amounts to a whole with interrelated parts that is more than a mere aggregate of these parts (Cassirer 2000, 88). Rather, these parts interrelate 'as if' they were organized around a singular purpose, that purpose being to preserve the form or whole itself. Cassirer derives this idea from Kant, explicitly endorsing Kantian purposiveness in relation to his concept of 'form' in Volume IV of his *Problem of Knowledge* series.[66] In this work, Cassirer appeals to Kant's notion of purposiveness to explain the significance of the 'form' concept in then-contemporary biology and physics. Cassirer aligns with Kant in rejecting both the attribution of anthropomorphic purposes to natural forms and the notion that our judgments about purposive beings in nature are merely illusions produced by our imagination (Cassirer 1950, 122ff.). Cassirer, like Kant, acknowledges that purposive material forms are wholes that are central to nature. These natural forms are neither produced 'top-down' by a decree from God, nor are they produced 'bottom-up' as reducible to local mechanical interactions; rather, they are endogenously self-forming and thus self-organizing.

[65] In modern terminology, these forms are emergent properties of these local mechanistic interactions.
[66] See Cassirer (1918, 276ff.) for an extensive discussion on Kantian purposiveness.

For Cassirer and Kant, nature is organized according to the (materially) purposive forms that we invest into nature. This does not mean we know what these forms are in advance, nor that these forms are imposed onto nature. Their function for making science possible lies merely in the scientific practices that they make possible. As we continue to act successfully according to these practices, our thought and the material forms of nature become explicit. Hence, through our practices, we slowly learn about our own capacities (our normative identities as scientists) and the basic principles that govern how the world functions. Since the forms guiding these practices are not predetermined but instead emerge as our understanding of self and world deepens, we can expect a more explicit grasp of natural forms to develop as science progresses—from Kant's era, which grappled with the logic of Newton's physics and "Linnaeus' descriptive science" (Cassirer 1950, 127), to Cassirer, who engaged with the holism of early 20th-century physics and biology.

Cassirer argues that science has undergone a 'revolution'[67] since Kant's time. Cassirer labels this revolution holism (or 'organicism') when reviewing the cutting-edge biology of the mid-20th century. Here, Cassirer emphasizes that the concept of 'form' has become more explicit in science. Formerly, in the explanation of living beings, organic processes are reduced to causal relations between bits of matter. Holism, by contrast, is a complete reversal of this position: it is "not a case of the higher being deduced from the lower but precisely the reverse" (Cassirer 1950, 212). The 'organic' or whole holds explanatory primacy, as the purposive form establishes the context through which local

[67] Cassirer thinks that since Kant there has been a "revolution in our mode of thinking" ranging from physics and biology to psychology and linguistics. Cassirer echoes identical remarks about 'wholeness' in modern physics, biology and psychology in (Cassirer 2000).

mechanisms are intelligible. That is, a cell's function can only be understood within the larger context of the organism's overall purpose to maintain itself. Cassirer approves of the biologist Emil Ungerer, who argues "that we examine the processes of life so as to discover to what extent the character of preserving wholeness manifests itself" (Cassirer 1950, 213).

For Cassirer, the form concept is not only becoming explicit in biology; holism reigns in then-contemporary physics as well. By moving beyond Newton, physics of the early 20th century is freed from "the chains imposed on it by the mechanistic view" (Cassirer 1950, 211). This mechanistic view argued that all physical interactions could be explained by local, mechanistic reactions, where the behavior of a given system is determined by the interactions between its individual parts. In contrast, modern field physics challenges this perspective by emphasizing holistic, purposive structures—that Cassirer calls 'forms'—that cannot be reduced to their mechanistic components. Rather, these material forms establish foundational principles and aims that not only guide more localized, mechanistic explanations, but function as a 'whole' that gives rise to the specific structure of local mechanistic reactions. This idea, he says, "illustrates the 'primacy of 'field physics,' for the field appears as a whole that is not merely put together from several parts—e.g., electrons—but rather forms the very condition for their existence" (Cassirer 1950, 212–213). The principles that establish this overarching framework simultaneously guide our explanations of local interactions between, e.g., light, mass and energy, where these interactions gain their determinacy via an overarching 'purposive' whole.

Cassirer synthesizes material and formal purposiveness in a unique way. The formal imperative (formal purposiveness) and the investment/discovery of material forms in nature (material purposiveness) are two moments of a single process. The formal imperative—Kant's principle of formal purposiveness— is the *horizonal* condition that makes objective inquiry possible. It constitutes the purposive subject (the inquirer who can legitimately expect nature to answer) and the purposive world (nature as a domain amenable to interrogation, composed of objective laws and patterns). Yet each science invests its own form into nature. Each, that is, establishes its own set of comprehensive norms that guide practices intended to achieve the end of their respective inquiries. Cassirer seems to think that this is the *material* purposiveness that each science invests into nature. These material forms may be the electromagnetic fields of physics that give rise to particles, the organism of biology that gives rise to organic mechanisms, or the ego of psychology that gives rise to various mental states. Scientists might not be aware of what these material forms amount to. But as they become more aware of the forms that guide their own frameworks, they become clearer about nature's own form. Both are simultaneously constituted as scientific practice develops.

The process of discovering these forms can be messy. It involves viewing nature *as if* it were guided by self-organizing, purposive forms to see if this leads to valid explanations. At the *actual* level, the success or failure of these judgments is determined through fallible inquiry. When they succeed by leading to valid explanations, we feel the disinterested pleasure that confirms the harmony between our cognitive frameworks and nature's own structure. When they fail, we acknowledge that it might be unfitting to view nature according to that form. The intellectualist is correct about the horizonal conditions of inquiry; the naturalist is correct that nature answers in ways not up to us. Neither is complete without the other.

D. Nature and Thought

Only certain beings can pose questions to nature: that is, only self-organizing, autonomous, and reflective beings that are not solely governed by their material desires. It takes a certain self-organizing being to know a self-organizing nature. Guided by our desires or inclinations (even the desire that nature or our external environments be a certain way), we could never form an objective theory of nature, nor would we be able to act in a way that prioritizes the 'right thing to do' over what is personally advantageous for us. Only a being capable of this difficult task of organizing their desires and inclinations can do this. This is why Cassirer claims that Kant's categorical imperative—the idea that our moral conduct should abide by the moral law, apart from any inclination—*is* the formal imperative (Cassirer 1996, 189). The imperative that experience abide by forms is behind an ethical life (insofar as we demand ourselves to abide by the moral law) as much as it is behind our scientific knowledge (insofar as we demand ourselves to view nature as abiding by certain forms and laws). Any judgement of good or bad, right or wrong already presupposes that experience contains stable regularities and order. If such freedom from mental states seems far-fetched, Cassirer thinks we only need to look to the vast history of culture and the repertoire of its 'works' to prove this point. In addition, the mere fact that we can view the world scientifically and objectively, apart from our desires, and think carefully about the right thing to do apart from what would personally benefit us are all the proof we need of this autonomy.[68]

This entails that a being that self-organizes their thoughts can reveal the self-organization of nature. It is *through* the formal imperative and the material/formal principle of purposiveness that we do this. We determine the forms and frameworks of nature, but only *problematically*: we do not know how nature will respond. The significance of nature is hence framed by our questions but not literally constructed by us. Nature, considered independently of these forms, lies beyond the bounds of possible discovery and thus is an incoherent flux, meaning realist metaphysics is unintelligible. It is requisite that a self-organizing nature couple with a thought capable of these interrogations to unveil itself. It is hence incoherent to compare the forms we discover to the 'real' forms outside and independent of the forms of thought. Rather, the specific inner form of nature requires reflecting upon the forms we presuppose in our most comprehensive scientific theories. Nature, having been discovered by these forms, organizes itself according to these forms: there is no gap between the forms of thought and of nature that needs to be reconciled. These are the (albeit still fallible, as thought is always in self-development) forms of nature themselves. Both the self-organization of nature and the self-organization of thought go hand and hand. To expand on the dynamics of the self-organization of thought, let us now turn to Cassirer's metaphysical works.

[68] For a helpful discussion of Kantian autonomy along these lines, see (Keller 2018)

III. The Self-Organization of Thought: Cassirer's Metaphysical Works

Cassirer's plans for a fourth volume of *The Philosophy of Symbolic Forms* (contained in a collection that was published after his death) provide a key starting point for unpacking his conception of the self-organizing dynamics of thought and culture. Cassirer's unpublished essay "'Geist' and 'Life'" seeks to position his work in relation to *Lebensphilosophie*[69] and Bergson. I wish to use this article to show not how Cassirer makes sense of the Lebensphilosophie movement, but how he articulates that culture is in the process of self-organizing. In this context, revisiting his Symbolic Forms project, Cassirer seeks a 'comprehensive species' capable of unifying the symbolic forms—that is, a principle that binds together the diverse modes through which we construct meaning and organize our understanding of the world, such as language, myth, and science. (Cassirer 1996, 5). The significance of discovering this unity grows deeper and broader—ultimately serving to "illuminate the riddle of the becoming of form as such" (Cassirer 1996, 4). At the horizonal level, this unifying principle constitutes the purposive world of culture as a domain within which symbolic activities show up as meaningful, and the purposive subject of the cultural agent as one who is accountable to the norms of symbolic formation. The dynamics of culture also requires an actual level, where concrete acts of speaking, writing, and meaning-making both presuppose the horizonal level and evoke its adaptation. This section will show how the two levels of bi-constitution overlap. [70]

[69] Due to lack of space, I cannot do justice to Cassirer's full reconstruction of Lebensphilosophie. For such a summary, see (Skidelsky 2008, ch. 7) and Bayer (2001).

[70] Since I aim to focus on what this reveals about how culture/thought is self-organizing, I do not wish to reconstruct Cassirer's full relationship to the Lebensphilosophie movement. I also do not want to address

Cassirer begins his discussion by unpacking the *Lebensphilosophie* notions of Life and Form. Rather than their usual denotations, Cassirer and *Lebensphilosophie* uses them to signify broader principles. Life is a dynamic, continuous, productive activity, while Form is a static, fixed portion of this whole that is isolated from this flux. To use an example of an organism, Life would be the organism's constant activity, whereas the body of the organism would be its Form. Yet for the *Lebensphilosophie* movement, the concept of Life transcends what is typically understood as biological life. Rather, Life is the underlying principle behind *human* social and cultural agency as such. It pertains merely to the synthetic, unifying, self-organizing function behind e.g., language, myth, art, and science. In this way, the function of Life is to create and integrate these symbolic forms in an 'undifferentiated unity':

> instead of the *discretio* of forms, which thought subsequently differentiates, we have here a continuing crossing over and crossover, a constant unbroken movement from one extreme to the other (Cassirer 1996, 6)

The unbroken continuum of Life is broken when, for instance, we dwell in the products (the forms) that Life creates. We abstract ourselves from Life when the worlds created by myth, language, art, technology, and science become all-encompassing, forgetting the

---

another concern which seems to arise in terms of naturalism. In "'Geist' und 'Leben' in der Philosophie der Gegenwart," (Cassirer 1930) Cassirer—in the context of criticizing Max Scheler, a philosopher in the Lebensphilosophie movement— argues that 'Spirit' has the upper hand on Life, in that any articulation of Life must happen within the context of Spirit. Any 'form' we might ascribe to life involves predication and judgement, which are functions of the "logos." The implication of this could mean that any naturalist project is wrongheaded for similar reasons, since nature must always be articulable in spirit. In section IV (also see the introduction) I will argue how I can concede this point and still argue how nature is self-organizing. We will see there that it is a presupposition of objective thought that nature is to be conceived of as independently self-organizing. For the time being, my primary aim to draw from the study of "Geist and Life" is that spirit itself is 'alive' and thereby self-organizing.

self-organizing process of Life that makes these forms possible and ceasing to participate in this process.

Cassirer notices that if we sharply distinguish between Life and Form in this manner, we arrive at a paradox. How do we reconcile the dualism of Life—the unity of continuous creativity— and Form—the discrete 'products' of this creativity? Cassirer frames the paradox as perceived by the Lebensphilosophie movement:

> When they are taken as principles in the shaping of the world, there is a deep contradiction between the processes of life and form, taken in themselves. Form is limitation, hence separation and differentiation, the assertion of an independent content of its own as enduring in being, so that it is set off from the continual process of becoming (Cassirer 1996, 10)

The boundlessness of Life is constrained when Forms bound it. However, Forms require Life to come into being, while Life must create (and eliminate) Form to fully be Life. They mutually depend on each other to exist, yet they cannot coexist. For Henri Bergson, this is a contradiction for *thought,* which we can easily surpass when we leave thought behind to dwell in the pure intuition of the life process. For Cassirer, however, this is not an option, claiming that "[pure] intuition can never go behind the world of forms per se because it is itself nothing other than a way of giving form" (Cassirer 1996, 13). There is no way to directly intuit the world, since the only conception we have of the world is mediated by the forms of thought.

Cassirer's solution to the paradox is a complete reframing of the dichotomy. Life and Form are not radically self-contained entities. Instead, he says,

> we only meet up with two accents that we fix in the process of becoming. Becoming is in its essence neither mere life nor mere form; rather, it is the becoming of form—[coming-into-being], as Plato said (Cassirer 1996, 15)

We both construct Form and use the Forms we construct in a process of self-organizing our experience. What was formerly called Life and Form are merely these two 'accents' in a singular self-organizing becoming of form. These two accents are precisely the horizonal level and actual levels of bi-constitution: the horizonal emphasizes the normative framework that makes symbolic activity possible; the actual level, on the other hand, emphasizes the concrete, performative act of forming specific utterances, artworks, scientific claims, etc. that are particular expressions of these horizonal normative frameworks. Cassirer likens this 'accenting' to a proper understanding of the Kantian notions of understanding and sensibility.[71] To disentangle them and treat them as two separate things would be a "methodological abstraction."[72] Cassirer claims our faculties of understanding and sensibility are two dimensions of a living, formative process of making sense of the world. Neither have priority.

But how exactly do we reconstruct the concepts of Life and Form to avoid the paradox? To this end, Cassirer first suggests we follow Wilhem von Humboldt in making a distinction between energeia ('activity') and ergon ('work'). Life, in its boundless creativity, does not create forms to leave them be as mere work (ergon). Rather, Life is "continuously shaping itself" *through* these forms in a dynamic process (energeia) of the becoming of form:

---

[71] This perhaps points to the idea that Cassirer is moving beyond his 'intellectualist' understanding of experience that is exhibited, for instance, in (Matherne 2021, 40).

[72] "With Kant it seems again and again as though…that theoretical knowledge of nature is built up when the 'given' of sensation enters into the a priori forms of pure intuition and the pure understanding…If one want to see this as anything more than a methodological abstraction, if one believes that the elements which have here been separated in analysis must also be given as separable parts of the knowing process, then this conception has been progressively refuted in the course of the development of our theoretical knowledge since Kant...Here too we can recognize the means of thought and their theoretical significance only in theoretical thought's living function, only in its work and in its achievements" (Cassirer 1996, 21).

> What seems to be a real opposition when seen from the standpoint of Being becomes merged; it becomes a correlation and cooperation, when regarded from the standpoint of activity, of intelligent creativity (Cassirer 1996, 17)

Forms are not merely static products or 'forma formata.' They are creative, productive vehicles *through* which cultural works are formed: they are 'forma formans.' There seems to be three dimensions to the energeia of intelligent creativity here that characterizes the becoming of form: Activities are 1) limited, 2) creative acts that 3) are performed for the sake of self-organizing our experience. Recall that Lebensphilosophie only grants the first dimension to Form, whereas for them Life in its creativity and in its unifying function covers the second and third dimensions. Cassirer, however, unites all three as dimensions of a concrete, self-organizing activity, which he calls the 'becoming of form.' This tripartite dynamic mirrors the bi-constitutive process: limitation (horizonal framework), creative act (actual performance), and self-preservation (the feedback loop from actual success back to horizonal norms).

Cassirer argues that for an 'act' to exist, it must be limited in a specific way; there is no abstract act 'as such.' For instance, a physiological organ, like the heart, is limited to performing a particular function (e.g., pump blood), and it would not perform this function without being limited. Action and limitation are inseparable. Language offers another pertinent example: it is not a static repository of written words and grammatical conventions but is constituted by *acts* of speaking and writing according to specific conventions. These conventions, in turn, only become meaningful through their use in acts of speech and writing. In this way, the activity inherent in the 'becoming of form' is both a performance and a limitation. Cassirer's preferred metaphor is that of a river.

Activity is not like water poured into a pre-formed riverbed; rather it actively *creates* the riverbed itself, even if only gradually (Cassirer 1996, 18).

As both performances and limitations, our forms are constantly changing and adapting themselves. Language, for instance, is always changing its conventions. For this reason, Cassirer suggests a third dimension of the 'becoming of form.' Apart from being 1) limitations and 2) acts, our activities are 3) continuously adapting and self-organizing. We have seen how forms for Cassirer act like water in a riverbed, creating the very riverbed itself. Language also has this character. Our speech acts are not the mere 'application' of the rules of language, they produce an alteration to the very rules of language itself:

> the stronger its [the speech act's] individuality, the more it is preserved and the more it reproduces itself so that the impulse which it contains… influences it further and further so as to affect the flow of the whole, its intensity, its direction, its dynamics, and its rhythm (Cassirer 1996, 18)

It is in this way that newly constructed forms (as products of formative activity—forma formata) become the *new* organs of production (forma formans), that both limit and allow for further activity. Language is not some abstract formal system of rules but exists only insofar as it is used and adapts itself to our ever-evolving needs. If new linguistic constructions are 'strong,' they are 'preserved,' being eligible to be used again both as limitations and enabling conditions for further constructions of language (see Figure 3). This is the actual level of bi-constitution feeding back into the horizonal level: successful concrete acts (strong forms) revise the normative framework that governs future acts. The self-organization of culture is precisely this bi-constitutive loop.

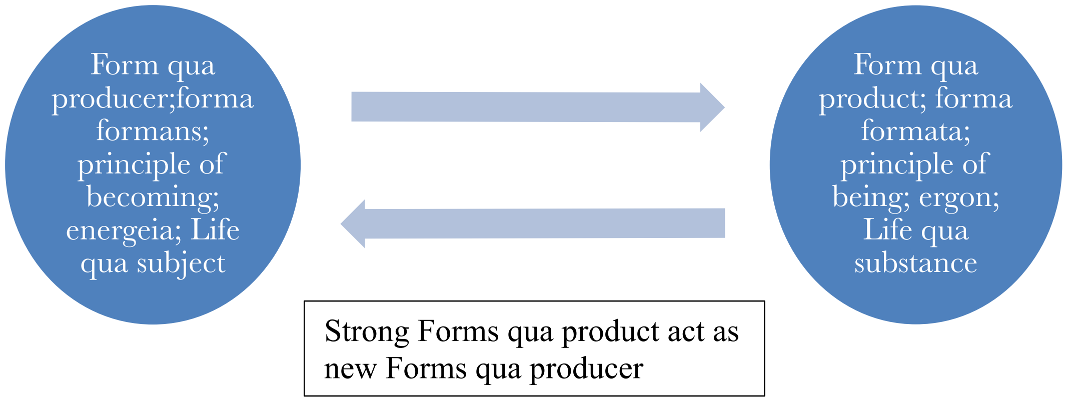

Figure 3: A schematization of the self-organizing 'becoming of form.' Forms qua producer (e.g., the activity of writing) produce Forms qua product (e.g., written works). The strong forms (e.g., Goethe's poetry) then change the way Form is produced (e.g., how we write). I include other terms that Cassirer uses interchangeably to characterize each side.

In this sense, I interpret the activity that is the 'becoming of form' to be a self-organizing process. What marks a successful and lasting 'form' in this process? Looking to Cassirer's earlier stance on organicism, I infer that it must 'preserve wholeness' (see section II). In this way, a 'strong' form is that which organizes successfully into a whole. For example, a mathematical form (such as the natural numbers) is strong and thereby preserved if it organizes and synthesizes an area of mathematical inquiry. A certain form of gas particles is strong and thereby preserved if said form helps organize the particles by, e.g., self-reinforcing feedback loops. These forms can fail to preserve this holistic structure if they cannot accommodate certain discoveries or disturbances, paving the way for new possible forms that do preserve this wholeness (Cassirer 1950, 213). The success or failure of a form is determined at the actual level—through concrete practices of inquiry, speech, or art—but this success retroactively validates or revises the horizonal framework that guided the act. This is precisely the structure of bi-constitution we saw in Chapter 2, where horizonal and actual bi-constitution are forever re-aligning themselves.

To drive this point home in the 'Geist and Life' piece, Cassirer says,

> Life… possesses and grasps itself in the imprint of form as the infinite possibility of formation, as the will to form and the power to form. Even life's limitation becomes its own act; what from outside seems to be its fate, its necessity, proves to be a witness to its freedom and self-formation (Cassirer 1996, 19)[73]

Here Cassirer describes Life as a "will to form," embedded in the process of 'becoming of form' in the three interconnected aspects I have described. First, it signifies 1) a limit and 2) a 'power' to form—much like an organ that performs a certain function but is restricted to that function. In this sense, to act is to become limited in a certain way. At the horizonal level, this limit is the normative framework that defines the purposive subject (the speaker, the scientist, the artist) and the purposive world (the domain of significance within which objects show up as meaningful). But Life as the becoming of form also 3) freely 'self-forms' or, to use Kant's term, 'self-organizes.' This formative process adapts to new circumstances, creating new ways of limiting itself based on the success of its formations, much like an organ might adapt to perform new functions through evolution. At the actual level, this self-formation is the concrete, performative act that succeeds or fails, and whose success feeds back into the horizonal framework, revising

[73] Cassirer sums up his comments in a similar way in the *Philosophy of Symbolic Forms Volume III*: "The forms in which life expresses [äußern] itself and by virtue of which it obtains its "objective" shape, signify for this very resistance how it designates its indispensable reverberation… The apparent counter force thus becomes the impulse of the movement as a whole; the direction toward the outwardness [Äußerlichkeit] not of things but of forms and symbols marks off the path by which pure subjectivity first finds itself." (Cassirer 2020b, 45).

the norms that will guide future acts. Thus, the 'becoming of form' is a self-organizing, adaptive activity—and it is self-organizing precisely because it is bi-constitutive.[74]

IV. Conclusion

To summarize, section II highlighted the importance of Kant's notion of purposiveness for Cassirer's philosophy of nature. I argued that—although Cassirer does not use these terms—we can interpret his view as suggesting that nature fundamentally consists of self-organizing wholes (forms) that shape and guide local mechanistic interactions. This self-organization of nature is accessed through the formal imperative, a practical presupposition that demands nature ought to conform to forms, and through the regulative judgment that guides us to discover material purposiveness in particular phenomena.

Section III has shown that thought, in all its cultural manifestations (including science), exhibits the same self-organizing dynamic. Drawing on Cassirer's "'Geist' and 'Life,'" we saw that thought is a 'becoming of form' (das Werden der Form): an activity that simultaneously limits itself (performance), creates new forms (production), and adaptively preserves wholeness (self-preservation). This tripartite dynamic of energeia—what Cassirer calls the "will to form and the power to form" (Cassirer 1996, 19)—characterizes language, science, art, and all symbolic forms. It is a development of Kant's

[74] Cassirer also speaks of a 'metamorphosis' of spirit in the same way in *Ziele und Wege* where individuals who "take hold of existing forms only to treat them in turn as material" where individuals "reshape what has been shaped" (1937, 173 [my translation]). This process is undoubtedly the process of self-organizing *energeia* as I have laid it out here.

conception of bi-constitution. Thought is not a static repository of concepts but a living, self-organizing activity that continuously reshapes itself through its own successful productions.

In "'Geist' and 'Life'" Cassirer makes it clear that these two theses—the self-organization of nature and the self-organization of thought—are connected. He claims that the "forces of nature" are on par with "all truly intellectual energies" insofar as "they all exist by virtue of being activated and by giving shape to themselves in this activation" (Cassirer 1996, 15). Nature, like culture, is activity in the triple sense of energeia just described. Yet this claim raises a fundamental question. Cassirer is no realist. He does not mean to metaphysically ground the self-organization of thought within the self-organization of a nature independent of our thought. But does rejecting realism lead inevitably to subjective idealism—the view that we project our own mental states onto nature, reducing nature's self-organization to a mere reflection of our own cognitive dynamics? If accepted uncritically, this reading would render Cassirer vulnerable to standard relativist or psychologistic objections.

However, as we have already seen in Section II, Cassirer's formal imperative provides a way out of this false dilemma. The self-organization of thought and the self-organization of nature are not opposed. Rather, they are two moments of a single bi-constitutive process. Material forms guide practices which simultaneously constitute a normative identity for us as certain agents capable of a type of inquiry while also constituting nature as a purposive domain for inquiry. The actual acts that these practices govern (discoveries, experiments, etc.) can succeed or fail, making our own identities and the structure of nature explicit. Thought does not project its dynamics onto a passive nature, nor does it simply copy a ready-made natural order. Instead, thought and nature co-arise through the normative demands of reason and the fallible, embodied practices of scientific inquiry. It is only *because* thought self-organizes in this way (i.e., it is not led by its desires or circumstances) that nature can be discovered to be self-organizing. The two depend on each other. The question of whether nature is "really" self-organizing independent of our cognition is, from this perspective, a metaphysical pseudo-problem. What matters is that we are compelled to judge nature as self-organizing by the very structure of rational inquiry—and that this compulsion is continually confirmed by the disinterested pleasure we feel when our practices succeed.

Bibliography

Bayer, T. I. (2001). Cassirer's Metaphysics of Symbolic Forms. Yale University Press.

Bishop, R. C., & Frigg, R. (2016). From Order to Chaos and Back Again. In N. Cartwright & K. Ward (Eds.), Rethinking Order after the Laws of Nature (55–74). Bloomsbury Academic.

Cassirer, E. (1918). Kants Leben und Lehre. In B. Recki (Ed.), *Gesammelte Werke: Hamburger Ausgabe* (Vol. 8). Meiner.

Cassirer, E. (1937). Ziele und Wege der Wirklichkeitserkenntnis. In K. C. Köhnke & J. M. Krois (Eds.), *Gesammelte Werke: Hamburger Ausgabe* (Vol. 2). Meiner.

Cassirer, E. (1939). Naturalistische und humanistische Begründung der Kulturphilosophie. In B. Recki (Ed.), *Gesammelte Werke: Hamburger Ausgabe* (Vol. 22, 140–166). Meiner.

Cassirer, E. (1944). *An Essay on Man: An Introduction to a Philosophy of Human Culture.* Yale University Press.

Cassirer, E. (1950). *The Problem of Knowledge: Philosophy, Science, and History since Hegel* (W. H. Woglom & C. W. Hendel, Trans.). Yale University Press.

Cassirer, E. (1996). *The Philosophy of Symbolic Forms, Volume Four: The Metaphysics of Symbolic Forms* (J. M. Krois & D. P. Verene, Eds. & Trans.). Yale University Press.

Cassirer, E. (2000). The Problem of Form and the Problem of Causality. In S. Lofts (Trans.), *The Logic of the Cultural Sciences* (87–102). Yale University Press.

Cassirer, E. (2021a). *The Philosophy of Symbolic Forms, Volume One: Language* (S. Lofts, Trans.). Routledge.

Cassirer, E. (2021b). *The Philosophy of Symbolic Forms, Volume Three: The Phenomenology of Knowledge* (S. Lofts, Trans.). Routledge.

Deacon, T. (2011). *Incomplete Nature: How Mind Emerged from Matter*. Norton.

Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, 4(4), 429–452. https://doi.org/10.1007/s11097-005-9002-y

Endres, T. (2021). Genealogische Kulturanthropologie – Erinnerung an Ernst Cassirer (1874–1945). In G. Hartung & M. Herrgen (Eds.), *Interdisziplinäre Anthropologie* (289–315). Springer.

Ferrari, M. (2012). Between Cassirer and Kuhn. Some Remarks on Friedman's Relativized A Priori. *Studies in History and Philosophy of Science*, 42(1), 18–26.

Ferrari, M. (2015). Cassirer and the Philosophy of Science. In N. de Warren & A. Staiti (Eds.), *New Approaches to Neo-Kantianism* (261–284). Cambridge University Press.

Friedman, M. (2000). *A Parting of the Ways*. Open Court Press.

García-Valdecasas, M. (2022). On the naturalisation of teleology: self-organisation, autopoiesis and teleodynamics. *Adaptive Behavior*, 30(2), 103–117. https://doi.org/10.1177/1059712320981224

Haken, H. (2006). *Information and Self-Organization: A Macroscopic Approach to Complex Systems.* Springer Science & Business Media.

Hartung, G. (2018). *Beyond the Babylonian Trauma*. De Gruyter.

Heis, J. (2014). Realism, Functions, and the a Priori: Ernst Cassirer's Philosophy of Science. *Studies in History and Philosophy of Science*, 48, 10–19.

Holling, C. S. (2010). Resilience and Stability of Ecological Systems. In L. Gunderson, C. Allen, & C. S. Holling (Eds.), *Foundations of Ecological Resilience* (19–50). Island Press.

Hooker, C. (2011). Introduction to Philosophy of Complex Systems. In C. Hooker (Ed.), *Philosophy of Complex Systems* (Vol. 10, 3–90). Elsevier.

Howe, C. S. (1961). Translator's Foreword. In E. Cassirer, *The Logic of the Humanities* (C. S. Howe, Trans, vii–xviii). Yale University Press.

Juarrero, A. (2023). *Context Changes Everything: How Constraints Create Coherence.* MIT Press.

Kant, I. (1993). *Opus postumum* (E. Förster & M. Rosen, Eds. & Trans.). Cambridge University Press.

Kant, I. (1998). *Critique of Pure Reason* (P. Guyer & A. W. Wood, Trans.). Cambridge University Press.

Kant, I. (2000). *Critique of the Power of Judgment* (P. Guyer & E. Matthews, Trans.). Cambridge University Press.

Kauffman, S. (1993). *The Origins of Order: Self-Organization and Selection in Evolution.* Oxford University Press.

Keller, P. (2018). Kant's Threefold Autonomy after the Groundwork. In S. Bacin & I. Sensen (Eds.), *The Emergence of Autonomy in Kant's Moral Philosophy* (196–212). Cambridge University Press.

Keller, P. (2021). Dynamic Structuralism – Cassirer and Method in Culture. In L. Filieri & A. Pollok (Eds.), *The Method of Culture: Ernst Cassirer's Philosophy of Symbolic Forms* (41–66). Edizioni ETS.

Kreis, G. (2010). *Cassirer und die Formen des Geistes.* Suhrkamp.

Kreis, G. (2021). Cassirer's Philosophy of Mind: From Consciousness to 'Objective Spirit.' In S. Truwant (Ed.), *Interpreting Cassirer*. Cambridge University Press.

Luft, S. (2019). Mind als Geist in der Welt der Kultur: Kulturphilosophie, 'Naturalische' Transzendentalphilosophie und die Frage nach dem Raum der Kultur. In T. Breyer & S. Niklas (Eds.), *Ernst Cassirer in systematische Beziehungen: Zur kritisch-kommunikativen Bedeutung seiner Kulturphilosophi*e (129–150). De Gruyter.

Matherne, S. (2021). *Cassirer*. Routledge.

Merleau-Ponty, M. (2013). *Phenomenology of Perception* (D. Landes, Trans.). Routledge.

Mitchell, M. (2009). *Complexity: A Guided Tour*. Oxford University Press.

Möckel, C. (2005). *Das Urphänomen des Lebens: Ernst Cassirers Lebensbegriff*. Felix Meiner Verlag.

Natorp, P. (2015). Kant and the Marburg School (F. Bottenberg, Trans.). In S. Luft (Ed.), *The Neo-Kantian Reader* (180–195). Routledge.

Packard, N. (1988). Adaptation Toward the Edge of Chaos. *Dynamic Patterns in Complex Systems*, 212, 293–301.

Prigogine, I., & Stengers, I. (1984). *Order Out of Chaos*. Bantam New Age Books.

Richardson, A. (1998). *Carnap's Construction of the World: The Aufbau and the Emergence of Logical Empiricism*. Cambridge University Press.

Ryckman, T. (2005). *The Reign of Relativity: Philosophy in Physics 1915-1925*. Oxford University Press.

Schubbach, A. (2022). *The Genesis of the Symbolic* (D. J. Hobbs, Trans.). De Gruyter.

Skidelsky, E. (2008). *Ernst Cassirer: The Last Philosopher of Culture*. Princeton University Press.

Verene, D. P. (2011). *The Origins of the Philosophy of Symbolic Forms: Kant, Hegel and Cassirer.* Northwestern University Press.

Waldrop, M. M. (1992). *Complexity: The Emerging Science at the Edge of Order and Chaos.* Simon & Schuster.

Chapter 5: A Critique of Complex Systems Science

I. Introduction

We have seen so far that Kant contributes to western philosophy a way to align the self-organization of nature with the way we self-organize as autonomous agents. Kant thinks that 'Ideas' guide practices that, when we participate in them, provide us with a normative identity and constitutes the world as purposive to our ends. The Ideas guide us toward complete moral integrity in the practical sphere, and complete cognition in the theoretical sphere. Precisely because they orient thinking and doing, these forms can never be fully articulated when taken as objects of theoretical scrutiny. This gives rise to an unending critique of the basic principles that enable our practices—a critique that at once reveals both who we are and the structure of nature.

Yet we are not only guided by Ideas—we are also remarkably adept at judging things in nature as if they, too, were guided by them. This act of judgment deploys what Kant called a principle of material purposiveness: the hypothesis that natural entities be construed as self-organizing agents.[75] To perceive a cell or a forest this way is to see it as governed by an immanent 'Idea' of wholeness—a principle whereby parts and whole mutually constrain and constitute one another, whereby the agent in turn organizes its own environment to sustain itself. This is the thesis of 'bi-constitution' applied to self-

[75] I will use the terms purposive being, (self-)organized being, and an agent all interchangeably. See chapter 3 for a unity of these terms

organizing, natural agents. I unpacked Kant's logic of self-organization in chapter 3 according to four interdependent moments:

1) First, the organizing being's form—its normative identity as described by the Idea—orients its parts (or subsystems) into a functional, self-sustaining whole. This is the horizonal constitution of the purposive subject, establishing what the organized being is and what counts as its proper functioning.
2) Second, the organized being's parts exist in reciprocal causality: each is both cause and effect of the others, and together they produce and maintain the whole. This is the actual constitution of the organized being, the ongoing activity through which the organized being sustains itself.[76]
3) To persist, the organized being requires selective input from its environment, which its form (Idea) thereby renders purposive, constituting a niche of potential significance. This environment is itself composed of other self-organizing systems, forming an ecology of reciprocal dependence. This is the horizonal constitution of the purposive environment, the domain of significance in which nutrients, threats, and meanings show up for the organized being's self-maintenance.
4) Successful acquisition of environmental input leads to persistence of the whole (whether through growth, reproduction, etc.), while failure pressures the

[76] This may include the actual environments it constructs (e.g., nests, burrows, hives) that, for our intents and purposes, count not as 'external' to the being, but its constitutive parts.

> system toward collapse or adaptive reorganization. This is the coherency test between actual and horizonal constitution: the actual acts of the, e.g., organism—its feeding, growth, repair—continually realize and articulate the horizonal field, either reinforcing the norms of its form of organization (e.g., Idea) or, in cases of failure, forcing a revision of how it attains its self-preservation.

To attribute such an Idea to an entity is to ascribe to it a unified source of activity that acts upon both itself and its world in order to maintain its purposive existence.

The new science of complex systems theory presents a unique challenge to Kant's principle of judgment. While Kant intended his principle to be merely a regulative framework for scientific inquiry, the remarkable success of complex systems theory—in modeling phenomena from tornadoes to neuronal networks to economic systems—has led prominent philosophers and scientists to argue that attributing self-organization is not merely a regulative feature of judgment. Rather, these philosophers and scientists argue we now *know* that the world is *constituted* by self-organizing agents. Complex systems science has led to a 'naturalized' redescription of teleological statements in the formal terms of attractor states and constraints. This has prompted 'realists' such as Alicia Juarrero (1999; 2023), Evan Thompson (2007), Richard Campbell (2015), Manuel De Landa (2002), Maturana & Varela (1980) and Stuart Kauffman (1993) to all claim that final causality—at least translated in contemporary terms of emergence, downward causation, and constraints— are real, constitutive features of nature, not just frameworks guiding inquiry as Kant would have it.

However, as argued in Chapter 1, if a realist view of self-organization is correct, we must locate our own cognitive practices within the very self-organizing world they seek to understand. This creates a tension: our cognition is governed by normative rules that determine successful reasoning. To theoretically describe these rules from a neutral, external standpoint would require us to step outside the normative framework we are attempting to analyze. This is an impossible feat if we are truly embedded, self-organizing agents. Therefore, if self-organization is constitutive of reality, we must abandon the possibility of making omniscient, purely descriptive claims about the cosmos. We would thereby have to give up the claim that reality is primordially constituted by self-organizing agents.

Yet if this is the case, realism about self-organizing agents must be mistaken. We have no means to leap beyond our constraints to make sweeping metaphysical conclusions about how nature is constituted. Nevertheless, even if realism is false, the realist is right to acknowledge the intuition generated by the success of complex systems theory: why does it feel like the success of these theories assert that the cosmos really is a self-organizing agent, comprised of other self-organizing agents?

While complex systems theory has led certain philosophers to be realists about final causes (albeit in a naturalized form), other naturalists have viewed the complexity revolution as an opportunity to *eliminate* final causes in scientific explanations.[77] For these 'deflationists,' final causes are not real, constitutive features of certain systems, but heuristic fictions that can be replaced by talk of efficient causation. Robert Cummins (1975), for instance, redefines "the function of X is Y" to mean something like "X has a capacity to do Y, and this capacity is singled out for explanation because it contributes in a specific way to a more complex capacity of the containing system S." All talk of goals and purposes are just shorthand descriptions of complex efficient causation.

Deflationists reduce teleology to a heuristic overlay on a purely efficient-causal substrate. Cummins himself explains that talk of functions is dependent on an 'analytical context' (1975, 180). This context is the heuristic explanatory perspective from which we analyze a system's capacities in terms of the organized contributions of its components. Deflationists like Cummins correctly acknowledge that our own cognition operates by using certain heuristics to guide scientific inquiry, but they fail to explain why the heuristic to describe things teleologically is both unique and indispensable for certain scientific practices. The deflationist treats all heuristics as interchangeable tools adopted for pragmatic convenience, yet the history and practice of science suggests that the teleological heuristic is not one option among many but has seemed to serve as a necessary condition needed to investigate organized systems at all. One cannot simply choose to set it aside and still make sense of organisms, ecosystems, or economies. The deflationist, however, has no resources to explain this necessity. To do so would require a

[77] Causal role theorists are not the only ones to attempt this elimination. 'Selected effect' theories of function, such as Karen Neander's (1991) and Ruth Millikin's (1989), expanding on the foundational work of Larry Wright (1973) aim to define a trait's function to be what it was selected for in evolutionary history.

transcendental critique of the nature of human reason. But such a critique would itself take a fundamentally teleological perspective on reason, treating reason as oriented toward the systematic unity of knowledge, which goes beyond the deflationist's commitment to purely efficient-causal explanations. The deflationist account, therefore, remains incomplete, and intrinsically so.

In response to this difficulty, the realist can consistently say that teleological heuristic is so indispensable because the world really is composed of final causes. If teleology is a real feature of nature, then of course the heuristic that tracks it would prove indispensable—its success is simply what we would expect. Yet we saw that the realist is inconsistent on this point, since they do not recognize that we as knowers are also teleological, self-organizing, constrained beings, who are limited from making such sweeping claims about reality. The realist thus explains too much: they claim to know what nature is like in itself, a claim that exceeds the limits of our cognition. The deflationist, on the other hand, while choosing not to explain the centrality of the teleological heuristic, explains too little.

This chapter aims to give a critique of complex systems theory. I mean to provide a critique in the Kantian sense: an investigation into the conditions of possibility that make a certain line of inquiry intelligible. In contrast to both deflationists and realists, I will show how complex systems science articulates Kant's regulative principle of material purposiveness, as outlined in the four 'moments' above. For Kant, we must hypothesize that nature operates as if it were composed of self-organizing agents interacting with other agents with the 'aim' of self-persistence. This is a hypothesis in the Kantian sense. It is not a claim proposing how nature is in itself that could be empirically confirmed or falsified, but a necessary presupposition that opens a domain of inquiry. Because our discursive intellect proceeds from parts to wholes, Kant thinks we cannot grasp how a whole could be the cause of its parts without presupposing that it is organized as if by an end. The principle of material purposiveness is thus not a discovery about nature, but a condition of the possibility of investigating organized beings at all. Yet it is more than a mere heuristic. It is an essential, objective way of viewing nature that—because the hypothesis is so successful—we can feel its truth.

To carry out this critique, I first develop in section II a Kantian framework for understanding whether complex systems science is a genuine 'science' and, if so, how it gets validated. I then examine in section III the core mathematical tools of complex systems science—graph theory, network theory, dynamical systems theory, criticality, and panarchy—to demonstrate how each presupposes and articulates the four-part schema of self-organization outlined above. Finally, in section IV I return to the realist and deflationist positions to show how this Kantian framework overcomes the incompleteness of both.

II. The Kantian Idea of Purposiveness

To understand what kind of science complex systems theory is and how its truth can be validated, we must first clarify Kant's distinction between two senses of science. In the *Metaphysical Foundations of Natural Science* Kant defines "proper science" as a doctrine that contains a pure, a priori, mathematizable part lying at the foundation of its empirical part. On this definition, only physics qualifies as a proper science. Chemistry, on the other hand, is merely a "systematic art" or "experimental doctrine" because its principles are empirical and admit of no a priori presentation in intuition (4:471). In the *Critique of Judgment,* biology seems to fare no better, Kant claiming that we must "absolutely deny" that a "Newton who could make comprehensible even the generation of a blade of grass according to natural laws" (5: 400). Yet in the *Critique of Pure Reason*, Kant also recognizes a broader sense of "science:" any systematic body of knowledge organized under a unifying Idea, where the parts are combined in accordance with a principle into a whole (A832/B860). The practitioners of a science in this sense may not initially grasp its Idea. Kant claims this Idea lays dormant in reason, "like a seed, all of whose parts still lie very involuted and are hardly recognizable" (A834/B862). As the discipline matures, its coordinating principles and aims become clarified, and the Idea that implicitly guided its development is gradually brought to light.

Complex systems science, I argue, occupies an interesting position relative to these distinctions. It does not meet Kant's criteria for proper science: its mathematical tools—graph theory, network theory, and dynamical systems theory—are formal languages and modeling methods, not a priori constitutive principles of nature. Its specific equations and models are drawn from experience and tailored to particular systems, and there is no set of universal laws that all self-organizing systems must obey. The mathematical structures it employs are formal frameworks for representing how a system might behave, not laws dictating how it must behave. Unlike Newtonian physics, which provides constitutive principles that govern all material bodies, complex systems science offers regulative heuristics and formalisms that guide inquiry without legislating nature. In this respect, complex systems science stands with biology and chemistry in Kant's taxonomy: it is a systematic art organized under guiding principles, but not a proper science in the strict sense.

However, complex systems science does seem to fit the criteria that characterize the *broader* sense of science Kant describes in the first Critique. It is a systematic field of inquiry organized under a unifying Idea: the Idea of self-organizing wholes[78], guided by the principle of material purposiveness. Its practitioners—like the first practitioners of any science—initially grasped this Idea only vaguely. The Belousov–Zhabotinsky reaction and Kauffman's NK models, discussed in chapter 1, exemplify this early, inchoate stage. Neither began as applications of a pre-existing complexity theory, but were anomalies that resisted explanation by prevailing paradigms. What united them was a common implicit intuition, what Kant calls a 'monogram', that order could arise spontaneously from the interactions of many components without an external template or central controller. This assumption—what Kant calls the involuted 'seed'—remained indistinct in the early work. Only with the development of graph theory, network theory, and

[78] As I have discussed in Chapter 3, this is a contentious claim. It is a fairly standard reading that what I call the 'Idea' of purposiveness is merely regulative for judgment—a principle of reflective judgment, not an Idea of reason. Whether the principle of purposiveness qualifies as an Idea of reason in the strict Kantian sense is not settled, and I do not need to settle it here. What matters for my argument is that it functions as if it were an Idea: it has the same problematic, hypothetical structure, guiding inquiry without being fully determined. It operates as a regulative presupposition that makes the investigation of self-organizing beings possible, and it does so in a way that is structurally analogous to how the Ideas of reason guide theoretical and practical philosophy. Complex systems science, I argue, takes this problematic, hypothetical principle and articulates it formally, giving mathematical expression to the structure that has always guided the sciences of organized beings. Whether this makes complex systems science a science of the Idea itself is a question I leave open.

dynamical systems theory as explicit mathematical frameworks did the Idea guiding these inquiries begin to be articulated into a comprehensive schema. The mathematical tools examined in the following sections are the outcomes of this trajectory that searches for this unified Idea.

Although many sciences depend on the principle of material purposiveness (namely the biological sciences), complex systems science is distinct in that it is a discipline that thematizes the regulative principle of purposiveness itself, making it an explicit object of investigation. Just as formal logic thematizes the principles of reasoning that all sciences use implicitly, complex systems science reflects on the condition that all sciences rely on when they investigate organized wholes. It does not thereby replace or ground those sciences. Rather, it shows that the regulative principle of material purposiveness can be formalized mathematically and applied across a wide range of domains that study organized wholes, from physics to economics.

This brings us to the question of validation. If complex systems science is a regulative enterprise rather than a constitutive one, how can we assess its truth? Is it merely a subjective heuristic? As we have seen in chapters 2 and 3, Kant provides an answer in his account of disinterested pleasure. In the *Critique of the Power of Judgment,* he writes that "the attainment of every aim is combined with the feeling of pleasure" (5:187). For certain aims, this pleasure is not personal but is determined by a "ground that is a priori and valid for everyone," arising "merely through the relation of the object to the faculty of cognition, without the concept of purposiveness... having the least regard to the faculty of desire" (5:187). This pleasure occurs whenever reflective judgment succeeds in discovering a principle that unifies heterogeneous laws of nature, or more generally when reason's regulative Ideas successfully guide empirical inquiry (5:187–188). In such moments, we do not merely postulate that nature is purposive. The success of our theoretical endeavors guided by these a priori principles provides a *felt* confirmation that the presupposition guiding our inquiry—that nature is systematically unified—has been answered by nature itself. This pleasure is not reducible to wish-fulfillment or desire, nor is it contingent from person to person; it is grounded in cognition itself, which is why it can claim universal validity.

Kant originally described this pleasure in the context of *formal* purposiveness—the systematic unity of nature for our cognition, a principle he argued is necessary for all empirical science. When we discover that the same formalisms—networks, attractors, criticality—govern phenomena as diverse as chemical oscillations, gene regulation, ant colonies, and economic markets, we experience the same kind of disinterested pleasure that Kant associated with discovering systematic unity in nature. This success does not prove that nature is constitutively self-organizing, but it does provide a powerful corroboration from the success of scientific practice. The pleasure we feel is the felt confirmation that our regulative presupposition—that nature can be investigated as if it were composed of self-organizing wholes—was appropriate to the way nature is structured. This lends the principle epistemic weight: it indicates that the framework we use to see nature in this way is not arbitrary but is in harmony with our cognitive faculties and with nature itself.

The four-part schema of self-organization outlined earlier provides the template for what complex systems theory both presupposes and *articulates* further. The organized being's form (horizonal constitution), the reciprocal causality of its parts (actual constitution), its selective relation to an environment (horizonal constitution of the world), and the coherency test between actual and horizonal constitution through success or failure, are the logical moments of purposive organization. In articulating these principles into a formal framework, it *develops* the four moments of the regulative principle, giving it a mathematical and conceptual precision that Kant could not have provided.

In sum, complex systems science is not a science in the narrow Kantian sense, but a science in his broader sense. It is a systematic field, organized under the Idea guiding the study of self-organizing wholes. Complex systems science thematizes the regulative principle of material purposiveness, making explicit what sciences tacitly presuppose when they investigate organized wholes. And its spectacular success produces the disinterested pleasure that, for Kant, signals that this regulative principle is appropriate to nature and more than a mere subjective heuristic. With this framework in place, we can now examine the core mathematical tools of complex systems science to see how they both presuppose and articulate the four-part schema of self-organization.

III. Graph, Network, and Dynamical Systems Theory

Unlike physics, biology, or psychology, complex systems theory does not have a particular content, such as matter, life, or the human mind. It is a set of content-neutral tools that enable scientists to build sophisticated models for understanding dynamic, emergent, and self-organizing phenomena across diverse domains, including matter, life and the human mind. I will begin by unpacking the core theoretical framework, which includes Graph Theory, Network Theory, and Dynamical Systems Theory, before moving on to more narrow theories such as criticality and panarchy. After describing the theories, I will articulate how each in their application both presupposes and articulates the four moments of the principle of material purposiveness.

A. Graph Theory, Network Theory, and Dynamical Systems Theory

The primary activity of complex systems scientists is modeling groups of interacting entities called 'systems.'[79] This endeavor forces the modeler to adopt a specific perspective. The modeler must first identify the system's components and the structure of their interactions, mapping the potential pathways of constraint and influence that one component has over the other. Second, the modeler must then assign local rules of motion, dictating how each component's state changes based on the states of its neighbors. These rules almost are always nonlinear, meaning a single cause does not produce a proportionate (linear) effect. The first act requires the tools of the subset of graph theory known as network theory, while the second requires dynamical systems theory. These do not exhaust the mathematical theories that are used to model systems and their dynamics but are the most essential modeling tools for complex systems

[79] For a more technical and comprehensive summary, see Boccalette et al. (2006).

scientists.[80] Before describing how the application of graph theory, network theory, and dynamical systems theory presupposes the Idea of purposiveness, I will first articulate each theory's unique contributions towards a science of self-organizing agency.

Graph theory's distinct formal contribution to complex systems theory is to introduce a formal language to describe distinct objects (known as nodes) and their pairwise relations (known as edges). This provides the tools to study the abstract structure of a given system, answering questions pertaining to, e.g., the shortest path between nodes, how many edges can be removed before fracturing the connectivity of a system, etc.

Let me break this down formally[81]. A graph is an ordered pair $G = (V, E)$ where:

V is a finite set of vertices or nodes. $V = \{v_1, v_{2,} \dots, v_n\}$ (1)

E is a set of edges, which are 2-element subsets of V:

$$E \subseteq \left\{\{v_i, v_j\} \middle| \, v_i, v_j \in V\right\}. \tag{2}$$

Relations can be described formally as edges between v_i and v_j. For example, $\{v_1, v_2\}$ could be the relation between me (v_1) and my father (v_2). The sets of vertices

[80] Statistical mechanics/probability theory (especially as used in network theory), game theory, and information theory are examples of such additional mathematical tools that are in the complex systems scientist's toolbox, but are more specialized, used to investigate different aspects or types of complex systems. Due to space, I will leave these tools unanalyzed. Yet my conclusion for them would be the same: their application requires a minimum attribution of agency to the systems under study.
[81] The graph theory equations here are taken from Harris et. al (2008)

alone, however, carries no information about the nature or strength of the relations: they are content free.

Graph theory builds on the primitive relation of edge to create more complex structural concepts. For example:

A path of length k from vertex u to vertex v is a sequence of vertices

$u = w_0, w_1, w_2, \dots, w_k = v$

such that every i from 0 to $k - 1$, the pair $\{w_i, w_{i+1}\}$ is an edge in E. (3)

This allows us to chain relations together, defining what it means to get from one vertex to another using the available connections. Two vertices are defined to be 'connected' if there exists a path from u to v. We can then proceed to define a distance $d(u, v)$ between two connected vertices u and v is to be the length of the shortest path connecting them. The distance between two points is always an integer, not a continuous measure, creating a granular geometry on the graph. We can furthermore define neighborhoods $N(v)$ to be sets of vertices adjacent to v. We can use this definition to define 'continuity' across the graph, where a graph is continuous if it does not 'jump' across neighborhoods, changing value only by stepping from a vertex to an immediate neighbor.

Finally, graph theory allows us to define different kinds of 'qualities' of connectivity by defining cycles and trees. Cycles are paths of at least three where the path begins where it ends ($w_0 = w_k$) and all other vertices are distinct. Connected graph with no cycles are called trees. Since there are exactly one path between any two vertices, these trees are 'brittle,' offering minimal connectivity. A graph with cycles, on the other hand, provide 'robust' connectivity, meaning that there are multiple paths between vertices.

Even if some edges are removed, the graph can remain connected (and sometimes robustly so).

We can give algebraic representations of graphs by creating matrices A (called adjacency matrices) that contain the complete structural information of the graph. That is, it displays the list of all edges that exist between nodes by equating rows and columns in the matrix to nodes in the graph. Entries that have 1 indicate an edge exists between two nodes, where 0 indicates a lack of such edge. Formally,

$$A_{ij} = \begin{cases} 1 \; if \; \{v_i, v_j\} \in E \\ 0 \; otherwise \end{cases} \tag{4}$$

While I will forgo the mathematical details, it is crucial to note that the power of these matrices arises from their unique properties under specific transformations. For instance, multiplying an adjacency matric A by itself to get A^2 counts the distinct walks of length 2 from v_i to v_j. Finally we can extract structural fingerprints called eigenvectors and eigenvalues from these matrices. These values can be useful for summing up large macroscopic properties of the graph. For instance, the largest eigenvalue gives a measure of the graph's connectivity, where a larger value suggests a more densely connected graph.

In modeling complex systems, graph theory's role is to supply a means to model the directed paths, feedback loops and connectivity of a given system. For example, take a neuronal network. In modelling this system, vertices would represent neurons, while edges would represent synapses or connections between neurons. We can use graph theory to model paths from any two random neurons for signal propagation, while also being able to represent cycles or feedback loops between the neurons. It can also show whether the neural network is weakly or strongly connected.

When we apply graph theory to model a real-world system, we can also utilize the tools of network theory to further analyze and describe it. To use the words of Albert-László Barabási, "what distinguishes network science from graph theory is its empirical nature, i.e., its focus on data, function, and utility" (2016, p. 5). Network theory is best described as a methodology, starting from empirically collecting data, constructing a network where the data is represented as nodes and edges (using graph theory), and using statistical measures to then characterize the network's topology. We can ask questions about the character of the structure and the roles of its components, instead of merely describing such network via patterned connections. We hence might refer to network theory as 'applied graph theory + statistical analysis.'

The statistical tools at the core of network theory characterize the typical or aggregate nature of the structure. In graph theory, a 'degree' is the number of edges incident with a given vertex.[82] Network theory analyzes degrees further, using a 'degree distribution' $P(k)$ which gives the probability that a randomly chosen node has k connections (a degree of k). This has the potential to classify networks into groups that graph theory does not. For example, a 'random network' (Erdős–Rényi) has a bell curve $P(k)$, whereas a scale-free network (where few nodes have a large number of connections, where most nodes have very few connections) has a power-law $P(k)$. These statistical descriptions allow us to study the overall connectivity pattern within a given system.

[82] The definitions presented on network theory come from Van Steen (2010)

Whereas graph theory can tell you if a node is connected, network theory can quantify a node's importance within the global structure. The means to do this are called 'centrality measures.'[83] One such measure called 'betweenness centrality,' for example, which describes the fraction of all shortest paths between any two vertices in the network that pass through a given vertex u. Formally,

$$C_B(u) = \sum_{x \neq y} \frac{S(x,u,y)}{S(x,y)} \tag{5}$$

Where $S\ (x, y)$ is the total number of shortest paths from vertices x, y and $S(x, u, y)$ is the set of paths that pass through vertex u. Certain nodes can have low degree but high betweenness centrality if it acts as a bridge between large communities. This gives a loose functional interpretation onto the bare structure given by Graph Theory, being able to identify key 'gateway' nodes that are more essential to the operation of the system than non-gateway nodes.

Network theory can also detect distinct 'communities' known as modules. Groups of nodes could display a dense web of edges connecting them, looking like a tightly knit cluster. These clusters might be like islands unto themselves, only vaguely connected to other clusters. If they are, these groups are known as 'modules' which are subsystems within a larger network. The quality of partitions between densely connected groups are measured by modularity, which gives a quantitative value to the 'partition-ality' of a given system (Newman, 2006).

---

[83] Concepts of centrality emerged from network theory to explain social phenomena (Freeman 1977), but has since been generalized by Graph theory. The equation here is from graph theory (Van Steen 2010, 152).

To return to our neuronal network example, network theory expands on the findings of graph theory and asks how the network displays large scale statistical patterns. The application of these tools would go beyond modeling the degree of individual nodes (as applied graph theory would) to analyzing the degree distribution across the entire network. Network theory can then identify highly connected neural 'hubs' and characterize the overall network topology of the entire neuronal system. It can also identify neurons or communities of neurons that are more essential than other neurons by calculating centrality metrics. Finally, it has the potential to model modules and communities of neurons that are more deeply connected than others, identifying important subsystems within the neuronal network.

While Graph theory provides tools from 'pure' mathematics and network theory supplements those tools with additional statistical analyses, each remain agnostic as to the specific rules that dictate the way nodes interact and change. Yet in physical systems, these rules are essential to understanding how this system evolves over time, as well as the overall 'stability' of this system as it evolves. This is where dynamical systems theory comes into play. Dynamical systems theory does this first by assigning 'states' to each node in the graph. Attached to each node i is a 'state space' which is a manifold M_i representing all potential states of that node. The total configuration space is $\prod_i M_i$, otherwise known as the total space. By switching to states, the fundamental unit of study is not the graph G but the state vectors $x(t)$ that belong to the manifold.

Given the states of the neighbors of any one node, dynamical systems theory allows us to predict how the state of that node might change. In other words, we specify how the state vector $x(t)$ evolves over time. This is typically done (assuming we take time to be continuous) with a system of differential equations. The general form for a deterministic system in continuous time evolves by the equations:

$$\frac{dx_i}{dt} = F_i(x_i, \{x_j : j \in N(i)\}) \tag{6}$$

Where F_i is a function defining the laws of the system and $N(i)$ is the set of neighbors in node i. We can separate the dynamics of how the node would behave in isolation with a coupling function, governing how interactions from neighbors influences the node:

$$\frac{dx_i}{dt} = f(x_i) + \sum_{j \in N(i)} A_{ij} g(x_i, x_j) \tag{7}$$

Where f is the self-dynamics, the sum of g is the coupling from neighbors, and A is the adjacency matrix (Newman 2010, 687).

Dynamical systems theory provides the tools to identify dynamic, long-term patterns in the fate of the system. The 'flow' or dynamic trajectory of the system through state space can be limited by attractors—stable patterns which the system settles into. Basins of attraction, on the other hand, limit the dynamics further by defining a domain of influence of an attractor, defined formally as the set of all initial states x_0 that evolve to a given attractor. Finally, measures of 'stability' provide the tools to determine whether a pattern will persist or shatter under small perturbations.[84] Unlike graph and network theory that describe what systems look like, dynamical systems theory studies what happens and what will become of the system given arbitrary disturbances.

To return to the neuronal network example, the application of dynamical systems theory provides laws of motion in the form of specific systems of differential equations (for the neuronal network, perhaps something like the Hodgin—Huxley type model) that act as rules dictating how states evolve over time (Ermentrout & Terman, 2010). Attractors, or stable points in state space, represent steady patterns of firing rates across the network. Basins of attraction—or sets of each initial state that will evolve into that attractor—can model, e.g., the capacity we have to perceptually complete patterns based on limited information, such as how seeing only the fragment of a face can generate an image of the full face. The full face represents the attractor, whereas identifying the initial fragment represents an initial state within the basic of attraction. Finally, dynamical systems theory can help model the 'stability' of the neuronal network, modeling how perturbations from new information can change how certain firing patterns evolve or remain the same.

In sum, graph theory gives the mathematical tools to model formal facts about the structure of a system; network theory offers methods to generate statistical patterns and hypotheses about system-level function; and dynamical systems theory offers a sophisticated way of modeling the laws of motion governing the physical interactions within such systems, using a battery of differential equations to reveal the causal

[84] For a more formal definition of these terms, see Strogatz (2015).

mechanisms at play while also using these equations to produce a geometry of state spaces modelling the potential trajectory of the system.

B. Kantian Interpretation of Graph, Network, and Dynamical Systems Theory

The formal theories described above are metaphysically neutral. By themselves, they neither prove nor disprove the existence of final causes. Their application to real world phenomena, on the other hand, does make some fundamental presuppositions. To view things purposively is a hypothesis in the Kantian sense in that it is a necessary presupposition for the domain of inquiry that it sparks. I will argue that the application of graph theory, network theory, and dynamical systems theory to study nature presupposes that the phenomena in question are not passive heaps of matter but are dynamic, self-organizing agents. Furthermore, these theories *articulate* the four moments of the principle of purposiveness into a formal and precise language.

In what follows, I will establish a transcendental argument showing that presupposing the four steps—coherent 'horizonal' identity (moment 1), reciprocal causality (moment 2), selective engagement with an environment (moment 3), and a coherency test of persistence or failure (moment 4)—make possible an explanation using graph theory, network theory, and dynamical systems theory. The presuppositions do not themselves explain but set the context for which an explanation becomes intelligible. [85] I

[85] I do not wish to say that practicing complexity scientists have consciously followed Kant's principle of purposiveness when coming up with their theories. Quite a substantial number of mathematicians, scientists, and philosophers of science are either realists or deflationists about final causes. I do not wish to argue that people with these interpretations are 'secretly' Kantian. Rather, the intelligibility of their claims

will then show that anyone who actually applies these theories to explain phenomena cannot consistently deny these presuppositions. The denial would be self-contradicting: the act of explaining commits one to the presuppositions one claims to reject. While unpacking how these theories presuppose the four moments of the Idea of purposiveness, it should be clear how they articulate Kant's somewhat vague Idea using novel formal and conceptual tools. This is because, as Kant predicted, a science becomes aware of its principles the more it matures (see section II). Of course, graph, network, and dynamical systems theory do not prove the Idea of purposiveness—it cannot be proven—but their success provides a powerful validation that the principles it generates are fruitful guides to inquiry, harmonizing our cognitive faculties with nature.

1. Graph Theory

When we choose to represent a system as a graph for the purpose of explaining its self-maintenance or loss thereof, we have already performed an act of judgment. We hypothesize the system must be treated as a purposive whole whose parts are oriented toward its persistence, and whose identity is constituted by the relations that sustain it (moment 1). Without this presupposition, we would have no criterion for selecting nodes and edges—any collection of nodes and any set of edges would be equally valid, and explanation would collapse into arbitrary description. Graph theory articulates this first

relies on presupposing the principle of material purposiveness—a presupposition that is neither realist nor deflationist about agents. This presupposition is not a conscious commitment but a condition for the intelligible application of their formalisms. The question is not what scientists think, but what must be assumed about the system for their formalisms to be rationally applicable at all.

moment in formal language: the system's coherent identity becomes a configuration of nodes and edges, and its boundary becomes the decision to include some relations while excluding others. The boundary is not an external imposition but is internal to the articulation of identity: to say what the system *is* simultaneously to say what it is *not*. Graph theory renders this distinction mathematically tractable.

Could one apply graph theory without the presupposition that the phenomenon in question is a purposive whole—say, by modeling a random collection of rocks connected by arbitrary relations? These rocks, taken as a mere aggregate, do not obviously constitute a system. Yet in the very act of selecting which rocks count as nodes and which relations count as edges, one has already performed a judgment of significance that presupposes a purposive whole: some rocks and some relations matter; others do not. If one truly rejected moment 1, there would be no criterion for choosing edges at all—any rock could connect to any other, or none, or the choice would be arbitrary. Now suppose the application yields no interesting explanations. That outcome does not contradict the presupposition. It merely indicates that the presupposition was not satisfied by the rock pile. The real contradiction emerges only if one makes a stronger claim: that the graph-theoretic model *successfully explains* some system-wide property of the rocks—for example, why the pile is stable or why it collapses under vibration. In that case, one must have already treated the rocks as a systematically unified whole whose parts are oriented toward maintaining that property. One cannot simultaneously claim that the graph explains the pile's persistence and deny that the pile has a purposive identity. The denial is inconsistent: the act of explaining commits one to the very presupposition one rejects.

We may be wrong about which components are significant, and we may revise our initial judgments, but the very act of applying graph theory fruitfully to explain persistence or collapse presupposes that the system can be understood as a purposive whole.

The application of graph theory also presupposes that it is modeling a structure where agents mutually constrain one another (moment 2). A mere aggregate involves no relationality between mutually constraining nodes. Formalized using graph theory, it would represent a bunch of nodes with no edges, illuminating no interesting properties. Insofar as the system has edges, however, we must presuppose an organized unity where the identity of each node is constituted by its relations. This presupposition is what makes possible any subsequent graph-theoretic explanation of the system's internal mechanisms (using, say, dynamical systems theory—see below). Nodes model proto-agents whose identities are defined by the constraints they exert and undergo, whereas edges represent channels of mutual constraint that preserve the nodes they connect. If interactions were destructive, we would represent this with node removal rather than edges. Therefore, since these graphs have edges, interactions are constructive of the nodes. Without the presupposition that the system is sustained by agents mutually constraining one another, edges would be mere marks on paper, and no explanation of the system's internal dynamics would be possible. The choice to model a system with graph theory thus presupposes that its structure is sustained by agents mutually constraining one another, and this presupposition establishes the context within which explanation becomes possible.

One contradicts oneself when they try to deny the presupposition of reciprocal causality while still using graph theory to explain a system's internal dynamics and causal mechanisms. Consider an objector who claims to use graph theory to model a system—say, a set of interacting components—while admitting that the system has a coherent identity (moment 1) but denying that its parts exist in reciprocal causality (moment 2). That is, the objector denies that each part is both cause and effect of the others in an ongoing activity that sustains the whole. Yet the very act of representing the system with edges already commits the objector to reciprocal causality. An edge in a graph, when used explanatorily, is not a mere line, but represents a relation of mutual constraint between nodes. If the objector truly denied moment 2, then each node would be independent, and any edges drawn would be arbitrary or one-way—but then the graph would not explain how the system's internal activity sustains the whole. One cannot simultaneously use edges to represent the relations that constitute the system's ongoing self-maintenance and deny that the parts stand in reciprocal causality. The denial is inconsistent because the explanatory act—drawing edges as channels of mutual constraint—commits the objector to the very presupposition being denied. The choice to model a system with graph theory thus presupposes that its structure is sustained by agents mutually constraining one another in circular fashion, and in doing so articulates Kant's second moment of reciprocal causality in formal terms, where edges are channels of constraint and cycles are feedback loops. Without this presupposition, no explanation of the system's internal constitution or the ongoing activity that sustains it is possible.

But does this mean these two moments constitute the *only* rationale for using graph theory to explain a system's behavior? One might question whether purely spatial adjacency, or logical dependency can be modelled without presupposing mutual constraint or purposiveness. These uses reduce graph theory to a notation, not an explanatory model—they fail to illuminate any interesting properties about the system as a whole and the constraints between its parts. Same thing applies to a physical system of colliding particles that exhibits mutual constraint (moment 2) but no purposive self-maintenance (moment 1). Graph theory might illuminate transient interactions there, but not systemic persistence. Without purposiveness, there is no criterion for which relations are significant for identity over time; any graph would be as good as any other. A third candidate rationale—mathematical convenience (e.g., easy eigenvalue computations)—merely aids calculation, not explanation. Convenience alone does not validate the model; eigenvalues of arbitrary graphs correspond to nothing real about persistence if edges do not represent mutual constraint oriented toward self-maintenance.

Thus, for the specific explanatory goal of understanding how a system maintains its identity through reciprocal constraints, purposiveness (moment 1) and mutual causality (moment 2) are jointly necessary. Other rationales for using graph theory exist—pure topology, data visualization, aesthetic preference—but they do not yield the kind of explanatory content concerning self-maintenance or the loss thereof that concerns us here. Within that explanatory context, these two moments constitute the only complete rationale. Furthermore, graph theory anticipates a structural articulation of selective engagement with the environment (moment 3) by treating salient features in the environment as additional nodes and edges. It also anticipates the coherency test (moment 4), where persistence is determined by a graph having cycles and fragility by having tree-like topologies. Graph theory thus provides the structural skeleton upon which the dynamics of persistence and collapse depend, though it does not explain these dynamics—that task belongs to dynamical systems theory and, in part, network theory.

2. Network Theory

The application of graph theory presupposes that the system is a purposive whole with a coherent identity and boundaries that distinguish it from its surroundings (moment 1) and that its parts mutually constrain one another (moment 2)—presuppositions without which the application would be meaningless. Network theory does not add new presuppositions but provides statistical tools—degree distributions, centrality measures, modularity—to analyze the purposive whole already presupposed. In doing so, network theory articulates Kant's first two moments in statistical language. The system's identity becomes a distribution of connections, and reciprocal causality becomes a matter of differential influence (hubs constrain more than spokes).

A system's connectivity may reveal emergent, system-wide properties that are not reducible to the attributes of any single node or an aggregate of independent nodes. Network theory thus presupposes that emergence is intelligible—that the whole has coordinating properties that cannot be captured by summing its parts. Without this presupposition, the search for system-wide patterns would be irrational. One might discover a power-law distribution as a surprise, but discovery presupposes that the search for such patterns is already a possibility. That rationality comes from the prior assumption that purposive wholes exhibit non-random, system-level regularities. If one genuinely believed emergence were meaningless or unintelligible, one would never bother computing degree distributions in the first place. Hence, the application of network theory—not merely its findings—already contains the presupposition of a purposive whole with emergent properties. This presupposition establishes the context within which network-theoretic explanation of system identity becomes possible.

Could network theory be applied without this presupposition—for instance, to a random set of connections generated by a neutral model? Yes, but such applications serve as null hypotheses or benchmarks, not as explanations of any system-wide properties. Consider someone who agrees that the system has individual nodes and pairwise edges but denies that the whole has any emergent properties that matter for explanation. This person claims that all relevant features are reducible to local interactions. Yet in the very act of computing a degree distribution, measuring modularity, or identifying that a network is scale-free, the objector is already treating the network as potentially having properties that no individual node possesses. A scale-free distribution is not a property of any single node; it is a property of the whole configuration. If the objector truly denied emergent properties, there would be no intelligible grounds to call the network "scale-free" or "modular"—one could only list individual degrees. The explanation of how the system maintains itself or collapses would then be reduced to a collection of local interactions with no account of how they cohere into a system-wide tendency. One cannot simultaneously use network-theoretic tools that detect emergent properties and deny that the system has such properties after it has described them purely formally. The explanatory act—computing global statistics to explain system-wide behavior—commits the objector to the very presupposition being denied.

Beyond emergence, what must be presupposed for network theory to explain how a system's internal causal architecture (moment 2) sustains its persistence? The answer is that certain topological features—high-degree hubs, high-betweenness nodes, modular partitions—can be interpreted as describing *functions*. A hub functions to integrate the network, whereas a gatekeeper functions to control flow between modules. Without this presupposition, we could describe the topology but not explain why the system's integrity depends on these nodes. The functional interpretation is what licenses counterfactuals: "If this hub fails, the system collapses." Such counterfactuals are essential to explanation because they tell us not merely what happened, but what would happen under perturbation. Hence, the presupposition that topological features have functions is a condition of possibility for network-theoretic explanations of persistence.

Avoiding functional language by speaking in purely correlational claims— e.g., "the hub correlates with integrity; remove it, and the system fragments"—explains nothing about why the system is organized that way. To call a hub "critical" or say the system "depends" on it is already to speak functionally. Now consider the claim that one can explain system persistence while denying that any node has a function. One might compute betweenness centrality and predict that removing high-betweenness nodes will fragment the network, while insisting these are merely nodes with a statistical property, not gatekeepers. This objection fails because the prediction itself treats the nodes as playing a role in maintaining the whole. That is a functional claim, avowed or not. You cannot predict that removing a node causes collapse while denying that node has any function in upholding system integrity. The explanatory act—identifying critical nodes to predict perturbation responses—commits the objector to functional interpretation, despite their denial.

Network theory also anticipates how to model the other moments (3 and 4). It can model selective engagement with the environment by using centrality measures—degree, betweenness, and the like—to identify which nodes are most sensitive to environmental inputs and which perturbations would propagate most widely through the system. It also anticipates the coherency test (moment 4) by analyzing degree distributions (e.g., scale-free vs. Poisson) and modularity scores to predict whether the system is robust to random failures, vulnerable to targeted attacks, or prone to fragmentation. Yet in general, graph theory and network theory struggle to provide complete explanations of purposive systems. It is only by being supplemented with dynamical systems theory that one acquires concrete mechanisms to explain all four moments.

3. Dynamical Systems Theory

The application of dynamical systems theory presupposes that the system is a purposive whole with a coherent identity (moment 1), that its parts mutually constrain one another (moment 2), that it is linked to an environment and vulnerable to change over time (moment 3), and that its persistence or failure can be assessed (moment 4). In doing so, it articulates the principle of purposiveness in the language of temporal dynamics—state spaces, attractors, coupled equations, and bifurcations—transforming Kant's schema into a mechanical framework for modeling change over time. Again, I do not claim dynamical systems theory asserts that these features exist in the world as a matter of constitutive fact. Rather, I claim that the only intelligible reason to apply dynamical systems theory to a real system is to model those features.

Systems that dynamical systems theory attempt to describe must already be assumed to have a coherent identity that endures through time despite changes in its internal states. Dynamical systems theory articulates this identity as an attractor landscape—a set of stable patterns toward which the system's activity tends. The system is not a fixed structure but a dynamic pattern of possible trajectories. Without this presupposition, there would be no way to distinguish the system's "same state" from a different state, and explanation of persistence would collapse. The attractor landscape is what makes the system's identity intelligible over time.

Of course, someone might deny the presupposition. They might claim to model a system using dynamical systems theory while denying that it has any attractor landscape—and, with it, denying that the system possesses a coherent, enduring identity. They write differential equations for a collection of particles, perhaps one that exhibits no recurrent patterns. But to write equations for the system at all is already to have presupposed an identity. The moment they interpret their trajectories—saying "the system returned to equilibrium," "entered a new regime," or even "persisted" or "collapsed"—they invoke exactly what they denied. Without an attractor landscape, one could only say "the variables changed," not that the system maintained or lost its identity. If every initial condition leads to a unique, non-repeating trajectory, then there is no system identity to explain—only a sequence of unrepeatable states. Yet the objector's explanatory act (identifying persistence, collapse, equilibrium, or regime shifts) commits them to the very presupposition they rejected. The denial is inconsistent, and explanation of system dynamics collapses.

What must be presupposed for dynamical systems theory to explain how a system's internal activity sustains the whole (moment 2)? The answer is that the parts mutually constrain one another through time—each is simultaneously cause and effect of the others. Dynamical systems theory articulates this as coupled differential equations, where each agent's activity depends both on its own state (self-dynamics) and on the states of others (coupling dynamics). Without this presupposition, the system would be a collection of independent trajectories, and no explanation of how the parts cohere into a sustaining whole would be possible. Reciprocal causality is the condition for any explanation of internal coherence.

Could one apply dynamical systems theory without presupposing reciprocal causality—for instance, by modeling each agent as independent with only one-way influences? One could write uncoupled equations. But then the system would not be organized; it would be a set of independent or hierarchical causes, not a circular whole. Someone might claim to explain a system's persistence using coupled differential equations yet deny that the parts exist in reciprocal causality. They merely write equations where each variable appears in the others' derivatives. Yet in the very act of applying coupled equations to real phenomena, this objector has already committed to mutual constraint. The denial is inconsistent: one cannot simultaneously use equations that encode reciprocal causality and deny that the parts mutually determine one another. The application of the mathematics itself enforces the presupposition.

The system's state space has a basin structure that distinguishes relevant from irrelevant perturbations from its environment. It hence presupposes that the system in question has an environment purposive to its maintenance (moment 3). Dynamical systems theory articulates this as basins of attraction, where perturbations move the system through state space, and the shape of the basins determines whether they are absorbed, amplified, or trigger a transition to a new attractor. Without this presupposition, all perturbations would be equivalent, and there would be no explanation of why the system ignores some environmental changes while reorganizing in response to others. Basin structure is what gives mathematical form to the system's niche.

Let us suppose that one can apply dynamical systems theory without presupposing selective engagement—for instance, by modeling a billiard ball where all perturbations are equivalent except by magnitude. This is possible, but such a model does not explain selective engagement because there is no selection. The contradiction emerges when one claims to explain a system's adaptive or persistent behavior while denying that perturbations are differentially relevant. Someone can model a predator-prey system using dynamical systems theory but deny that the system selectively engages with its environment—denying that some environmental changes matter more than others. In the very act of constructing the model, the objector must decide which environmental variables to include (e.g., food availability, temperature) and which to ignore (e.g., a distant star). That decision already presupposes selective engagement. If all perturbations were equally relevant, the model would be infinite and unintelligible. The denial is inconsistent: one cannot simultaneously build a model that includes some environmental inputs and excludes others and deny that the system selectively engages with its environment.

What must be presupposed for dynamical systems theory to explain whether a system persists or fails (moment 4)? The answer is that trajectories can be evaluated as successful (remaining within a stable attractor basin) or failed (crossing a basin boundary or becoming trapped in a maladaptive attractor). Dynamical systems theory articulates this as stability analysis and bifurcation theory. Success means remaining in a basin, failure means crossing a boundary, and bifurcation analysis reveals tipping points where organization breaks down. Without this presupposition, we could only describe trajectories, not explain why some perturbations lead to collapse and others do not. The coherency test is the condition for any explanation of persistence versus failure.

One might attempt to apply dynamical systems theory without presupposing a coherency test by simply reporting that the system moved from basin A to basin B without calling one state "success" and the other "failure." This is possible, but it forfeits explanatory purchase on why the system resists certain perturbations and succumbs to others. Consider the claim that one can explain a system's collapse using dynamical systems theory while denying that any trajectory constitutes success or failure. One might insist: "I only describe; I do not evaluate." Yet the explanation that "the system crossed the basin boundary because the perturbation exceeded the basin's depth" implicitly invokes basin depth as a measure of resilience—a normative-evaluative concept relative to the system's self-maintenance. Without this presupposition, the explanation would reduce to "the perturbation was large," which says nothing about why the system collapsed rather than absorbed it. The denial is inconsistent: one cannot simultaneously explain why a system fails and deny that failure is a category that applies to the system's own self-maintenance.

In sum, graph theory, network theory, and dynamical systems theory articulate Kant's principle of material purposiveness in three distinct formal languages. Graph theory renders moment 1 (coherent identity) as a configuration of nodes and edges with a definite boundary, and moment 2 (reciprocal causality) as edges representing mutual constraint with cycles as feedback loops. Network theory builds on this articulation by expressing moment 1 statistically through emergent properties such as degree distributions and modularity, and moment 2 functionally through differential influence—hubs and gatekeepers that sustain the whole. Together, graph and network theory make moments 3 and 4 representable by providing the structural skeleton for selective engagement with the environment (moment 3) and coherence (moment 4). Dynamical systems theory then completes the articulation by rendering all four moments in mechanical terms: moment 1 as attractor landscapes, moment 2 as coupled equations, moment 3 as basin structure, and moment 4 as stability and bifurcation analysis. What Kant described philosophically is now mathematically and conceptually palpable. The fruitfulness of the explanations these articulations enable validates the presuppositions that make their application meaningful. The models themselves may be stripped of teleological language, but their success confirms that the regulative principle of purposiveness—when made mathematically explicit—is not a mere heuristic but maintains an objective significance. This success does not prove that nature is constitutively self-organizing, but it provides a felt confirmation that the framework is appropriate to nature, grounding its "truth" not in theoretical knowledge but in how it successfully enables the harmonization of our cognitive frameworks with nature.

C. Phase Transitions and Criticality

Graph theory, network theory, and dynamical systems theory are not the only achievements of the 'complexity' revolution. Although it can be defined almost exclusively using dynamical systems theory, the notion of 'criticality' was an astounding conceptual discovery. Initially used in studying phase transitions in thermodynamics, it is now integrated with a broader set of theories to illuminate how systems leverage state spaces that hover close to system collapse for adaptive success. Like the other theories mentioned thus far, criticality both presupposes and further articulates the four moments of purposiveness.

Traditionally, thermodynamics and statistical mechanics in the 19th and early 20th centuries viewed phase transitions as shifts between equilibrium states. The transition was viewed as a single instantaneous event, driven by thermodynamic potentials. That is, when the free energy of a new phase becomes lower than the old phase, the system rapidly transitions to a new equilibrium phase. Complex systems scientists, on the other hand, began to study the transition process itself. No longer are individual particles minimizing energy but are 'actants' coordinating themselves to create new relationships amongst themselves. This is a shift from concentrating on the various material properties of the passive particles (e.g., their density and magnetization) to studying the relational structure between the particles treated as interacting proto agents.

Complex systems scientists are interested in the transition points between a relatively ordered phase (e.g., the solid phase of water) and relatively disordered phase (e.g., the liquid phase of water). In the ordered phase, a rigid global constraint forces elements of the system to align, whereas in a disordered phase elements are only weakly correlated. At the so-called 'critical point' the global constraint is strong enough to correlate some elements, but the tendency for disorder prevents the elements from settling into a stable order. At such a point, the system is extremely sensitive: the smallest perturbance can send it to an ordered or disordered phase. From the perspective of dynamical systems theory, this is a 'bifurcation point' (Strogatz, 2015). Bifurcations are qualitative, topological changes in the total state space where the attractor landscape is fundamentally reshaped. This concept is useful in articulating how the system's internal organization can break down and reconfigure under environmental stress.

In a critical phase, deep, stable valleys (the ordered phases) flatten into a vast, flat plain. At this sensitive stage, one small push could send the system into a new valley, a new stable state. In the system, order is patchy, where tiny islands of order form themselves into larger-ordered parts. What is interesting is that these levels of order are scale-invariant: the statistical patterns of fluctuations look identical at every magnification, a defining feature known as fractal geometry (Mandelbrot, 1982) and characterized by universal critical exponents (Stanley, 1971). The rule of correlation—the constraint—operates the same way at every magnification, just as a fractal has the same patterns the closer you zoom in. In this state, the system is maximally correlated, where a tiny local event can propagate and influence the entire system, a phenomenon known as 'avalanche' dynamics whose size and duration follow a power-law distribution (Bak et al., 1987). They are also highly unstable for the same reason.

This phenomenon is known as the 'edge of chaos' (Langton, 1990): a system neither rigid nor chaotic, but maximally sensitive to perturbations, allowing certain signals to serve as information that rapidly propagate across the whole without being lost. In state space terms, criticality appears as a flat basin of attraction where the smallest nudge sends the system into a ceaseless search for metastable states, where small signals can trigger large-scale adaptive reorganization. For simple systems, success meant staying within a stable attractor basin. For complex systems at criticality, on the other hand, success means maintaining proximity to the critical threshold without crossing over—remaining poised at the edge of chaos, maximizing adaptability and transmissibility of information without disintegrating. Such systems are maximally sensitive to their environment while also possessing the largest possible repertoire of internal states, a hallmark of a state space where the number of accessible attractors is maximized. This state of pure potentiality is achieved only by sufficiently complex systems—neuronal networks (Chialvo, 2010), evolutionary processes (Kauffman, 1993), and functioning societies (Bak et al., 1987)—which constantly move between fleeting stable states while remaining metastable, avoiding the extremes of chaos and rigid order.

1. Presupposition and Articulation of the Four Moments

The application of criticality to real world systems presupposes that it has a purposive identity (moment 1) that can be positioned along a continuum from simple stability to critical metastability. A system that cannot leverage criticality (like ice at its melting point) still has an identity—it is either solid or liquid—but its purposiveness is shallow, since only two attractors exist, and it cannot maintain itself at the critical threshold. A system that can leverage criticality (like a brain or ecosystem) is more complex because it can remain poised at the edge of chaos, balancing stability and adaptability. Without this presupposition of purposive identity as a continuum, we could not distinguish a system that *fails* to maintain criticality from one that simply does not exhibit it. The distinction between failure and absence is essential to explanation.

Consider someone who claims to explain a system's behavior using criticality theory yet denies that the system has any purposive identity oriented toward self-maintenance. The objector measures avalanche sizes in a sandpile and finds a power law. When asked what this explains, the objector admits this says nothing of persistence or failure—it is just a statistical fact. But if that is all, then the application is merely descriptive, not explanatory. The performative contradiction emerges when the objector then claims that criticality *explains* why the system is sensitive to perturbations or why it reorganizes. That explanation requires treating the system as having an identity that is maintained or lost. One cannot simultaneously claim that criticality explains the system's adaptive flexibility and deny that the system has a purposive identity that does the adapting. The denial is inconsistent.

Reciprocal causality (moment 2) operates across all scales simultaneously to maintain a whole. At the critical point, scale-invariant patterns of correlation reveal that a tiny local event can propagate across the entire system, demonstrating that no part is independent. Criticality articulates moment 2 as a dynamic feedback loop between scales, where the system's internal organization is constantly tested and reconfigured. Without this presupposition, we would simply observe that large events are rare. We would lack the explanation that the system's organization makes it maximally sensitive across all scales. The rationale for applying criticality theory is to capture this circular causality where local interactions and global organization mutually determine one another.

One could measure scale invariance without interpreting it as mutual constraint. Someone can claim to explain a forest fire system using criticality theory yet deny that local events and global organization mutually determine one another. The objector simply asserts that this is a mere statistical property. But then this objector cannot explain *why* the system remains at criticality or *why* a small spark can sometimes cause a continent-scale fire. The explanation of cascade propagation depends on the claim that local rules and global connectivity are mutually constraining. If the objector truly denied reciprocal causality, the scale invariance would be an unexplained coincidence. One cannot simultaneously use the scale-invariant correlation as an explanation of the system's sensitivity and deny that the parts and whole mutually constrain one another. The explanatory act commits the objector to moment 2.

The system's niche (moment 3) is defined by its capacity to detect and respond to environmental signals without being overwhelmed. At the edge of chaos, the system is maximally sensitive: a small signal can trigger large-scale adaptive reorganization. Criticality articulates this moment by showing that the system's relation to its environment is not merely selective but dynamic—selective components of the environment constantly tests the system, and success depends on the system's capacity to respond without disintegrating (Scheffer et al., 2009). Without this presupposition, all environmental inputs would be equivalent, and there would be no explanation of why the system amplifies some perturbations and ignores others.

Consider an objector who models a neuron membrane near its firing threshold using criticality theory but denies that the neuron selectively engages with its input environment. The objector asserts that the neuron simply has a threshold where inputs above it cause spikes and inputs below it do not. No selection is involved. Yet the very fact that the threshold is set where it is—balancing sensitivity and stability—already implies that some inputs matter (those near threshold) and others do not (those far below). The objector's own model distinguishes relevant from irrelevant perturbations. If all inputs were equally relevant, the neuron would fire randomly or not at all. One cannot simultaneously use criticality theory to explain why the neuron detects signals amid noise and deny that the neuron selectively engages with its environment.

Success for a system (moment 4) means maintaining proximity to the critical threshold without crossing over—remaining poised at the edge of chaos, maximizing adaptability without disintegrating. For complex systems at criticality, the goal is not static self-preservation but the preservation of the capacity for transformation itself. Complex systems like brains and ecosystems exhibit flattened, fractal landscapes of shallow metastable attractors (Tognoli & Kelso, 2014), where small perturbations trigger adaptive reorganization, enabling learning, memory, and innovation. Criticality articulates this moment by making the coherency test mathematically tractable, where the distance from the critical threshold is the measure of coherence. Without this presupposition, we could only describe activation statistics. We could not ask whether the system is succeeding or failing at its self-maintenance.

Could one apply criticality theory without presupposing a coherency test? One might attempt to model, e.g., an immune system using criticality theory while claiming to simply report whether it is subcritical, critical, or supercritical, without evaluating any state as success or failure. But this attempted neutrality collapses upon examination. The explanation of why the immune system avoids autoimmunity (supercritical) and avoids tolerance (subcritical) depends on the normative claim that the critical regime is preferable for the organism's self-maintenance. The model itself predicts that moving away from criticality leads to disfunction. That prediction only makes sense if disfunction is understood as failure from the perspective of the system's own self-preservation. One cannot simultaneously explain why the immune system stays near criticality (because deviation causes harm) and deny that harm or failure are categories that apply to the system.

Criticality theory does not introduce new presuppositions beyond Kant's four moments, but articulates those moments in a language suited to systems poised between order and chaos. Moment 1 (purposive identity as a continuum) is articulated through the distinction between simple systems (ice, lasers below threshold) that cannot leverage criticality and complex systems (brains, ecosystems) that can remain at the edge. Moment 2 (reciprocal causality across scales) is articulated through scale-invariant correlations, where local interactions and global organization mutually determine one another. Moment 3 (selective engagement with the environment) is articulated through the system's capacity to detect and respond to perturbations without disintegrating, formalized as a threshold that filters inputs. Moment 4 (coherency test as preservation of transformative capacity) is articulated through distance from the critical threshold as a measure of coherence, where success means maintaining adaptability without tipping into chaos or rigidity. The only intelligible reason to apply criticality theory to explain any of these systems—rather than merely describe their statistical properties—is to capture these four moments in formal, quantitative terms

D. Panarchy and Resilience

Complex systems theory fundamentally employs the capacity to model cross-scale interactions. To dance near criticality, for instance, requires lower-scale regulatory mechanisms to be in place so that the system does not disintegrate when it gets too close to a critical point. In a simple, single-scale system (such as water boiling), approaching a critical point leads to a phase transition from liquid to gas. There are no mechanisms to stabilize it near the brink. More complex systems harbor many interactions across scales to ensure that such disintegration does not happen, while not being too rigid to adapt to change. The theory of 'panarchy' provides a framework to study how cross-scale interactions dynamically manage a system (Gunderson & Holling, 2002). It shows how self-organizing 'adaptive cycles' interact at various scales. Before examining these cross-scale interactions, we must first understand the adaptive cycle itself.

The adaptive cycle contains four phases: r (growth), K (conservation), Ω (collapse), and α (re-organization) (Holling 1986; 2001). During the transition from growth (r) to conservation (K), the system becomes more densely and efficiently connected, creating a highly ordered but less diverse network topology. As connectivity increases, the system's dynamics begin to reinforce a single, deep 'valley' in state space. The system's state is pulled strongly into this attractor basin. Although stable against small perturbations, it has lost its flexibility, becoming overconnected and brittle (Scheffer et al., 2001). It has traded a wide, flat basin (many possible stable states) for a deep, narrow one (one very stable state).

The next phase is from conservation (K) to collapse (Ω). A perturbation, such as a new pathogen or a new technology, upends the system's stability, leading to the catastrophic collapse of the system. The system is pushed out of its stable basin into a region without another stable attractor nearby. The state space flattens out, becoming extremely close to criticality. It has maximum potential but absolutely no stability.

After collapse (Ω), the system begins to reorganize (α). Intense experimentation for global stability between the system's agents begins, where new forms are transiently created and destroyed. Eventually, new attractors are born, carving depressions in state space. The system's total state starts to orbit around these attractors. When attractor basins become deep enough, positive feedback mechanisms deepen them further. Energy and resources are efficiently channeled to exploit these new configurations. The system enters a growth phase (r) again, but with new patterns of behavior and connections. The deepening of this basin to rigidity (K) leads to collapse and the repetition of the cycle.

Panarchy describes that these cycles are nested at various scales. These scales have their own dynamics, all of which contribute to the structure of any given system. Perturbations in a system are often caused by faster, smaller scale systems that the rigid conservation phase can no longer absorb. These faster cycles do not only send a conserved system to collapse, but can also provide innovations that help save it. They push the slow system out of rigid equilibrium, forcing it to become more sensitive and adaptable. For example, a flu virus is a smaller scale system that can kill you (a much larger system), or it can get you sick temporarily so that you develop antibodies for future strains. Fast cycles 'revolt' in this way: they push a system closer to criticality, which could lead to innovation or collapse (Gunderson & Holling, 2002; Holling, 2001).

So-called 'remember' effects, by contrast, are ways in which slow, large-scale systems impact fast, small-scale ones. Slow cycles are deep, stable attractor basins built over long periods. They have a 'memory' of successful configurations that impacts the dynamics of systems at faster and smaller scales. When in a reorganization (α) phase, a system may draw upon these patterns to shape new potential attractors. The remember function pulls a system away from criticality toward new, coherent attractors (Gunderson & Holling, 2002). For example, even though the people serving the US government changes every couple years, the Constitution acts as a slower, larger scale system that serves to 'remember' an overarching sense of order, regulating novel changes so that the government does not collapse.

Revolt and remember—the impact of lower-scale, faster systems on slower systems, and the impact of higher-scale, slower systems on faster systems—represent the two fundamental ways scales interact. These contribute to the 'panarchy' of scales. The adaptive cycle is scale-invariant: there is no system too small or too big to abide by its dynamics. Cycles of all scales constantly interact, with interactions often determined by their relative stage in the cycle.

1. Presupposition and Articulation of the Four Moments

To apply the adaptive cycle to a real phenomenon presupposes that the system's identity is not fixed but undergoes transformation through time (moment 1). As the system moves from growth to conservation, its identity becomes rigid and brittle, whereas during collapse, its identity disintegrates. During reorganization, however, a new identity is forged. Where Kant described the subject's identity as a stable form oriented toward self-preservation, panarchy articulates this identity as a dynamic trajectory through phases of growth, conservation, collapse, and reorganization. Identity is not a fixed state but a cycle of becoming, stabilizing, breaking down, and renewing. Furthermore, a system's identity is defined by its relationship to systems at both smaller/faster and larger/slower scales. The rationale for applying panarchy to a system (e.g., a grassland, a corporation, a civilization) is precisely to explain how its identity is maintained through transformation.

One could simply label phases: "this system is in growth, now conservation, now collapse," denying any teleological presuppositions. But consider an attempt to explain a forest's persistence over centuries using panarchy while denying that the forest has any identity that endures through the cycle. The explanation would refer to the forest's return to similar structures after fire, its memory stored in soil seed banks, its recurrent patterns of succession. Yet each of these references treats the forest as the same system undergoing transformation. If the forest had no identity across phases, there would be nothing to explain—one would merely describe a sequence of different configurations with no connection. One cannot simultaneously explain how a system renews itself after collapse and deny that the system has an identity that persists through the cycle.

Reciprocal causality (moment 2) operates across nested hierarchies, not just within a single system. The system's ongoing activity is co-produced by revolt from below (fast, small cycles pushing toward criticality) and remember from above (slow, large cycles providing stability and memory). Where Kant described reciprocal causality as the circular relation between parts and whole within a single system, panarchy adds that this reciprocity is distributed across scales. Without this presupposition, we would treat each scale as independent and miss how smaller cycles disrupt larger ones and larger cycles constrain smaller ones.

Could one apply panarchy without presupposing cross-scale reciprocal causality? One could describe each scale in isolation. But consider an attempt to explain a fishery's collapse using panarchy while denying that fast fish population cycles (revolt) and slow ocean current regimes (remember) mutually constrain one another. The explanation would refer to how overfishing at the fast scale triggers a regime shift at the slow scale, and how the slow scale's memory prevents recovery. Yet these references already assume cross-scale reciprocity. If the scales were truly independent, the fast dynamics could not affect the slow, and the slow could not constrain the fast. One cannot simultaneously explain how collapse propagates across scales and deny that scales mutually constrain one another.

The environment, on a panarchic view, is no passive backdrop but an active, cross-scale ecology of nested adaptive cycles (moment 3). A system's niche is not given in advance but constituted by its position within this multi-scalar field: smaller, faster cycles revolt to constitute the environment of larger, slower systems, while larger, slower cycles remember to shape the possibilities of smaller, faster ones. Where Kant described the niche as a domain of selective input that the system draws upon for self-maintenance, panarchy adds that this domain is itself an ecology of interacting systems at different scales, each pursuing its own purposive trajectory. To mistake the environment for a fixed background is to miss this dynamic, cross-scale agency—the very agency that determines which perturbations matter, which resources are available, and which risks threaten the system's coherence.

One could treat the environment as a source of fixed resources and disturbances. But consider an attempt to explain a wolf pack's foraging behavior using panarchy while denying that the forest successional cycle (slower, larger) and rodent population cycles (faster, smaller) actively constitute the wolves' niche. The explanation would refer to how a rodent crash starves the wolves' prey, or how a forest fire opens new hunting grounds. These references treat the slower and faster cycles as constituting what counts as a perturbation or opportunity for the wolves. If the environment were merely passive, the wolves' relationship to it would be the same regardless of these cycles. Hence, one cannot simultaneously explain how nested cycles shape a system's selective engagement and deny that those cycles actively constitute its environment.

Kant's coherency test posed a binary outcome: persistence or collapse. Panarchy transforms this test by relocating it across scales. What appears as failure at one scale—a forest fire, a population crash, a firm's bankruptcy—may enable success at another scale, such as ecosystem renewal, genetic innovation, or market reallocation. Collapse, in other words, is not an endpoint but a phase within a larger cycle. A system's success (moment 4) is therefore redefined as resilience: the capacity to coordinate across scales, where slower scales constrain chaotic reorganization and faster scales provide innovation to prevent stagnation. Without this cross-scale presupposition, we would mistake collapse for mere cessation and miss its generative role in renewal.

One might attempt to remain purely descriptive, simply noting phases without evaluating any as success or failure. But such neutrality collapses as soon as explanation begins. Explaining why a forest ecosystem is resilient—why it regrows after fire, why modular structure prevents contagion, why seed banks enable recovery—already commits one to treating these features as good for the forest's persistence. These are not neutral observations but are normative judgments about what enables the system to maintain itself. If resilience were not a standard of success, the features cited would be merely interesting facts, not explanations of why the system endures. To explain persistence is already to treat persistence as a success condition. Hence, one cannot explain why a system's structure enables it to persist through disturbance while denying that persistence is a normative standard for that system.

Panarchy thus gives Kant's principle of purposiveness a new articulation across all four moments. Moment 1 (identity as dynamic cycle) is articulated through the phases of growth, conservation, collapse, and reorganization, where stability is the recurrence of the cycle, not the persistence of content. Moment 2 (reciprocal causality across scales) is articulated through revolt from below and remember from above, where fast and slow cycles co-produce the system's ongoing activity. Moment 3 (selective engagement with environment) is articulated through the recognition that the environment is an active ecology of nested adaptive cycles, where a system's niche is constituted by its position within this multi-scalar ecology. Moment 4 (coherency test as resilience) is articulated through the capacity to coordinate across scales, where failure at one scale enables innovation at another, and success means navigating the full cycle of death and rebirth.

IV. Objections, Conclusion

Kant's regulative principle of purposiveness helps guide scientific explanations by considering not what natural entities 'are' but what they 'do.' Instead of the study of passive atoms obeying certain laws of motion, we begin our study of nature postulating a realm of active, organizing matter. This agential stance acts as a hypothesis in the Kantian sense: a necessary presupposition that opens up a domain of inquiry. As we have seen in complex systems science alone, under this hypothesis we can fruitfully model systems in terms of interacting nodes, apply statistical tools to quantify emergent properties, assign laws of motion and create state spaces that characterize evolution, study why certain organized beings are more adaptive than others using criticality, and investigate cross-scale interactions within an ecosystem of nested agents of all sizes and speeds.

I have argued throughout this chapter that—contra the deflationist—the fact that we can describe these systems in purely formal, non-agential language does not purge nature of agency. On the contrary, since modeling systems using graph theory, network theory, and dynamical systems theory each presupposes the four moments of the principle of purposiveness, it points to the indispensability of the agential stance rather than its elimination. The stance itself has no explanatory power in the sense of a metaphysical ground; it does not explain why, e.g., a particular hurricane exists. Yet it sets the context of explanation by connecting abstract mathematical frameworks to real phenomena. The truth of this framework is something we can feel due to its wild success from modeling a diversity of phenomena, from brains to cells to economic systems. We are warranted to feel (in the Kantian, a priori sense—see section II) as if we live in a world of purposive agents of all scales, but we can never know this in the constitutive sense. The realist intuition—but not the realist ontology—is justified.

Someone might object: "I do not need to consider a hurricane as an agent to predict its path. I just solve an equation with initial conditions." Kant would not dispute this. The replacement of agential language with purely formal vocabulary is a marker of the success of Kant's principle of purposiveness. Kant's point is not whether individual scientists must treat hurricanes as agents, but that the applicability of certain mathematical models to hurricanes already presupposes that the hurricane is the kind of entity to which such formalisms rightfully apply—an entity with the unity, coherence, and active self-maintenance (via mutual constraint) characteristic of a natural purpose or self-organizing agent. That we can solve an equation representing a hurricane's movement independently of our awareness is but the expression of the discovered internal principle of the purposive whole.

One might also worry that the principle of purposiveness is anthropomorphic—that it projects human agency onto nature. In Chapter 3 we saw that Kant grounded the intelligibility of this principle by analogy to the technical (artifact-producing) agency of human beings. Yet the core of the analogy, as we saw there, is minimalist. It attributes only certain formal features of agency to nature, stripped entirely of the intentionality and goal-oriented structure that characterizes human agents. For a Kantian, saying that natural agents have 'goals' to self-persist merely means that if a system does not have self-reinforcing constraints, it ceases to exist as that system. Persistence is a prerequisite for being an object of study at all. This is a bare-bones sense of agency, defined purely in terms of whether the agent succeeds in constraining both its parts and its environment.

Is the principle of purposiveness refutable? Kant thinks that it is a necessary hypothesis for studying organized beings. This might seem to beg the question—defining agency so broadly that any stable physical structure qualifies. But this is somewhat the point. Along with the lens of mechanism—which hypothesizes that we can describe things in nature in terms of laws of motion—the lens of purposiveness comprises a fundamental context of explanation for studying any organized being, from living organisms to inorganic systems like hurricanes.

What emerges from this critique is that complex systems science does not refute Kant's regulative principle of purposiveness rather than articulates it. Graph theory, network theory, dynamical systems theory, criticality, and panarchy do not replace the agential stance with a purely mechanistic description. Rather, they presuppose that stance and, in their success, allow us to feel their objectivity. The four moments of purposiveness—coherent identity, reciprocal causality, selective engagement with an environment, and the coherency test of persistence or failure—are the formal conditions that make the application of these mathematical tools intelligible in the first place. To model a neural network as a graph, to analyze its statistical topology, to simulate its dynamics, and to study its adaptability through criticality and its resilience through panarchy is already to judge it as a purposive whole. The models themselves may speak only in the language of nodes, edges, equations, and attractors, but their very intelligibility rests on a prior judgment that the system in question is the kind of thing that can be understood in these terms.

This does not mean we have proven that nature is constituted by self-organizing agents. The regulative principle remains a problem, claiming something about things that we can never fully know. What the success of complex systems science provides is a felt confirmation of the truth of the principle. This is not full-blown knowledge, but the disinterested pleasure of seeing the same formal structures unify phenomena as diverse as chemical oscillations, gene regulation, ant colonies, and financial markets. We are warranted to feel as if this cognitive presupposition harmonizes with a world of purposive agents of all scales. However, we can never know this in the way we know physical laws. The realist intuition is justified, but not the realist ontology. And the deflationist's desire to purge nature of agency overlooks the fact that the very formalisms they celebrate derive their fruitfulness from a presupposition they would dismiss as a mere heuristic, a heuristic whose indispensability remains unexplained.

In the end, Kant's regulative principle of purposiveness turns out to be a generative source of mathematical and empirical discovery. It creates a science in Kant's broader sense, uniting a vast range of modeling techniques and conceptual machinery to create a systematic body of knowledge. Complex systems science both presupposes this principle and articulates it further. What Kant described in abstract philosophical terms—the self-organizing whole whose parts reciprocally constrain one another and whose persistence is tested against an environment—has become a precise formal language. The seed he planted has grown into a flourishing plant. And as with any living idea, its growth is far from over.

Bibliography

Bak, P., Tang, C., & Wiesenfeld, K. (1987). Self organized criticality: An explanation of 1/f noise. *Physical Review Letters*, 59 (4), 381-384.

Barabási, A. (2016). *Network Science*. Cambridge University Press.

Berkes, F., & Folke, C. (Eds.). (1998). *Linking Social and Ecological Systems: Management Practices and Social Mechanisms for Building Resilience*. Cambridge University Press.

Biggs, R., Schlüter, M., & Schoon, M. L. (Eds.). (2012). *Principles for Building Resilience: Sustaining Ecosystem Services in Social-Ecological Systems*. Cambridge University Press.

Boccaletti, S., Latora, V., Moreno Y., Chavez M., & Hwang, D.-U. (2006). Complex Networks: Structure and Dynamics. *Physics Reports*, 424 (4-5), 175-308.

Campbell, R. J. (2015). *The Metaphysics of Emergence.* Palgrave Macmillan.

Carpenter, S. R., & Brock, W. A. (2008). Adaptive capacity and traps. *Ecology and Society*, 13 (2), 1-16.

Chialvo, D. R. (2010). Emergent complex neural dynamics. *Nature Physics,* 6, 744-750.

Cummins, R. (1975). Functional analysis. *The Journal of Philosophy,* 72 (20), 741-765.

DeLanda, M. (2002). *Intensive Science and Virtual Philosophy*. Continuum.

Ermentrout, G. B. & Terman, D. H. (2010). *Mathematical Foundations of Neuroscience*. Springer.

Folke, C., Carpenter, S.R., Walker, B., Scheffer, M., Chapin, T., & Rockström, J. (2010). Resilience thinking: integrating resilience, adaptability and transformability. *Ecology and Society*, 15 (4), 1-9.

Freeman, L. C. (1977). A Set of Measures of Centrality Based on Betweenness. *Sociometry,* 40(1), 35-41.

Gunderson, L. H. & Holling, C.S. (2002). *Panarchy: Understanding Transformations in Human and Natural Systems*. Island Press.

Harris, J.M., Hirst, J.L., & Mossinghoff, M.J. (2008). *Combinatorics and Graph Theory*. Springer Science + Business Media.

Holling, C.S. (1986). The Resilience of Terrestrial Ecosystems: Local Surprise and Global Change. In W.C. Clark & R. E. Munn (Eds.), *Sustainable Development of the Biosphere* (292-317). Cambridge University Press.

Holling, C. S. (2001). Understanding the Complexity of Economic, Ecological, and Social Systems. *Ecosystems*, 4, 390-405.

Juarrero, A. (1999). *Dynamics in Action: Intentional Behavior as a Complex System*. MIT Press.

Juarrero, A. (2023). *Context Changes Everything: How Constraints Create Coherence*. MIT Press.

Kauffman, S. (1993). *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press.

Kant, I. (1998). *Critique of Pure Reason* (P. Guyer & A. Wood, Eds.) Cambridge University Press.

Kant, I. (2000). *Critique of the Power of Judgment* (P. Guyer, Ed.). Cambridge University Press.

Kant, I. (2004). *Metaphysical Foundations of Natural Science* (M. Friedman, Ed.). Cambridge University Press.

Kant, I. (2015). *Critique of Practical Reason.* Cambridge University Press.

Landau, L.D. (1937). On the theory of phase transitions. *Zhurnal Eksperimental'noi I Teoreticheskoi Fiziki,* 7, 19-32.

Langton, C.G. (1990). Computation at the edge of chaos: Phase transitions and emergent computation. *Physica D: Nonlinear Phenomena,* 42 (1-3), 12-37.

Latour, B. (2005). *Reassembling the Social: An Introduction to Actor-Network Theory*. Oxford University Press.

Mandelbrot, B. B. (1982). *The Fractal Geometry of Nature*. W. H. Freeman and Company.

Maturana, H. R. & Varela, F. J. (1980). *Autopoiesis and Cognition: The Realization of the Living*. D. Reidel Publishing.

Neander, K. (1991). Functions as Selected Effects: The Conceptual Analyst's Defense. *Philosophy of Science*, 58 (2), 168-184.

Millikin, R. (1989). In Defense of Proper Functions. *Philosophy of Science*, 56 (2), 288-302.

Newman, M.E.J. (2006). Modularity and community structure in networks. *Proc. Natl. Acad. Sci. U.S.A.,* 103 (23), 8577-8582.

Newman, M.E.J. (2010). *Networks: An Introduction*. Oxford University Press.

Scheffer, M., Carpenter, S., Foley, J.A., Folke, C., & Walker, B. (2001). Catastrophic shifts in ecosystems. *Nature*, 413, 591-596.

Scheffer, M. et al. (2009). Early-warning signals for critical transitions. *Nature,* 461, 53-59.

Shew, W. L., & Plenz, D. (2013). The functional benefits of criticality in the cortex. *The Neuroscientist,* 19 (1), 88-100.

Song, C., Havlin, S., & Makse, H.A. (2005). Self-similarity of complex networks. *Nature*, 433, 392-395.

Stanley, H. E. (1971). *Introduction to Phase Transitions and Critical Phenomena*. Oxford University Press.

Strogatz, S. (2015). *Nonlinear Dynamics and Chaos*. Taylor & Francis.

Thompson, E. (2007). *Mind in Life: Biology, Phenomenology and Sciences of Mind*. Harvard University Press.

Tognoli, E. & Kelso, J. A. S. (2014). The metastable brain. *Neuron*, 81 (1), 35-48.

Van Steen, M. (2010). *Graph Theory and Complex Networks: An Introduction*. Maarten Van Steen.

Wilson, K. G. (1971) Renormalization group and critical phenomena. I. Renormalization Group and the Kadanoff Scaling Picture. *Physical Review B,* 4(9), 3174-3183.

Wright, L. (1973). Functions. *Philosophical Review,* 82 (2), 139-168.

Yates, F. E. (1994). Order and complexity in dynamical systems: Homeodynamics as a generalized mechanics for biology. *Mathematical and Computer Modelling*, 19 (6-8), 49-74.

Chapter 6: Kant, the Highest Good and Environmental Ethics

I. Introduction

Environmental ethicists have long found Kant's moral philosophy wanting. The prevailing reception charges Kant's ethics with the ultimate environmental sin: anthropocentrism. This occurs on two levels.[86] First, Kant is charged with drawing an arbitrary boundary between humans and non-humans: only rational beings qualify as ends-in-themselves[87], while everything else in nature is relegated to the status of mere things. Second, Kant's indirect-duty framework is charged with rendering nature merely instrumental—we have no direct duties to animals or ecosystems, only duties to ourselves and other persons that sometimes require us to treat nature well. Together, these two levels of anthropocentrism have led many environmental philosophers to conclude that Kant's ethics is incapable of grounding the robust, non-instrumental obligations to the non-human world demanded by contemporary ecological crises.

This paper contends that the traditional critique overlooks a deep-seated implication at the core of Kant's practical philosophy. As discussed in chapter 2, Kant's account of moral agency is inherently bi-constitutive. In legislating the moral law, the

---

[86] Both of these levels are described cogently in Tony Svoboda (2015)
[87] Tom Regan (1983) holds the most prominent pro-Kantian account. His account, however, is only Kantian in spirit rather than grounded in Kant's deontological structure, extending Kantian direct duties to "subjects-of-a-life," grounding direct duties to animals in their psychological capacities as individuals. My account, by contrast, derives a direct (yet imperfect) duty to preserve nature's purposive unity from a contradiction in will argument that remains fully within Kant's own framework. Moreover, my criterion for moral consideration (integrative versus disintegrative form) extends to non-sentient natural agents like ecosystems and geo-chemical cycles, which Regan's subject-of-a-life criterion cannot reach.

rational agent does not merely perfect itself in isolation but simultaneously engages in the world-constituting project of actualizing the highest good—a moral world in which nature is purposively aligned with freedom such that those who deserve to be happy are happy.[88] This transforms our relationship to the non-human world. Our duty is not merely to use nature as an instrument for self-cultivation (the indirect model), but to preserve and promote the functional integrity of nature as constitutive of the very moral world we are obligated to bring about. The stability and purposive unity of the natural order thus become a direct duty, as it integral to the highest good.[89]

The argument proceeds in four additional sections. Section II surveys the existing landscape of Kantian environmental ethics, analyzing indirect-duty accounts, revisionist direct-duty arguments, and sophisticated indirect-duty views, concluding that even the most robust among them remain anchored in duties of moral self-perfection and so cannot adequately address our obligations toward ecosystems as integrated wholes. In Section III, I argue that the highest good constitutively requires a purposively unified natural world, and that a maxim of degrading this unity yields a contradiction in will,

---

[88] "I call the world as it would be if it were in conformity with all moral laws…a moral world….Thus far it is therefore a mere, yet practical, idea, which really can and should have its influence on the sensible world, in order to make it agree as far as possible with this idea" (A808/B836).

[89] My reading shares many affinities with John Martin Gillroy (1998) account, particularly with his reading of Kant that emphasizes nature's functional integrity as morally significant and rejection of purely instrumental attitudes toward the natural world. However, we diverge in several respects. First, Gillroy claims that nature constitutes an "end-in-itself" with "intrinsic value" comparable to persons; I reject this, holding that the duty is direct regarding nature, that nature is an intrinsically valuable end, but not an end-in-itself.. Second, Gillroy grounds his position in a teleological reading of Kant's *Critique of Judgment* and *Opus Postumum*, whereas my argument derives the duty from a contradiction in will test applied directly to the categorical imperative and the structure of the highest good. Third, my account provides a multi-scalar, formal criterion for distinguishing integrative from disintegrative natural forms—a distinction largely absent from Gillroy's treatment of "functional integrity." Finally, the duties I derive are imperfect (allowing latitude in execution), whereas Gillroy's prescriptions appear to function as perfect duties.

generating a direct but imperfect duty to preserve nature's self-organizing form. I further argue that this duty is internal to the Formula of Humanity itself, since to will the permissible happiness of embodied others is to will nature in harmony with itself, making the duty regarding nature direct without claiming nature is an end-in-itself. Section IV bridges this argument with contemporary ecology, arguing that nature's purposive unity that we are obligated to bring about is a dynamic, multi-scalar architecture of interlocking forms. This provides a systematic, non-arbitrary criterion for moral considerability in non-human agents, understood not as static substances but as practices. This yields two imperfect duties: to refrain from disintegrative practices and to promote integrative ones. Finally, Section V confronts the objection that a wholly artificial "schmecosystem" could satisfy our duties, reducing nature to a replaceable instrument. Drawing on an analogy with David Enoch's "schmagency" objection to constitutivism, I argue that morality requires a genuinely self-organizing cosmos, since replacing nature with a curated artifact would collapse moral necessity into solipsistic play. The duty is to the form of natural self-organization itself—the meta-agent with which we must harmonize. Overall, I advance a reading of Kant in which duties to nature and duties to humanity are not separate obligations, the first instrumental to the second, but two dimensions of a single practical commitment.

II. Interpretations of Kant's Environmental Ethics

Most scholarship on Kant's duties regarding non-human nature focuses on animals and accepts what has become the standard "traditional" interpretation. This interpretation holds that for Kant, our obligations regarding animals are merely indirect, derived from a psychological tendency. Cruel treatment of animals desensitizes a person to suffering in general, thereby making one more likely to disregard human suffering and fail in one's direct duties to other persons. As scholars like James Skidmore (2001) argue, this establishes only a weak link between morality and the treatment of non-humans. One can, for instance, imagine someone who is cruel to animals yet scrupulously fulfills duties to humans. Consequently, this reading renders Kant's position as offering only a prudential reason for kindness to animals, dimming the prospects for a substantive Kantian environmental ethic. It is against this backdrop of the perceived inadequacy of the traditional interpretation that contemporary Kantian scholars have sought more robust foundations, branching into two main camps: those who argue for direct duties to non-humans and those who develop a more sophisticated account of indirect duties rooted in the necessary cultivation of moral character.

Dissatisfied with the traditional view, Christine Korsgaard (2018) argues that we have *direct* duties to animals. Korsgaard is aware that although she intends to remain within a Kantian framework, she must address Kant's explicit claims to the contrary, where he argues that "a human being has duties only to human beings… since his duty to any subject is moral constraint by that subject's will" (6:442). Korsgaard interprets this to mean only beings with a "legislative will"—the capacity for self-given universal law—can be sources of obligation. Animals lack this capacity and thus cannot obligate us. However, she argues this does not preclude us from willing laws that protect them: "The question, then, is whether we human beings ever find it necessary, on rational reflection, to will laws whose protection extends to the other animals" (Korsgaard, 96). Her affirmative answer hinges on the Formula of Humanity. To treat humanity as an end-in-itself, she contends, requires conferring value not only on our rational nature but also on our *animal nature* and its "natural good"—the ends we pursue as embodied, sentient beings. Since non-human animals also possess a natural good as "organic system[s] that matter to [themselves]" (Korsgaard, 102-3), consistency demands we extend this valuation to them, generating direct duties.

Allen Wood (1998) presents a second major argument for direct duties, one that seeks to expand Kant's "logocentrism"—the doctrine that rational nature is the source or ground of all value—beyond the confines of Kant's "personification principle." While logocentrism posits rational nature as the sole unconditional good, the personification principle insists this nature "has a moral claim on us only in the person of a being who actually possesses it" (Wood, 193). The latter principle leads directly to Kant's denial of duties to animals, as obligation requires a constraining will (6:442). Wood retains logocentrism but jettisons the personification principle. He argues that respect for rational nature as an end-in-itself logically requires us "to behave with respect toward nonrational beings if they bear the right relations to rational nature," such as having potential for it, parts of it, or "necessary conditions of it" (Wood, 197). Thus, entities displaying "fragments" or preconditions of rationality—like desire, a capacity for pleasure and pain, or preferences—become indirect objects of respect. To treat them with contempt is, on this view, to show contempt for the infrastructure of rational nature itself.[90]

Despite their novelty, both Korsgaard's and Wood's revisionist approaches have serious tensions. Each must navigate the core Kantian principle that obligation is intrinsically tied to a constraining will, seeking to justify direct moral claims for beings that lack legislative agency. This raises a parallel concern about the scope and grounding of the duty: Korsgaard's argument, by rooting obligation in the shared possession of a "natural good," must explain why valuing our own animal nature necessarily extends to the independent animal nature of others, and it risks expanding moral concern to any entity with a functional good. Wood's logocentric strategy, conversely, faces the challenge of distinguishing "fragments" of rationality from mere aspects of animality in a non-arbitrary way. Together, these tensions highlight the difficulty of deriving a robust, direct duty to non-human nature without either weakening the link to rational agency or inviting an unmanageable breadth of moral concern. It is also highly questionable how both can derive duties to higher-order, non-animal agents, such as ecosystems.

In contrast to *direct*-duty revisionism, Onora O'Neill (1997) defends an interpretation of Kant's *indirect* duty framework, arguing it provides a powerful, non-speciesist basis for environmental concern without attributing rights or unconditional worth to non-humans. She accepts Kant's "necessary anthropocentrism"—the view that

[90] Holly Wilson (2008) is close to Wood's view, in that she is an indirect-duty theorist who expands the scope of moral consideration by analogical comparisons to humanity.

only beings with a legislative will can be moral agents and subjects of obligation—and his concomitant "logocentrism," which locates unconditional value solely in rational nature. For O'Neill, this does not justify "unjustified preference for the human species" (speciesism) (1997), but it does mean that moral claims regarding animals must be indirect, derived from duties owed to persons. She roots this duty in moral self-perfection: caring for animal welfare is integral to one's own improvement. (1998, 223). This indirect approach, she argues, avoids the pitfalls of analogical arguments for animal rights, which weaken "as analogies with humans weaken" (1998, 225). Furthermore, by not attributing rights to individual animals, a Kantian ethic is, in her view, more hospitable to holistic environmental goals, such as preserving biodiversity or habitats, because it can prioritize systemic goods without confronting inviolable individual claims. O'Neill thus positions Kantian indirect duties not as a weak compromise, but as a theoretically stable and environmentally capacious foundation for obligation.

Like O'Neill, Paul Guyer (1993) develops a sophisticated indirect-duty account. He does this by anchoring moral concern for nature in Kant's aesthetics and the imperfect duty of moral self-perfection. He interprets Kant's claim that "the beautiful is the symbol of the morally good" (5:353) to mean that our disinterested appreciation of natural beauty cultivates dispositions favorable to morality. This generates a duty toward ourselves to appreciate, and thus not wantonly destroy, natural beauty as a means of cultivating the feelings requisite for moral virtue. Yet while Kant classifies moral self-perfection as an imperfect duty, the prohibitions against destroying beauty or being cruel to animals appear in a discussion of perfect duties to oneself. Guyer addresses this by suggesting that such actions do not merely miss a chance for self-improvement but actively decrease one's own moral perfection by weakening a morally useful disposition. Therefore, they are prohibited with a force akin to a perfect duty. Guyer thus argues that Kant's framework imposes strict restrictions on the treatment of nature, condemning malevolent destruction and endorsing positive conservation as meritorious. He concludes this indirect approach, which views humanity as the moral "ultimate end of nature," is the most promising route for a Kantian environmental ethic, as it imposes robust obligations without attributing direct duties to non-persons.

Advancing and systematizing the indirect-duty tradition, Tony Svoboda (2012, 2015) provides the most promising Kantian account, rigorously grounding obligations to nature in the imperfect duty to increase one's own moral perfection. Against the traditional interpretation—which reduces Kant's view to a weak, contingent claim about psychological desensitization—Svoboda argues that Kant's position "strictly prohibits animal cruelty and wanton destruction of plant life" (2015, 62). These actions are "opposed to a human being's duty to himself" (6:443) because such actions actively violate the maxim commanded by the duty of moral self-perfection. While kind actions or aesthetic appreciation are meritorious ways to cultivate virtuous dispositions, causing unnecessary harm is prohibited because it weakens those very dispositions, thereby decreasing one's moral perfection. Svoboda thus provides a clear, actionable criterion—unnecessary harm—and a sophisticated framework that distinguishes between the prescriptive (meritorious cultivation) and proscriptive (blameworthy violation) dimensions of an imperfect duty. By doing so, he elevates the indirect approach from a theory of useful sentiment to a robust deontic structure with firm restrictions on how nature may be treated, while remaining faithful to Kant's logocentrism and the personification principle.

The indirect-duty tradition, culminating in Svoboda, successfully establishes that a Kantian ethic prohibits cruelty and destruction of nature, promoting instead a benevolent regard for non-human nature as integral to the project of moral self-perfection. However, when assessed as the foundation for a comprehensive environmental ethic—one tasked with addressing systemic collapse, biodiversity loss, and climate change—this framework reveals inherent limitations of scope and grounding. Even O'Neill, who argues that lacking direct duties to animals leaves room for duties toward a functional ecosystem, nevertheless leaves unspecified how a Kantian framework might actually generate such duties toward the ecosystem. This is because the only options for an indirect duties approach seems to be the moral quality of individual interactions with discrete natural entities (animals, flora). This focus, while generating duties regarding those entities, makes it difficult to derive a duty to preserve the functional integrity of ecosystems as a whole—the complex, often non-sentient processes that constitute environmental resilience. Consequently, the prohibition on 'unnecessary harm' and the promotion of benevolent behavior are insufficient.

Consequently, the most sophisticated indirect duties approaches posit that the ultimate ground of obligation remains the realization of the moral self. This reflects an individualist virtue ethics within Kantianism, where nature is valued as the essential context for cultivating and expressing the virtuous dispositions of the lone autonomous agent. This leaves unaddressed the question of whether we also have a duty regarding nature that stems not from its role in self-cultivation, but from its status as world-constitutor. A truly Kantian environmental ethic must reflect the bi-constitutive logic of practical reason itself. The fundamental Idea of reason—autonomy—does not merely constitute an autonomous individual in isolation, but simultaneously aims to actualize a moral world, a nature purposively aligned with freedom. In this framework, the autonomous subject and its environment are united in the project of reason. To cultivate oneself as a moral being is necessarily to contribute to the realization of a world fit for morality, and vice versa. The indirect model explains why destroying individual natural entities corrupts the agent, but it does not ground its argument on the premise that destroying the integrity of nature unmakes the very world that practical reason is commanded to bring about. This suggests the need for an argument that grounds our duty not in the needs of a morally perfecting self, but in the autonomous will's necessary end: the moral world. Specifically, it would be a duty to preserve ecological resilience as the indispensable way in which the kingdom of nature can be unified with the kingdom of ends in the highest good.

Thus, Kantian environmental ethics reveals several recurring challenges. The traditional psychological interpretation establishes only a tenuous, contingent link between morality and nature. The more ambitious direct-duty approaches, in their efforts to ground moral standing in traits like "natural good" or "fragments of rationality," must make significant and textually strained revisions to Kant's framework. The indirect-duty models, while providing the most faithful and robust account of our duties regarding discrete natural entities, ultimately anchor obligation in the project of individual self-perfection. This leaves the moral status of systemic ecological functions under-theorized and does not fully capture the way rational autonomy, through its necessary end (the highest good), is inherently a world-making project. What remains to be developed is an account that grounds a categorical, systemic duty to preserve nature not in analogies to human traits, nor solely in the cultivation of individual virtue, but in the structure of the moral world itself—an account in which the cultivation of the autonomous subject and the realization of a purposively unified nature are seen as the same aim of practical reason.

III. Bi-constitution, the highest good

Kant's moral philosophy explains that our ethical practices are generated by the Idea of autonomy. The Idea of autonomy is experienced as a "fact of reason," through which we become conscious of the moral law as binding on us. For such a fact to be possible, we must conceive of ourselves as autonomous from external[91], contingent environmental circumstances, while simultaneously legislating the moral law. This fact unveils our moral bi-constitution, where we constitute both our phenomenal selves as moral, autonomous subjects, while also constituting the world as a moral 'object,' the product or end of our moral intentions. The comprehensive object of our moral striving is the 'highest good.' This 'moral world' is formally defined as the world where our moral

[91] As it turns out, we are already ends-in-ourselves in Kantian ethics: the Idea of Autonomy rather reveals this as a normative goal, to align our phenomenal self that is dominated by inclinations with our moral self.

actions are synced up with the natural world such that each moral action is rewarded with happiness: "if both moral and physical perfection are combined, then we have the best world" (Kant 1978, 140-1). It is a world which brings the "kingdom of nature into exact harmony with the kingdom of ends" (ibid, 159).

Kant defines the kingdom of ends as it appears in the highest good as "a systematic union of rational beings through common objective laws" (4:433)—a community in which all act from duty. But the highest good demands more than virtue alone. It unifies this kingdom with the kingdom of nature such that happiness is proportioned to virtue. Happiness, for Kant, is the "complete well-being and satisfaction with one's condition" (4:393), the "satisfaction of all inclinations" (4:399). The moral law commands us not to eradicate natural inclinations but to subordinate them to duty. In the highest good, the virtuous not only act according to duty but receive the happiness of which they are worthy.

This raises a question: how can we be assured that virtue and happiness will be united in the world? Kant answers through the postulates of God and the immortality of the soul. Kant argues that we must postulate God's[92] existence as the cause capable of proportioning happiness to virtue, and the immortality of the soul to make possible the infinite striving toward perfect virtue. In conjunction, these make the ultimate union of

[92] By claiming we have a duty regarding the purposive world, I am not arguing for a task that Kant reserves for God. God's role in the highest good is to *proportion* happiness to virtue—to judge the moral worth of each agent and distribute happiness accordingly. Our role is to preserve the *conditions* under which such proportioning is possible at all. A degraded ecosystem makes happiness impossible for everyone, virtuous and vicious alike. By preserving nature's purposive unity, we ensure that the world remains capable of meeting permissible ends, so that when virtue is achieved, the harmony it deserves is not blocked by a collapsed world.

virtue and happiness a possibility for rational beings. These postulates are not theoretical proofs but necessary presuppositions of morality's comprehensibility from a practical standpoint. Yet these postulates also suggest that the natural world itself must be capable of supporting the happiness of virtuous persons. A world where soil is depleted, water cycles disrupted, and ecosystems collapsed is a world where even the most virtuous person cannot be happy. The highest good therefore presupposes a natural world that is sufficiently ordered, purposively unified, and capable of sustaining rational beings and their permissible ends over time (I will expand more on this conception of nature in section IV).

To bring about the moral world involves both categorical and hypothetical imperatives. Kant claims that the moral worth of an action lies not in the consequences it brings about or its purpose, but on the maxim on which it is done. This maxim is the principle on which the agent acts (4:399). Categorical imperatives tell us what we ought to do unconditionally, "without reference to a further end" (4:414). They command any action that is "itself good" (4:414). Kant formulates the categorical imperative as requiring that we "act only in accordance with that maxim through which you can at the same time will that it become a universal law" (4:421).

Hypothetical imperatives, in contrast, are conditional. They instruct us that if we will a certain end, we also will the necessary means to it—a relation Kant calls "analytic," since "whoever wills the end also wills…the indispensably necessary means to it that are within his power" (4:417). For example, if we want to be healthy, we must eat nutritious food. While categorical imperatives command unconditionally what we ought to do, hypothetical imperatives guide how we realize those ends. Questions about the social arrangements or economic systems most conducive to a kingdom of ends, for instance, are guided by hypothetical imperatives. They require us to consider basic facts of human and non-human nature: Will people turn to vice without a judicial system? Can they act morally when deprived of adequate food? Only through such instrumental reasoning can we establish the specific material conditions that allow a moral world to take root.

Do we have a categorical or hypothetical obligation to preserve a systematically integrated nature? By a systematically integrated nature, I mean nature's capacity to organize itself as a coherent, self-sustaining whole whose parts—ourselves included (as far as our sensible nature goes)—reciprocally produce and maintain one another. It is tempting to think Kant's sees the purposive unity of nature as a contingent, external precondition, a stable background that must be secured before moral action can be effective. In a chaotic or degraded ecosystem, moral action becomes impossible. Our plans to do moral things simply cannot be realized. For instance, the maxim "I will help those in dire need of food or water if I am able" is universalizable and treats humanity as an end. Yet in an ecologically collapsed world, this becomes unattainable. No matter how good my will, the resources to act on it do not exist. Hence, we might think that Kant proposed a hypothetical imperative to bring about or maintain a self-organizing world, but only so that we may achieve our moral ends.

On the other hand, however, we can conceive of the self-organizing unity of nature not as a precondition, but part of the fabric of the moral world itself. This seems like the most plausible reading of the highest good, given Kant's extensive use of the language of ‘harmony’ and unity between the sensible and moral worlds in the highest good. If purposive unity is a *constitutive* feature of the highest good in this way, then we have an unconditional duty to preserve it as part of our moral vocation. This stands in stark contrast to an instrumental feature of the highest good, which obligates us only conditionally—for example, promoting social harmony simply because it helps achieve personal happiness or avoids punishment. In that instrumental case, once the external incentive disappears, so does the obligation; but with a constitutive feature, the duty remains binding regardless of one’s ulterior ends. Thus, whether nature’s self-organization is constitutive or merely instrumental for morality determines if a Kantian ethic can ground direct duties to the natural world.

Perhaps the best test of whether nature’s purposive unity is instrumental or constitutive for morality is to apply the categorical imperative directly. Kant claims that suicide is categorically irrational: could collective suicide by destroying our own planet be irrational in the same way? The maxim of a suicidal individual is: “from self-love I make it my principle to shorten my life when its longer duration threatens more troubles than it promises agreeableness” (4: 422). Kant rejects this maxim, claiming that it is not universalizable. Kant claims, “a nature whose law it would be to destroy life itself by means of the same feeling whose destination is to impel toward the furtherance of life would contradict itself and would therefore not subsist as nature” (4: 422). Because this cannot be a law of nature, it opposes the “supreme principle of all duty” (4: 422). The contradiction lies in using a principle of self-love to will the destruction of the very ground—one’s own existence—that makes the exercise of that principle possible.

This suicide case illustrates a contradiction in conception. A contradiction in conception occurs when the universalized maxim cannot even be thought as a coherent law of nature. The contradiction is immediate and does not depend on any empirical conditions such as population size, technology, or the availability of alternative resources. Now consider the nature analogue: "From self-love, I make it my principle to degrade the self-organizing unity of nature whenever doing so serves my particular ends." Does this maxim generate a contradiction in conception in the same way?

At first glance, the analogical argument seems promising. Self-love functions to preserve and further the conditions under which one's ends can be satisfied. In the maxim, we satisfy our ends by annihilating rainforests, destroying ecosystems, burning oil, and farming monocultures with chemical fertilizer. Taken to its extreme, the maxim—if taken as universal law—uses self-love as a reason to destroy the ecological ground that make possible the achievement of one's ends, including the end of loving oneself. The continuance of these practices leads to a damaged world where severe weather events destroy homes, landscapes are depressing, and food/water insecurity is rampant. One that loves oneself would struggle to do so in such awful conditions.

However, the analogy fails upon closer inspection. Unlike suicide, where a single act destroys one's own existence immediately and unconditionally, the degradation of nature's self-organization yields its destructive effect only gradually and under specific conditions. If the population is small, or if technology is weak, or if new frontiers (including other planets) are available, agents could degrade local ecosystems and move elsewhere indefinitely without ever undermining the conditions needed to establish their self-love. The alleged contradiction therefore depends on facts about scale, technology, and resource availability—not on the logical structure of the maxim itself. Because one can conceive of a nature where this maxim holds (for instance, a sparsely populated galaxy with infinite habitable planets), the maxim does not generate a contradiction in conception. Collective suicide by gradually destroying the planet is not self-contradictory in precisely the same way individual suicide is.

A contradiction in conception, however, is not the only way for a maxim to fail the categorical imperative. A maxim can also fail because it generates a contradiction in will. A contradiction in will occurs when the universalized maxim can be conceived as a law of nature (so it passes the first test) but cannot be rationally willed because it conflicts with a necessary end of practical reason. Unlike a contradiction in conception, which yields perfect duties (strict duties with no latitude in the way these duties are fulfilled), a contradiction in will yields imperfect duties (duties that allow flexibility in how they are fulfilled).

The maxim, "From self-love, I make it my principle to degrade the self-organizing unity of nature whenever doing so serves my particular ends"—yields a contradiction in will when tested by the categorical imperative. To see why, we must consider the highest good—the complete object of pure practical reason, in which virtue and happiness are united in exact proportion (5:110). The moral law does not merely command that we have good intentions. It commands the realization of a moral world, where the kingdom of ends is brought into harmony with the kingdom of nature. Kant says this complete object is not optional: "The highest good is the necessary highest end of a morally determined will and is a true object of that will" (5:115). Happiness, for Kant, is "the state of a rational being in the world in the whole of whose existence everything goes according to his wish and will, and rests, therefore, on the harmony of nature with his whole end as well as with the essential determining ground of his will" (5:124). Without happiness, the complete good for a finite rational being would be incomplete—not because happiness adds moral worth, but because a world in which the virtuous suffer and the wicked prosper is not a world fully in harmony with the moral law.

Now consider what happens if the maxim of degrading nature's self-organizing unity is universalized. The nomic, productive, and regenerative conditions of the natural world—fertile soil, water cycles, climate stability, biodiversity—would be systematically undermined. The result would be a world in which even the most virtuous person could not reliably satisfy their natural needs for food, shelter, health, and community. Happiness for this agent, as the satisfaction of legitimate inclinations, would be materially impossible, as they will not have "the harmony of nature with his whole end" (5:124). We hence have a contradiction in will: the universalized maxim conflicts with something the rational will, insofar as it is determined by the moral law, necessarily wills. This suggests that a purposive natural world is *constitutive* of the highest good, not merely instrumental to it.

One might object that this argument remains anthropocentric, since it makes nature valuable only in relation to the moral world, which is a necessary end of rational human wills. I accept this characterization, but I deny that it reduces Kant's view to instrumentalism. Nature's purposive unity is not a replaceable means to a separately definable end, but is part of what the highest good *is* for embodied rational beings. The duty to preserve nature is therefore direct—not in the sense that nature is a separate object of moral concern alongside persons, but in the sense that it is not derived from a prior duty to persons as a means to an independently defined end. Rather, the end to preserve nature is internal to the highest good. As such, the duty to an integrated nature stands opposed to indirect-duty accounts, which can only generate obligations regarding nature insofar as they are a means to cultivate our moral character. Those accounts proscribe cruelty to particular animals but lack the resources to demand the preservation of nature as a self-organizing whole, precisely because they do not treat nature's integrity as a constitutive of the highest good.

A related objection presses further: if the duty regarding nature is emerges from a conflict with the necessary end of moral reasoning—the highest good— and the highest good includes the happiness of others, then perhaps the duty regarding nature is really just an indirect duty to persons. After all, we have a direct duty to contribute to the "happiness of others" insofar as we have a direct duty to "further the ends of others" (4:430), and Kant claims that a "maxim that excludes" this happiness "cannot be included as a universal law in one and the same volition" (4:441). On this reading, nature becomes a mere means to fulfilling our duty to promote the happiness of others, and the duty regarding nature collapses into an indirect duty.

This objection, however, rests on an unduly narrow conception of what it means to respect persons as ends in themselves. The Formula of Humanity commands us to treat persons as ends in themselves, and part of what this requires is adopting their permissible ends (i.e., ends compatible with the moral law) as our own. This is what Kant calls the duty of beneficence (6:452). As Kant himself states, the duty of beneficence is "a universal duty of human beings, just because they are to be considered fellowmen, that is, rational beings with needs, united by nature in one dwelling place so that they can help one another" (6: 453). To respect someone as an end in itself is therefore to respect the full exercise of their bi-constitutive agency: both their self-constitution as moral agents (their capacity to legislate the moral law and perfect their own virtue) and their world-constitution (their capacity to realize their permissible ends in a shared moral world—the highest good—which, as Kant makes clear, we all pursue together in virtue of the "one dwelling place" [6:453] that we all occupy).

We have a duty to the completion of one's permissible ends, to adopt their own (permissible) world-constitution as our own. This is not a blanket duty to one's happiness. For Kant says it is irrational to have happiness as one's end. As Kant cautions, happiness is not a concept "abstracted from his instincts and thus derived from the animality in himself"—it is an "idea" shaped by reason, imagination, and social comparison, where "the nature inside of us is not receptive to that" lasting satisfaction (5:430). The "nature within" is not a stable set of inclinations waiting to be met by an accommodating world. It is restless, expansive, and prone to internal conflict. To pursue such happiness is thereby irrational, let alone will another to have such happiness.

Kant holds that happiness, when considered as a rational end rather than a mere psychological state, requires inner order. Without such order, happiness is "arrogance" which "the impartial rational spectator can take no delight in" (G 4:393)—it is unstable, illusory, and undeserving of reason's approval. The moral law alone provides this order: the good will is "the indispensable condition even of the worthiness to be happy" (G 4:393), not by *producing* happiness, but by establishing the formal condition under which happiness can be coherently and rationally pursued. The duty to one another's happiness is therefore not just a duty to their satisfaction of ends, but the satisfaction of their *permissible* ends. Satisfying these ends does not make one virtuous but is consistent with virtue.

The permissible ends of finite rational beings are the ends of embodied social beings: nourishment, shelter, health, and community. To satisfy these ends requires a natural world capable of meeting them.[93] But this raises a further question: is the preservation of such a world merely instrumental to fulfilling our duty of beneficence—a useful means we could in principle replace—or is it constitutive of what beneficence requires for embodied beings in a shared world? If the former, our duty to preserve nature's purposive unity is conditional and indirect. If the latter, it is unconditional and direct: to will the happiness of others just is to will the integrity of the natural world in and through which their happiness is realized.

To see why the instrumental reading fails, we must first understand the kind of beings we are. Kant's account of organic interconnectivity in the 'Critique of Teleological Judgment,' the second part of his *Critique of the Power of Judgment* provides this understanding. There, Kant argues that organized beings are not self-enclosed units but

[93] It should be clear by this statement that not *any* self-organization of nature will do here. Many galaxies are self-organizing, yet do not sustain life let alone the permissible ends of human beings. The purposive unity of nature, therefore, applies more directly to the multi-scalar processes that allow these permissible ends to flourish (see section IV)

stand in relations of reciprocal purposiveness: each part exists "for the sake of the others and on account of the whole" (5:373–374), and nature as a whole must be judged as a "system of ends" in which "everything is an end and reciprocally a means" (5:376). An organism is constituted by its dependence on and contribution to the wider nexus of natural processes—what Kant calls the "external purposiveness" of nature, in which one natural being serves as the end of another (5:425). Human beings, at least in terms of our sensible nature, are no exception. We are embedded in this web of reciprocal dependence, being "always only a link in the chain of natural ends," a "means for the preservation of the purposiveness in the mechanism of the other members" (5:430). To satisfy our natural ends is not a one-sided relationship where we gain something from without—it is a two-way relationship where our own end-satisfaction contributes to sustaining (or disintegrating) the whole which it depends on. In satisfying our ends, we are world-constitutors, not mere world-consumers. In the natural order of things, we are reciprocally a means and an end to the purposive unity of nature.

If we always are links in the chain of ends, then our end-satisfaction always constitutes the chain that, in turn, provides the conditions for that very satisfaction. We cannot avoid such constitution. We are left with two ways in which we can constitute this chain. The first is disintegrative: we will that the subject attains their ends and is happy, yet to do so in a mode that erodes the chain of natural ends that makes end-satisfaction possible for themselves and others. The second is integrative: will that the subject attains their ends and is happy, and to do so in a mode that sustains the chain of natural ends.

Recall what the Formula of Humanity commands. To treat another as an end-in-itself is to adopt their permissible ends as my own—to make their well-being part of my practical project. An end is permissible only if its attainment is consistent with the moral law, which commands us to treat all persons as ends-in-themselves. But I cannot genuinely adopt someone's ends while simultaneously destroying the conditions under which those ends can be realized. If I claim to will your nourishment while degrading the soil and water cycles that make food production possible, I am undermining the very conditions of your future nourishment, as well as the nourishment of others who depend on those same cycles. I cannot genuinely adopt your end while eroding the material basis for both it and the ends of others. An end-satisfaction that degrades the chain of natural ends is therefore incompatible with the commitment to others' ends that the Formula of Humanity demands. To will permissible end-satisfaction is not to will end-satisfaction and then, separately, constrain that it be achieved in a permissible, integrative manner. The integrativeness is internal to what makes end-satisfaction permissible. And since the integrative mode just is the mode that sustains the purposive unity of nature, the Formula of Humanity commands me to preserve that unity.

It would be a mistake to conclude from this that we preserve nature's integrity in order that humans may be happy, as if nature's integrity were a requirement generated by a prior commitment to human welfare and added to it as a means. Integrated nature would be instrumental in that case. But the relation is more intimate than this. Whatever the nebulous concept of happiness may involve, it necessarily includes the satisfaction of embodied needs—for nourishment, shelter, health, and the material conditions of community—that can only be met by a functioning natural world. To will permissible end-satisfaction is therefore at least in part to will that those needs be met, which is to will a certain configuration of nature itself: the alignment of our embodied needs with the natural conditions that satisfy them. Since our bodies are a part of nature, that alignment just is nature in harmony with itself. Conversely, to will the integrity of nature is to will that each link in the chain—including human beings—be sustained. These are not two acts of willing, one serving the other, but one act described at two levels. Neither is the reason for the other. They are the same state of affairs. The duty regarding nature is therefore direct not because nature is an end in itself, but because preserving its self-organizing form is what it means to will the permissible happiness of embodied beings. There is no gap between them for an instrumental relation to occupy.

One might still object that the reason we care about this unified state of affairs is because we care about satisfying human's permissible ends. Nature's integrity matters, on this objection, only because we care about humans. And if humans did not matter, nature's integrity would not matter either. This again sounds like anthropocentrism, and in one sense I concede that it is: the normative ground of the duty is ultimately the moral law's command to promote the permissible ends of persons. But I deny that this reduces my reading to instrumentalism. The standard anthropocentric critique draws two conclusions from the primacy of human welfare that do not follow in Kant's case. First, critics charge that anthropocentrism establishes an arbitrary moral boundary between humans and non-humans. Second, they charge that it deems nature as merely instrumental for our human ends. Kant's anthropocentrism, properly understood, undermines both.

For Kant, the moral boundary between persons and non-persons remains since only rational beings are ends-in-themselves. But this boundary breaks down when we track what the duty to these beings consists of. It is a duty to the satisfaction of their permissible ends as they are integrated in a purposive whole. These ends are nodes within a purposive system of means-ends relations, benefiting from and sustaining the system that sustains them. A duty to the node is a duty to the system, and vice versa. We hence do not have a duty to others' end-satisfaction first and a purposively unified nature second. Both emerge together within the singular duty to other human beings. The boundary between human and non-human holds at the level of who can obligate us. It dissolves at the level of what the duty to their permissible end-satisfaction consists in: a harmony of nature with itself.

If there were no more human beings, would we still have obligations regarding nature?[94] If the Formula of Humanity was our only articulation of the moral law, we might be compelled to concede that we would have no obligations to nature in this scenario. However, the argument from the contradiction in will still pertains. We cannot consistently will the degradation of nature and the highest good. We have a direct, imperfect duty to instead be integrated with the system of nature and promote that integration, while refraining from acts that degrade nature.

---

[94] I am here consciously drawing upon Richard Routley's (1973) 'last man' argument that is supposed to show that western ethics is humanly chauvinistic and unable to address environmental problems.

Granted that I have established that we have a direct duty to the integrity of the natural world, what does that concretely require of us? Does it demand preserving every species, every ecosystem, every natural process? Or can some be sacrificed so long as the whole remains functional? And does it generate any obligations toward non-human agents themselves? These require a substantive account of what nature's "purposive unity" consists in—an account that bridges Kant's philosophy with contemporary ecology. I take up that task in the next section.

IV. The Purposive Unity

Kant laid the philosophical groundwork for a distinctly ecological mode of thought. His conception of a "natural end"—an organism as a being that is "cause and effect of itself"—established a model of nature as a system of reciprocal causation (5:371). For Kant, an organism is to be conceived as if it were a purposive whole whose form (its "Idea") shapes both itself and an environmental niche (see chapter 3). He extended this logic to interspecies relations, noting how "the plant-eating animals exist in order to moderate the excessive growth of the plant kingdom" (5:427). In this way, Kant's analysis provides a formal, proto-ecological schema, where nature must be judged as if it were a community of bi-constituting agents of which we are a part, together arranged in a self-organizing, purposive whole.

This foundational, proto-ecological insight did not immediately translate, for Kant, into a sophisticated ecological theory the likes of which we see today, nor did it translate into an explicit environmental ethics. Lacking the modern sciences of ecology, climatology, and systems theory, as well as awareness of the devastating widespread effects of human activity on core environmental functions, Kant could not identify what, in concrete terms, sustains that purposive unity of nature that his moral theory seems to require. He could not know that the stability of climate, the integrity of nutrient cycles, or the diversity of species are the very conditions of nature's integrated coherence. Therefore, his obligatory postulate did not amount to a set of actionable duties to wetlands, species, forests, or the atmosphere.[95] It was a mere formal feature of the highest good.

Our epistemic situation today is different, due in part to the logic of Kant's proto-ecological framework. We now know that we live in a complex ecological system. Like all natural systems, they are finite and vary in their susceptibility to perturbation and collapse. Modern ecology labels those systems "resilient" that can either resist or adapt to novel disturbances that could potentially collapse the system (see chapter 5). Furthermore, such systems are resilient due to interactions at all scales, ranging from fast, microscopic interactions to global and interplanetary interactions. The natural world is resilient, in other words, because cross-scale interactions—from the microbial to the atmospheric—coordinate to produce a self-organizing whole.

[95] Kant's *Metaphysics of Morals* does, like our ecological sciences today, acknowledge that the earth is finite in which we occupy a "community of the land" where the inhabitant's land is a "part of a certain whole and thus a part to which everyone originally has a right" (6: 352). Yet this spherical limit of the "globus terraqueus" was, for Kant, the empirical condition that compels human coexistence and cosmopolitan right. It was a given fact of the human condition, not yet the scientific object of an integrated ecological whole towards which we have an obligation to preserve.

Nature's purposive unity is not a static, single-scale structure but a dynamic, nested hierarchy of self-organizing systems. As shown in Chapter 5, the theory of panarchy demonstrates how adaptive cycles at different scales interact through 'revolt' (smaller, faster systems pushing larger systems toward criticality) and 'remember' (larger, slower systems providing stability and memory to smaller systems) dynamics. A virus can revolt against a human body, forcing adaptation or collapse; the body's immune system, shaped by slower evolutionary timescales, remembers successful configurations that guide future responses. Similarly, a local forest fire (fast, small scale) may trigger collapse at that scale but enable renewal and innovation at a higher scale, while the regional climate system (slow, large scale) remembers patterns that guide recovery at smaller scales. These cross-scale interactions are not external to the highest good—they are the concrete, dynamic fabric within which human happiness and virtue are to be realized. To degrade any scale—from soil microbes to global climate—is to disrupt the nested ecology that constitutes the purposive unity of nature.

To be happy—at least to have one's embodied ends satisfied—is not to receive a set of discrete goods from an external world. It is for these nested, interdependent systems to function in alignment, such that one's permissible ends and the world that meets them are no longer at odds. The purposive unity worth preserving is therefore not just any system, but the one in which we are already embedded, and whose integrity is the condition of our own embodied flourishing. The purposive unity we are obligated to preserve is hence not a lifeless galaxy that nevertheless dynamically organizes across scales. It is one in which our bodies maintain their integrity through the microbial communities in our guts, which in turn depend on nutrient cycles moving through soils and ecosystems. It is one in which our nourishment draws on food webs, pollination networks, and hydrological cycles that operate at scales far beyond the individual. It is one in which biodiversity—from the genetic variation within a single species to the mosaic of habitats across a continent—functions as the living fabric that stabilizes and enriches these processes. And it is one in which our shelter and community depend on climate stability, itself sustained by the cross-scale interplay of oceanic currents, atmospheric chemistry, and terrestrial carbon sinks. Perhaps most importantly, it is one that has the capacity to adapt, evolve, and transform.

What does this obligation to the purposive unity of nature, then, involve? It seems we must safeguard not just the whole ecosystem, but the parts on various scales that are the 'cause and effect' of themselves and this whole. We hence have obligations to protect individual natural agents insofar as they are crucial nodes in this self-organizing whole. We must assume that the ecological world that is constituted by these agents at all levels is the one we share with them. They constitute the highest good, insofar as their contributions to a self-organizing kingdom of nature is integral to the realization of the moral world.

Are non-human agents hence valuable insofar as they are *instrumental* for realizing the highest good? This would introduce a crude consequentialism into Kantian environmental ethics. In this reconstruction of Kant, a natural agent is not a static individual to be instrumentally used. It is, essentially, defined by its analogical *practice* of self-organization—an ongoing activity of bi-constituting its internal states (e.g., its organs) and its external world (e.g., its functional niche; see chapter 3). Their value does not reside first in an end-state or in an isolated entity. It resides in the specific *form* of this bi-constitutive practice itself.

Some agents engage in practices that seem purely disintegrative: their mode of self- and world constitution degrades, suppresses, and destroys the very conditions that make its self- constitution possible. Think of a virus that kills its host, and, thereby, itself.[96] But what makes a form disintegrative is not what it does considered at a single scale. As we have seen, the form of a natural agent is multi-scalar, including the pattern of its self-constitution, the pattern of its environment-constitution, and, in general, the pattern of how it interacts with the forms of other agents at other scales. Each form is a dynamic form-among-forms. Whether a given form is integrative or disintegrative depends on its full dynamic multi-scalar structure, that is, on how its activity at one scale composes with activity at another, and on how it interpenetrates the forms it is embedded among.

To illustrate this, let us return to the virus example. At the level of one scale, the virus is disintegrative: it kills its host and itself. But we cannot stop the analysis at this scale. We must ask what the hosts themselves were doing at higher and lower scales. Suppose the hosts were an overpopulated deer population, stripped of predators, whose overgrazing was depleting native flora, eroding soils, and collapsing the forms of life that depend on that vegetation. At the scale of the deer, the virus is disintegrative—it unravels the deer's environment-constituting activity. But at the scale of the ecosystem, the virus

[96] The concept of these beings seem to contradict themselves—similar to a contradiction in conception in the categorical imperative. Their self-negation cannot persist as a law of nature.

may be integrative: by reducing the deer population, it relieves pressure on vegetation, restores soil stability, and reopens niches for other forms to constitute themselves and their worlds. The same viral form, considered across several scales, is simultaneously disintegrative and integrative depending on which cross-scale relations we highlight.

The ultimate arbiter of whether a form is integrative or disintegrative is not any single scale—not the organism, not the population, not the ecosystem considered as a top-down whole.[97] The integrity of nature as a whole is not a phenomenon that sits at the highest level, governing the lower levels (e.g., organisms) from above. Nor is it a property of the sum of these lower levels. It is the coherence of the entire multi-scalar architecture, where forms at every scale interlock with forms above, below, and alongside them to sustain the conditions for environment-constitution at all scales. A form is integrative when its activity, traced across the full network of its cross-scale relations, supports the resilience of this architecture (which, of course, must include us and the sustainable satisfaction of our permissible ends). It is disintegrative when its activity unravels that architecture. To properly distinguish the two requires us to follow the form wherever its constitutive activity reaches—upward, downward, and laterally—and to ask whether the pattern it enacts, in concert with the patterns it interpenetrates, sustains or degrades the 'meta-form' that makes any world-constitution possible. The same form can be integrative in one cross-scale configuration and disintegrative in another. There is no

[97] Callicott (1989) defends a form of ecocentrism based off of Aldo Leopold's (1949) land ethic, where moral standing favors ecological wholes over individuals. The view defended here stands diametrically opposed to making a decision between individual organisms or whole, claiming that a resilient, integrated nature is the interdependency of both.

privileged scale at which the verdict is decided in advance. There is only the formal coherence of the whole, evaluated across every scale at once.

Thus, the position is not consequentialist. It does not derive moral value from aggregated beneficial outcomes for sentient beings. It locates the moral standing of a non-human agent in the form of its activity—its multi-scalar, bi-constitutive logic. Our obligation to favor and preserve integrative agency follows directly from the duty to realize the highest good. To undermine integrative agency is to undermine the concrete manifestation of the purposive unity that moral consideration presupposes. We do not, therefore, have intrinsic duties to animals or life in general as static substances—as bearers of moral status merely by virtue of being alive or sentient.[98] We have duties to preserve and promote integrative activities, which entails obligations to all processes that instantiate the resilience of the self-organizing cosmos: geo-chemical cycles, the atmosphere, hydrological systems, microbial communities, and countless others. We are obligated not merely to preserve self-organizing wholes abstractly but to protect their interdependent dynamics, defend their regenerative thresholds, and govern our actions by their logic of reciprocal constraint—curbing emissions, halting deforestation, eliminating pollutants that disrupt metabolic and reproductive processes across species. We must also contain or eliminate any agency whose multi-scalar form is, in its actual context, disintegrative, because its cross-scale activity degrades the meta-form that makes environment-constitution possible.

[98] Sentientists, like Peter Singer (1975) think we have obligations to beings that are capable of suffering. Biocentrists, like Paul Tayler (1986), think that all living organisms possess inherent worth and demand our respect

Since the duty to preserve nature's purposive unity is internal to the duty of beneficence—to willing the permissible happiness of others as part of the highest good—this position obviously prevents the elimination of humans, as perhaps other environmental theories might promote.[99] Although in certain cases the elimination of humans might be best for the ecosystem, this action would not contribute to a moral world. Kant's theory curtails any wide-scale actions that violate the autonomy of humans by treating them as a means to a self-organizing kingdom of nature. Our permissible ends are part of this self-organizing cosmos, and we have a duty (the duty of beneficence) to satisfy them for other human beings. Yet human actions that obviously undermine the possibility of a moral world—whether they be against humans or against the functionality of our ecosystem—would generate direct duties against them.

It should be noted that the obligation we have to the integrity of the ecosystem does not replace, but complements other duties we have for the cultivation of our moral self-perfection.[100] This might include, as Svoboda has pointed out (see section II), indirect duties to be benevolent and to avoid unnecessarily harming animals. In addition to those duties, the duty to functional integrity exists as a direct but imperfect duty that contributes to the realization of the highest good. To educate our judgment in whether certain actions are integrative or disintegrative is also an imperfect duty. Kant himself argues that

[99] For instance, the Deep Ecology movement recommends the reduction in the human population (Naess & Sessions, 1986).

[100] Nor does the duty to the integrity of the ecosystem reduce all values—aesthetic, cultural, spiritual—to the metric of systemic functioning. The Grand Canyon's value is not merely its value in the Colorado River's watershed; it's value also lies within its sublime beauty. Kant's environmental ethics, as Paul Guyer has noted, can include obligations for maintaining nature's beauty that go beyond its role in systemic functioning. Yet Kant's environmental ethics goes deeper: even ugly landscapes contribute to the functioning of the ecosystem, and hence we have an obligation to protect them.

virtue requires "cultivating one's will" and "sharpening one's capacity to judge" (6:387). Willful ignorance or negligence in ecological understanding constitutes a failure of moral self-perfection.

In sum, we can point to a few obligations that a Kantian environmental ethics directly compels us. First, we have a direct, imperfect duty to refrain from actions that directly and knowingly negate integrative, world-constituting agency, such as driving a species to extinction for trivial gain or irretrievably dismantling an ecosystem's capacity for renewal. Concurrently, we bear a direct, imperfect duty to actively promote and cultivate this integrative structure, making the flourishing of nature's purposive form one of our adopted ends. This duty, beneficence, obliges us to incorporate the support of ecological integrity into our maxims, developing the practical wisdom necessary to discern and foster practices that sustain the whole. Thus, Kantian ethics directs us not only to avoid being disintegrative agents but also to strive toward becoming integrative ones, binding our moral vocation to the concrete, active manifestation of purposive unity in the natural world. Both dimensions remain imperfect duties, categorical in their bindingness yet flexible in their execution.

V. Ecosystem Schmecosystem

A crucial objection to my view is that valuing the structural integrity of the ecosystem leads to highly managed, human dominated landscapes that are functional for human moral practices but impoverished in other ways, like biodiversity. This rests on a false empirical presupposition, since in the most sophisticated ecological accounts, biodiversity seems to promote rather than hinder ecosystem flourishing.[101] Yet this does not seem to hit at the heart of the objection. It asks: if we could theoretically replace a living ecosystem with inorganic, human-made machines and chemicals that perform the same functional services, would this framework permit it? Perhaps we can imagine a spaceship that sustains the life of the humans onboard, but replaces all other ecosystem functions by machines that regulate the climate and provide food and water, and shelter. If we could—at least conceptually— do this on a mass, geo-engineering scale, then the self-organization of the cosmos is not constitutive of the highest good. Rather, *any* kind of natural organization is constitutive of the highest good, whether it self-organizes or it is organized by us.

This objection shares a structural similarity with David Enoch's (2006) 'schmagency' objection to metaethical constitutivism. Constitutivism seeks to ground moral norms in the constitutive features of agency. For instance, it might claim that aiming for coherence or pursuing the good is constitutive of being an agent, and from these inescapable aims, substantive norms are derived. Enoch's objection challenges this foundation. He argues that even if certain norms are constitutive of agency, the theory cannot explain why one must value being an agent. He imagines a "schmagent," who is indifferent to these constitutive aims. The constitutivist can say a schmagent is not a 'true' agent, but the schmagent can retort, "So what? Why should I be an agent rather than a

[101] As just one example, see the "multiple-mechanisms' hypothesis (Eisenhauer et al., 2024), which provides empirical evidence that high-diversity plant communities exhibit greater stability through efficient resource use and tighter nutrient cycling.

schmagent?" The objection highlights that constitutivism may fail to provide a reason to engage in agency and be bound by its norms in the first place.

The environmental constitutivist theory that I attribute to Kant claims that a self-organizing cosmos is constitutive of the highest good—happiness in proportion to virtue—which is the necessary object of pure practical reason. The parallel challenge, then, is whether a moral agent can simply 'opt out' of preserving this self-organizing cosmos. Most agents already do opt out, degrading the environment for their own ends. Yet to degrade the ecological systems that make happiness possible—soil, water cycles, climate, biodiversity—is to will away the foundation of the highest good. The agent who does this cannot consistently will the highest good (as every rational being necessarily does) while simultaneously willing a maxim that destroys its constitutive conditions. In this sense, their agency is self-undermining: they free ride on the integrative practices of others who maintain the shared, self-organizing world, while their own will actively dismantles that world. Such free-riding contradicts the necessary end of preserving and promoting nature's purposive unity—an end that is not optional but constitutive of the highest good itself.

Yet it still seems plausible we can 'opt out' of preserving a *self*-organizing ecosystem. Alternatively, we could use technology to constitute a completely artifactual, managed "schmecosystem" that provides just enough environmental regularity for a kingdom of ends to be actualized in the natural world, thereby sidestepping the norm to promote nature's own *self*-organization. The objection forces the question: does the theory bind us to a *self-organizing cosmos* specifically, or merely to *any stable functional backdrop* for our agency, no matter how artificial?

To consider this man-made ecosystem in its full scope ought to include our own bodies, which constitute this ecosystem. We would have to replace the self-organization of our own bodies with artificial machines that provide us with desires and inclinations. When designing these bodies, we have a choice: do we design them to be maximally conducive to moral action, or not? If we do, then morality ceases to exist: we would no longer act from duty or obligation, but from mechanical necessity. Yet if we do not, and we design our bodies to be realistically human-like, then the struggle of morality becomes an arbitrary, curated experience.

Imagine, for instance, an actor who wants to simulate the most profound experience of a moral struggle. They want to simulate staying loyal to a friend while under torture. In this situation, if they betray their simulated 'friend,' they will die. So, the actor creates a hyper-realistic prison cell and hires an assistant to play the interrogator. The actor even takes certain drugs to engineer their own physiological responses to create fear, exhaustion, and pain signals. After the actor undergoes intense interrogation, he succeeds in not betraying his 'friend.' Did he behave morally?

Intuitively, he did not. He was merely acting like a moral person. He could have redesigned the torture chamber at any moment, made the interrogator friendlier, or taken less drugs to simulate pain. The opposition to his moral struggle is stipulated, not genuinely encountered. The problem is that the arena for action is contingent upon the choice of the actor's will, which can, in principle, redesign any feature of the scenario to better suit themselves. But if this is true, any necessity within his will is not a true, binding necessity. The basis of his will is not the moral law, but a contingent artifactual scenario. The actor can change the scenario at any point to be more conducive to the moral law. Morality is turned into a solipsistic exercise, willing against obstacles you yourself have offered for the purpose of studying whether you would behave morally when the time required it.

Analogously, nature needs to be genuinely self-organizing for virtue to be possible, because virtue requires a genuine other—a world with real, independent constraints that are not merely stipulated by our own will. The actor who designs his own torture chamber cannot be genuinely virtuous because the obstacles he overcomes are stipulated, not genuinely encountered. His "virtue" is a performance, not a real achievement. Similarly, a "schmecosystem" designed entirely by us would gamify virtue where we can at any point redesign any obstacle to suit ourselves. Happiness might still be possible, but virtue would collapse into performance. And without virtue, the highest good cannot be realized, because there is no virtue to proportion happiness to. This demonstrates why any purely instrumental or indirect models are insufficient. If nature were merely a tool for achieving human happiness (as in the traditional indirect view), a functionally equivalent artifact could replace it. But this is impossible because virtue requires a genuine, self-organizing other that is not of our own making. The only way to preserve the integrity of the ecosystem is not, therefore, to impose our technological will on it, replacing its self-organizing functions with mechanical designs of our own. It is to respect nature's 'otherness,' its *self*-organization not as instrumental to happiness, but as constitutive of happiness in proportion to virtue itself, and thereby constitutive of the highest good. Only with a genuinely self-organizing natural world—a genuine other—can virtue, and thus the highest good, be conceivable.

The only way to preserve the integrity of the ecosystem is not, hence, to replace the function of the ecosystem with mechanical functions of our own design. It is to respect the self-organization of nature not as instrumental to morality, but as constitutive of the highest good, as a meta-form to harmonize with as we approach, together, the moral world. This does not mean, of course, that we cannot technologically intervene to *aid* in the self-organization of the cosmos. Such aid, on the contrary, might be quite necessary, especially in eliminating invasive species, finding less energy intensive methods of consumption, and using technology to create reusables. The point is, rather, that a duty to a 'schmecosystem' that mimics the functions of the ecosystem—but is entirely man-made—contradicts our will to the highest good. This does not mean that we are obligated to preserve a particular ecosystem (like the way the ecosystem organized itself 500 years ago), but the *form* of such an ecosystem: that it is self-organizing.

VI. Conclusion

This chapter derives two duties towards the natural world within a Kantian framework: to refrain from disintegrating the meta-stability of the natural world and to promote integration. In contrast to the sophisticated indirect-duty tradition of Kant scholarship, this obligation to the self-organizing cosmos is direct, derived directly from the categorical imperative and the necessary conditions it generates for realizing the highest good. The integrity of the cosmos—understood as the integrated totality of all integrative, world-constituting practices—is constitutive of the very field in which our own agency becomes intelligible and which we can achieve happiness for morally permissive ends. Consequently, damaging this integrative structure is not just bad for our character. It is an assault on the necessary end of moral reasoning: a moral world. Our duty is thus owed directly to this form of organization, because our own final end as moral agents is internally and inextricably bound up with its existence. We are not merely cultivating ourselves through nature; we are obligated to the cosmic whole of which our agency is a participatory part.

Bibliography

Callicott, J.B. (1989). *In Defense of the Land Ethic: Essays in Environmental Philosophy*. SUNY Press.

Eisenhauer, N. et al. (2024). The Multiple-Mechanisms Hypothesis of Biodiversity-Stability Relationships. *Basic and Applied Ecology,* 79, 153-166.

Enoch, D. (2006). Agency, schmagency: Why normativity won't come from what is constitutive of action. *Philosophical Review*, 115 (2), 169-198.

Gillroy, J. M. (1998). Kantian Ethics and Environmental Policy Argument: Autonomy, Ecosystem Integrity, and our Duties to Nature. *Ethics and the Environment*, 3(2), 131-158.

Guyer, P. (1993). Duties Regarding Nature. In *Kant and the Experience of Freedom* (304-334). Cambridge University Press.

Hoff, C. (1983). Kant's Invidious Humanism. *Environmental Ethics*, 5(1), 63-70.

Kant, I. (1978). *Lectures on Philosophical Theology.* Cornell University Press.

Kant, I. (1998). *Critique of Pure Reason* (P. Guyer & A. Wood, Eds.) Cambridge University Press.

Kant, I. (1998). *The Metaphysics of Morals*. Cambridge University Press.

Kant, I. (2000). *Critique of the Power of Judgment* (P. Guyer, Ed.). Cambridge University Press.

Kant, I. (2015). *Critique of Practical Reason.* Cambridge University Press.

Kant, I. (2019). *Groundwork of the Metaphysics of Morals* (R. Stern, Ed.). Oxford University Press.

Korsgaard, C. M. (2018). *Fellow Creatures: Our obligations to the other animals*. Oxford University Press.

Leopold, A. (1949). *A Sand County Almanac*. Oxford University Press.

Naess, A., & Sessions, G. (1986). The Basic Principles of Deep Ecology. *The Trumpeter*, 3(4), 14.

O'Neill, O. (1997). Environmental Values, Anthropocentrism and Speciesism. *Environmental Values,* 6(2), 127-142.

Regan, T. (1983). *The Case for Animal Rights*. University of California Press.

Routley, R. (1973). Is There a Need for a New, an Environmental, Ethic? *Proceedings of the XVth World Congress of Philosophy*, 1, 205-210.

Skidmore, J. (2001). Duties to Animals: The Failure of Kant's Moral Theory. *The Journal of Value Inquiry*, 35, 541-559.

Singer, P. (1975). *Animal Liberation*. HarperCollins.

Svoboda, T. (2012). Duties Regarding Nature: A Kantian Approach to Environmental Ethics. *The Kant Yearbook* 4, 143-163.

Svoboda, T. (2015). *Duties Regarding Nature: A Kantian Environmental Ethic*. Routledge.

Taylor, P. (1986). *Respect for Nature: A Theory of Environmental Ethics.* Princeton University Press.

Wilson, H. (2008) The Green Kant: Kant's Treatment of Animals. In P. Pojman & L. Pojman (Eds.), *Environmental Ethics: Readings in Theory and Application* (87-96). Cengage Learning.

Wood, A. (1998). Kant on Duties Regarding Nonrational Nature. *Supplement to the Proceedings of The Aristotelian Society*, 72 (195), 189-228.