5

THE GREEDINESS OF SCALES

[T]here is throughout nature something mocking, something that leads us on, but arrives nowhere; keeps no faith with us. All promise outruns the performance. We live in a system of approximations . . . We are encamped in nature, not domesticated.

Emerson

(i)

The central objective of this essay is to highlight an important form of computational architecture called multiscalar modeling. However, I’ll begin with some familiar philosophical considerations to which such structures will prove directly relevant. In his celebrated “On What There Is,” W. V. Quine claims that whether a specific individual (e.g. Susie) embraces the existence of a group of objects of type $\varphi$ (very big sets, for example) should be adjudicated upon the basis of whether the overall “theory of the world” that Susie accepts contains quantificational claims of the form $(\exists x) (\varphi \land \psi)$ (= some $\varphi$’s are $\psi$) where $\psi$ captures additional assumptions that Susie makes about the $\varphi$. In everyday speech, however, we commonly hide the crucial $\varphi$ component within various forms of sorted quantifiers (such as “for some set” or “someone”). But it is easy, Quine advises, to regiment Susie’s discourse into a logically connected whole by employing a single species of unsorted quantifier, wherein her collected ontological posits become syntactically manifest. Every student of philosophy learns how to make the appropriate adjustments in an elementary logic course—one simply supplies each $\psi$ with an extension that is null outside the bounds of $\varphi$ and then inserts an appropriate explicit mention of $\varphi$. Susie’s ontological choices can then be read off her newly regimented syntax. The resulting voluntarism with respect to ontological

2 In W.V.O. Quine, From a Logical Point of View (Cambridge: Harvard University Press, 1980).
issues that this approach facilitates offers many advantages and has been widely accepted.3

Rather than calling this process regimentation, let’s dub it amalgamation, for the technique brings previously isolated subdomains into interactive community.

How should we do this? The standard answer just reviewed: supply each \( \psi \) with an extension \( \psi^* \) that is null outside the bounds of \( \varphi \), and then frame the unsorted claim (\( 3x \) \(( \varphi \& \psi^*)\)). But there’s a potential glitch. Suppose that the sub-domains we propose to amalgamate in this manner appear to talk of common objects \( a_1, a_2, \ldots \) of which inconsistent \( \psi \) claims have already been made within the localized discourses. If so, we won’t be able to produce the required \( \psi^* \)’s without engaging in some more drastic form of reformulation.

In fact, this amalgamation difficulty arises quite frequently within the discourses of science. In this essay, I’ll discuss a cluster of obstacles that trace to the employment of differential equation models. I’ll call these puzzles of scale because they revolve around the fractured and unequal ways in which physical information is registered at different choices of characteristic size. Our primary question will be: how should we think of Quine’s recipe for discerning an “ontology” in such a setting?

Illustrated are some of the characteristic size scales pertinent to a steel beam. At the highest levels, steel stretches and compresses by a simple set of Hookean rules down to about 10 \( \mu \text{m} \), at which point the grain structure within steel becomes important and its components whose stretch and compress according to a more complicated set of rules than larger-scale steel. In turn, each of these component grains contains a number of laminate layers which rub against one another in complicated

---

3 Quine’s recommendations have struck most of us as a great improvement over earlier non-voluntarist approaches to ontology in which Susie finds herself saddled with commitments to large cardinals or unicorns simply because Bertrand Russell or A. A. Meinong have philosophically opined that she must, no matter how earnestly she strived to avoid such entrapments.
ways. And on it goes until we reach the tiny crystal lattices of the molecular level, whose orderly patterns are interrupted by higher-scale irregularities called dislocations.

Let us now consider the centrality of differential equations within physics. Their requirements regulate behaviors nominally occurring at an infinitesimal level. Such specifications are often reached by scaling higher-level behaviors downward until a simpler infinitesimal level is reached (this approach is codified in a familiar slogan: “physics is simpler in the small”). But with steel we clearly overreach in these extrapolations—its behaviors stop scaling in a regular manner at about 10 μm. As illustrated in the diagram, small sections of steel behave almost identically at all length scales above this level, but below this cutoff the component grain becomes important and spoils the regular scaling assumptions that previously apply. To capture behaviors on a smaller scale accurately, we must model its grain in a more elaborate, laminate-based manner. Some useful jargon in these respects: the smallest length at which scale scaling symmetries become valid is called an RVE (= Representative Volume Element). A full understanding of steel requires that a substantial hierarchy of these successive RVE sub-models be examined, and today’s scientists know quite a lot about what happens on each of these levels. But problems of data amalgamation prevent practitioners from profiting from this collective knowledge in a straightforward way, due to the problem of the greediness of scales.

What is this difficulty? Each RVE scale-focused modeling will utilize differential equations, and their descriptive demands inherently reach down to the infinitesimal level. But these demands often differ. The result: under amalgamation direct descriptive conflicts will arise in the same vocabulary with respect to the properties that steel displays on small-scale levels.

Here’s an example. At low scale levels, crystalline materials consist of blocks of perfect lattice glued together along randomly oriented boundaries (such structures form when small crystals cool around different nucleating centers within a melt). Due to the scrambled orientations of these blocks, RVE behaviors above the level of these conglomerations will generally be isotropic (= the material responds by the same rules in whatever direction it is pulled). This simple form of higher-scale response supports a modeling in which the stretching and compression behaviors of largish hunks of metal are governed by two simple parameters (Young’s modulus E and the shear modulus μ). But the tiny slivers of crystal within these conglomerates will not stretch and compress in this simple manner, and RVE modelings appropriate to these tiny structures require five or six elastic moduli to capture their anisotropic behaviors. None of this would seem surprising,
except that the differential equations appropriate to these two levels of sub-model require that their parochial rules for stretching and compression reach extend to straining to a zero length scale (we’ll examine whether this presumption can be easily evaded soon). But these divergent extensions generate an immediate syntactic disharmony. A differential equation model must hog all of the lower-size scales available to reach the infinitesimal level at which these equations articulate their stipulations. The unhappy result is a collection of inconsistent claims about the very same small parts of our steel beam. Accordingly, two scientists, each working with good models of steel appropriate to a selected scale level, cannot blithely feed the combined results of their researches into a computer, for it will only generate inferential gibberish due to the syntactic inconsistencies within the lower-scale ranges in which their differential equation requirements overlap. This state of affairs presents exactly the difficulties of amalgamation that we posed with respect to a Quinean perspective on “ontology” at the beginning of this essay.

But this inconsistency of sub-model problem is worse than a mere philosophical puzzle, for it reflects substantive technical obstacles that can hinder modeling success with respect to complex materials. Indeed, many substantive advances in computer simulation within the past twenty years trace to sophisticated policies for evading these descriptive clashes, some of which we shall survey here. The task of integrating science’s various forms of localized modeling knowledge in a successful fashion has been evocatively labeled “the tyranny of scales problem” by J. T. Oden’s research group. Here is an extract from the technical report in which this phrase is introduced:

---

As I will remark later, our greediness problem is closely allied to the celebrated “two tables” debate outlined in Appendix I to Essay 1. It is amusing to view Wilfrid Sellars’ notorious “pink ice cube argument” in this light (Science, Perception and Reality (London: Routledge and Kegan Paul, 1966), p. 26):

*The manifest ice cube presents itself to us as something which is pink through and through, as a pink continuum, all the regions of which, however small, are pink. It presents itself to us as ultimately homogeneous; and an ice cube variegated in color is, though not homogeneous in its specific color, “ultimately homogeneous,” in the sense to which I am calling attention, with respect to the generic trait of being colored.*

“Manifest” is Sellars’ term for “a plain man’s conception,” and he argues that our ordinary concept of pink demands that its applications must scale continuously downward to an infinitesimal level, in contrast to the story provided within “the scientific image.” Most commentators (rightly) reject this a priori requirement as implausible and unmotivated. But observe that a formal analog of Sellars’ “clash between images” does befit the scale-hogging requirements of “mass” and “stress” within standard continuum physics; to reach the differential equations level they must scale their applications downward in exactly the manner that Sellars alleges.
In many ways, all that we know about the physical universe and about the design and functioning of engineering systems has been partitioned according to categories of scale... Virtually all simulation methods known at the beginning of the twenty-first century were valid only for limited ranges of spatial and temporal scales. Those conventional methods, however, cannot cope with physical phenomena operating across large ranges of scale—12 orders of magnitude in time scales, such as in the modeling of protein folding or 10 orders of magnitude in spatial scales, such as in the design of advanced materials. At those ranges, the power of the tyranny of scales renders useless virtually all conventional methods.³

They conclude with this remark:

Confounding matters further, the principal physics governing events often changes with scale, so that the models themselves must change in structure as the ramifications of events pass from one scale to another.

This is a muted way of expressing the inconsistency concerns that I have highlighted as a greediness-of-scales conflict.

A modern multiscalar modeling scheme resolves these disharmonies by allowing each layer of RVE sub-modeling to fill out its descriptive agenda within its own localized and semi-autonomous kingdom (such descriptive patches are sometimes designated as “protectorates” for this reason⁶). But these protectorates can then communicate with one another only through coded messages that are called homogenizations. These division-of-linguistic-labor techniques result in novel forms of explanatory architecture that are worthy of philosophical attention.

(ii)

At this juncture, philosophers may scoff, “Oh, those inconsistencies merely represent the practical setbacks that arise from not modeling the material in a fully bottom-up fashion beginning at the atomic level (or lower). But in principle we can do so and, from that perspective, Oden’s tyranny-of-scales woes prove philosophically irrelevant.” Later in the essay, I’ll complain about these “in principle” excuses in some detail, but, for the time being, let’s simply ignore these false promises and proceed onward, for we shall find that there is more to worry about in these “mere practicalities” than these writers recognize.

Oden’s phrase “principal physics” invokes the notion of the dominant behaviors appearing on a characteristic size scale, and this basic notion will prove central to our discussion.

⁶ For terminological variety, I shall sometimes borrow a term from Bob Batterman (who borrowed it from the physicists): “protectorate,” to signal the semi-self enclosed “worlds” framed by the dominant behaviors available at a characteristic scale of size and time.
Probably the easiest (if not the most general) manner of understanding “capturing dominant behavior upon a scale size” is to consider the natural energy cascades or capacities for transmitting coherent work that characterize a material. Let us bang on our steel beam with a hammer, thereby injecting strong pulses of macroscopic energy into the steel. If nothing else interferes, traveling waves will spread out across the beam in both directions. These waves will retain a coherent structure in the sense that its parts will display an evolving shape that strongly reflects the character of the original hammering (after awhile, dispersive effects will break down this coherence, but we’ll return to such considerations later). If no energy degradation occurs along the way, a traveling wave that reaches the far end of the beam will retain the same capacity to move its environment in an assigned direction as did the parent pulse when it was first injected into the beam (in the physicists’ jargon, the transmission of work capacity inserted by the hammering will be perfect\(^7\)). However, if we instead pin our beam on both ends in a manner in which a violin string is held fixed by its bridge and nut, an advantageous new form of descriptive opportunity arises due to the way in which our traveling waves are continually reflected back into the interior by the endpoints. After a certain initial relaxation time, these reflected waves will often coalesce into standing wave structures, which represent coordinated-across-the-entire-beam patterns of movement that wiggle vertically independently of one another for long spans of time.

\(^7\) Observe that the rigidity constraints invoked in the two main examples of Essay 7 trade upon this absence of significant degradation in work capacity.
Traveling waves, in contrast, move horizontally and are more localized in scale. If our beam is thin, perfectly symmetrical, and homogeneous, these standing waves will appear as the familiar sine wave patterns pictured, familiar from the elementary tonal analysis of a violin string (also as a repeated example within this book!). When the beam is thicker, less homogeneous but still symmetrical, less familiar forms of standing wave pattern $\phi_i$ will appear, which mathematicians call the eigenfunction modes of the system in question. These emergent $\phi_i$ modes act as energy traps in the sense that their individual wigglings preserve the same packets of energy over appreciable spans of time. The tonalities of string instruments and woodwinds generally trace to the manner in which their sine wave “standing wave” components harmonize pleasingly with one another. The comparable patterns within a normal steel beam possess the same energy trapping capacities, but these traits sound more clankly than musical when excited.

When a system’s dominant behaviors factor into independent behavioral modes like $\phi_i$, through endpoint-induced coordination, a great opportunity for descriptive simplification opens before us, because our target’s large-scale behaviors can be neatly captured within a much smaller range of variables (the $\phi_i$) than the position coordinates ($x$) we must still employ for its traveling wave behaviors. Locating these advantageous descriptive opportunities comprises the central strategy behind (generalized) Fourier analysis and proved central within many of the greatest breakthroughs in nineteenth-century physics (Kelvin called it a “great mathematical poem” for that reason). See Essay 9 for more on this.

Scientists generally associate these descriptive opportunities with a so-called characteristic scale length $\Delta L$. Once standing wave behaviors become prominent within our rod, an appropriate assignment of $\Delta L$ encompasses the entire extent of the beam, because the movements of every local portion have become locked into coordinated correlation with every other section (a Fourier $\phi_i$ registration does not decompose a system into smaller spatial pieces, but merely into energetically isolated components). On the other hand, within behavioral regimes where traveling waves remain dominant, the characteristic

---

8 The $\phi_i$ are called the spectrum of the problem, and our standing wave modes are called eigenfunctions of the differential equation modeling (or, more properly, its eigenfunction movements, if we undo the separation of variables factorization that uncovers eigenfunctions in the proper sense of the term).

9 From W. Thomson (= Lord Kelvin) and P. G. Tait, Elements of Natural Philosophy, Vol. 1 (Cambridge: Cambridge University Press, 2010), p. 28:

Fourier’s Theorem is not only one of the most beautiful results of modern analysis, but it is said to furnish an indispensable instrument in the treatment of nearly every recondite question in modern physics.

10 Associated with these characteristic lengths are the natural time measures associated with each $\phi_i$ mode (its period of vibration, the relaxation time that must pass before the $\phi_i$ pattern becomes strongly apparent, and the lifetime over which its dominance lasts). In the case of a violin string, these three characteristic times ask: What is the frequency of the $\phi_i$ overtone? How long must we wait for the initial transients to die away? How long will the $\phi_i$ overtone persist?

11 The reaction-stablized-by-diffusion manner in which so-called solitons emerge as dominant objects within a fiber-optic cable is quite instructive, for these gizmos do not transport internal energy in the simple manner of a regular elastic traveling wave but sustain themselves as coordinated alignments of voltages. We must continually pump energy into the system to render these patterns persistent as the soliton merrily whisks along.
length $\Delta L$ is smaller, determined by the breadth of the region in which the largest part of the disturbance occurs (we typically cut off any longer ranging tails).

However, these standing wave patterns must be regarded as dominant structures only because insidious factors such as friction inevitably operate to corrode their behavioral prominence—viz. our beam will stop wiggling once friction or small nonlinearities drain away all of the energies captured within its $\phi_i$. And this decline generalizes; dominant behaviors rarely stay dominant forever. The kingdoms of mighty potentates crumble into rubble. Placid Californians turn murderous when a hot Santa Ana blows in from the desert. Prevailing light patterns lapse into feeble sub-dominance when a Stokes line is crossed.

Relentlessly of all, all-conquering friction grinds every coordinated motion into incoherent heat in the fullness of time. Nonetheless, from the viewpoint of descriptive effectiveness, we want to locate the perfected dominant behavior patterns first, before we turn to the disruptive factors that eventually spoil their transitory triumphs. Devising layered modeling policies that take these disruptive factors into account without simultaneously inducing a computational explosion of a tyranny-of-scales stripe is a central objective of all multiscale methods.

Within the natural world itself, the persistence of dominant behavior structures is crucial to the interactions between physical systems, despite these imperfections. A mayfly (vide Essay 9) catches its food in little river inlets that trap flotsam-bearing waters in dominant behavior whirlpools. Wine glasses resonate dangerously to the chief Fourier components of a piercing soprano. Speculators can plan their stock purchases through factoring Wall Street behaviors into dominant behavior trends and sub-dominant behavior noise. Reliable patterns of interaction between physical systems often rest upon fundamental issues of coherent control (e.g. how ably can the work capacity instilled by hammering be transmitted along the length of a steel beam without losing its original directional capacity?). In a world containing perfectly rigid bars, none of the injected pulse’s original coherence would be lost in the transport: one can hammer on a target as effectively employing the rod as intermediary as if we directly applied our hammer to it. But the assumption of perfect transmission is

---

12 This is a strongly Duhemian lesson from Essay 4; it is also supported by the “volcano” considerations of Essay 6.

13 Compare Descartes:

*I shall observe that while a stone cannot pass to another place in one and the same moment, because it is a body, yet a force similar to that which moves the stone is communicated exactly instantaneously as it passes unencumbered from one object to
merely a “dominant behavior” characterization of our rod; in real life, some minute portion of our initial effort will disperse into less useful forms of energetic effect that appear as dominant behavior alterations upon a smaller size scale (for example, the minute plastic distortions that eventually ruin a steel beam’s ability to regain its macroscopic shape still represent significant patterns of organized recrystallization when considered at the scale of the component grain). Accordingly, most materials display characteristic varieties of energetic cascade along which effects applied upon a large-size scale trickle downward to induce locally significant alterations upon smaller RVE scales. When enough of these small-scale effects accumulate as “damage,” the material’s original higher-scale behavior patterns can shift significantly, as occurs when we run a large locomotive repeatedly over a steel rail. In fact, the particular multiscalar modeling we shall study patterns its computational architecture on the natural energetic cascade found within a piece of railroad rail.

I again stress the fact that these characteristic lengths $\Delta L$ usually depend upon the manner in which the target system has been confined by its natural environmental boundaries, such as the pinned endpoints in the case of our beam and string. The longer characteristic lengths of standing wave behaviors stem from the directive orders

another... For instance, if I move one end of a stick of whatever length, I easily understand that the power by which that part of the stick is moved necessarily moves also all its other parts at the same moment, because the force passes unencumbered and is not imprisoned in any body; e.g., a stone, which bears it along. (“Rules for the Direction of the Mind” in The Philosophical Writings of Descartes, Vol. 1 (Cambridge: Cambridge University Press, 1985), p. 33)
supplied at the confining endpoints, which induce an across-the-entire-system coordination in the manner of a galley master in a slave ship. When we first begin to bow our violin or hammer on our steel beam, we inject a sequence of traveling wave disturbances that move towards the endpoints, where they become reflected backwards (and upside down) back into the interior of the medium. After many reflections of this character, our beam (often) begins to wiggle almost exclusively in a vertical direction, and the characteristic time $\Delta T_r$ required for this behavioral reorientation comprises the relaxation time mentioned above. But it is the confining endpoints that supply the chief mechanism that supports the convenient computational opportunities offered by $\varphi$-variable description.

Aside: philosophers often talk loosely of the “laws that govern wave behavior” in a manner that ignores the crucial role that boundary arrangements (and other forms of side condition) play in inducing the characteristic behaviors under discussion. But it is exactly the side conditions that operate as the culpable taskmasters within our beam’s circumstances. Formal unclarity on these points seems central to the closed-orbit-of-kind-terms misapprehensions criticized in Essay 6.

Of course, we cannot change the macroscopic condition of a large-scale material without altering the states of all of the smaller RVE units it contains. But we can still categorize these adjustments along a scale that ranges from the completely harmonious to the completely destructive. Consider our earlier diagram of what occurs when coherent work, in the form of a compression wave, is transmitted along a steel beam. If the bar implements the behaviors characteristic of upper-scale elasticity perfectly, the wave can travel from one extremity to the other and completely retain its starting capacities for performing coherent work (viz. moving other objects about). In reality, however, upper-scale elasticity is merely a dominant rather than perfect trait, and some of the wave’s original coherence becomes lost as subdominant effects gradually weaken the strength of the pulse. If the underlying metal remains unaltered, the increasing energetic incoherence within our wave will eventually descend into molecular wiggling (i.e. heat) without inflicting any permanent damage upon the intervening layers of RVE structure. When trickle-down processes of this non-damaging character transpire, each intervening RVE structure will shear away from

---

14 Why? The reflections are not due to any physical processes registered within the interior modeling equations themselves but stem from the quite distinct binding processes active within the endpoint regions. The fact that conventional “boundary conditions” report on these substantial physical factors in a fashion that is highly condensed in comparison to the interior description I call an unequal policy of data registration.

For accuracy’s sake, we should acknowledge that, within real violin strings, the posited relaxation amongst the traveling waves never fully occurs within the string itself, but only within the comparatively sluggish responses of the attached instrument and air. In this fashion, the attribution of a conventional overtone spectrum to the string represents an interesting form of property transference that we will not attempt to explore here.
its relaxed state configuration temporarily as the pulse passes through, but will regain its unstrained configuration soon thereafter. In viscoelastic materials, this recovery can take an appreciable length of time. Despite these delays, no permanent damage on a higher RVE scale will have occurred, and all loss in wave coherence will eventually show up as an increased temperature. In such circumstances, a beam’s lower-scale behaviors remain in harmony with the dominant activities we usually witness on a macroscopic scale.\(^\text{15}\)

However, under stronger blows, the intervening RVE units will distort so severely that they are pushed into altered crystal formations, from which they cannot fully recover. Exceeding their normal energetic barriers, they will partially “forget” their preferences for a rest state configuration. We then say that plastic damage has occurred in the material and that it no longer qualifies as “perfectly elastic.” When energetic cascades of this damaging character occur, the responses at intermediate RVE levels will have fallen out of harmony with the perfect elasticity that we normally expect within a bar of steel. A key source of a multiscalar scheme’s computational efficiency lies in its policy of investigating lower-scale behaviors in detail only as the prospect of unharmful, intermediate-scale damage threatens.

A material’s upper-scale toughness—its resistance to fracture—supplies an important intermediate case. Materials like steel contain a large number of so-called dislocations.

\(^{15}\) When traveling waves eventually congeal into standing wave patterns, the cooperative harmony between scales still requires that the microscopic steel crystals shear vertically with complete elastic recovery. In that manner, no capacity for performing coherent work gets lost. But a subtler form of informational degradation occurs because a traveling wave “remembers” the directionality of its original hammering better than a standing wave. Subtleties in the notion of “pressure” trace to this distinction, as the appendix to Essay 6 observes. We should also recognize that the physical factors that make initial traveling waves eventually spread out into standing wave patterns are rarely registered within the pertinent differential equation model itself. Instead, our mathematical interest in the eigenfunction solutions \(\theta\) is tacitly predicated upon the extraneous empirical assumption that the energetic shift into \(\theta\) modes will occur quickly (in point of fact, this energetic relaxation does not fully transpire within realistic violin strings, whose vibrations maintain a traveling pulse character for appreciable spans of time—see Essay 6 for more on this). Careful students of methodology should notice the silent manner in which important physical processes are here “modeled”—the only linguistic indicator of the complicated factors behind standing wave consolidation is found within the fact that we seek solutions within a restricted class of eigenfunctions. This is an instance of “uneven data registration” at its most extreme—the only “registration” of the relevant processes lies in an inferential practice. Sometimes the proper articulation of an associated inverse problem forces these tacit physical presumptions into brighter daylight.
lines dwelling at an RVE size about ten to twenty times larger than the bond lengths of the molecular lattices in which these structures dwell. From a lower-scale point of view, these “lines” appear as insignificant irregularities within the crystal lattice, but they collect together into line-like structures at a higher RVE scale. Afflicted with a large-scale blow such as our hammering, these lines display a remarkable capacity to move as a coherent group across the lattices in which they live. Accordingly, some of the work we inject into the metal through our hammering pushes these tiny line irregularities across their matrix backgrounds and does not immediately convert to heat. From a macroscopic point of view, such adjustments prove a Very Good Thing because their easy-to-achieve movements shield the underlying molecular bonds from the shearing distortions they would otherwise experience if the full impetus of the original blow had been allowed to reach their bonding sites directly. The net result is that RVE units containing a plentitude of dislocations generally retain their dominant upper-scale behaviors far longer than they could if the dislocations weren’t there, due to the fact that the dislocations significantly lessen the danger of fracture at the molecular lattice level. As a consequence, a steel containing an abundance of dislocations proves considerably tougher than a steel that lacks them and the perfect iron whiskers grown in outer space (which contain fewer dislocations) are very brittle. For such reasons doth the village blacksmith repeatedly fold and hammer his iron, without recognizing that the chief objective of his labors is that of installing a large population of dislocations.

Secretly, however, these little dislocation movements represent a form of structural damage on a minute scale—their statistical distribution becomes slightly altered by our hammer pulse. Normally, we factor this “damage” into our upper-scale expectations by assigning the material a macroscopic toughness. After many repeated blows, this microscopic damage gradually accumulates to a point at which it manifests itself as a macroscopic change in behavior, such as a marked decrease in toughness. History-dependent phenomena of this character are called hysteresis and will become central in the multiscalar model we construct later on.

These natural scale-to-scale interconnections are what I had in mind when I wrote of energy cascades earlier. When we press upon a solid material’s exterior in a macroscopic manner (e.g. hammering), the energies injected into the system rarely manifest themselves entirely as heat at the molecular level but instead pass through an elaborate
hierarchy of intermediate dominant behavior RVE structures that trap the infused energy locally for significant periods of time. Lewis F. Richardson’s little verse\textsuperscript{16} on fluid turbulence captures the basic situation effectively:

\begin{quote}
Big whorls have little whorls \\
That feed on their velocity, \\
And little whorls have lesser whorls \\
And so on to viscosity.
\end{quote}

That is, if we stir a liquid rapidly with a paddle, we insert a large packet of energy into its interior. The chief effects of these large-scale manipulations do not directly affect the molecular level immediately, but instead push the fluid around in large, coherently structured whirlpools (Richardson’s “big whorls”). After a certain interval of time, frictional effects generate smaller eddies between the large-scale movements that gradually rob the large whirlpools of their entrapped momentum (“little whorls feed on their velocity”). Only after passing down a long cascade of these scale-centered effects does the coherent energy of our original stirring become converted into the randomly oriented pushes and pulls at the molecular level that we call “heat” (“and so on to viscosity”). If continual paddling keeps the water turbulent, persistent hierarchies of coherent structure will endure in some statistical variety of a steady-state condition, whose precise characterization eludes us to this day.\textsuperscript{17}

In short, our talk of characteristic scales and times arises as the linguistic shadow of the energetic cascades that occur within complex materials. Most forms of multiscale modeling attempt to imitate these natural hierarchies within their computational architectures. This is the lesson that physicists such as Errico Presutti intend when he writes:

\begin{quote}
Each event has its own characteristic scale, and, even though theoretically we improve by successive blow-ups [= examining the phenomenon on a lower characteristic scale], in fact we may be losing the true meaning of the phenomenon.\textsuperscript{18}
\end{quote}

However, his articulation inadvertently suggests that our hierarchies of characteristic scales merely reflect epistemic limitations upon our representational capacities, an


\textsuperscript{17} For an intriguing expression of skepticism, see Arkady Tsinober, \textit{The Essence of Turbulence as a Physical Phenomenon} (Dordrecht: Springer, 2014).

impression that I’ve tried to counteract by stressing the direct correspondence of dominant behaviors to objective issues of energetic transfer and degradation.¹⁹

Technical comment: I can now correct a slight infelicity in the way that I introduced the term “dominant behavior,” wherein I appealed to the simple behavioral rules that can emerge as descriptively effective at a chosen scale level. Here “simplicity” is partially in the eyes of the beholder, and by complicating the upper-scale “rules,” more effects can be characterized as “dominant” that are not so addressed within simpler treatments. For example, I have characterized a well-made steel bar as “dominantly elastic” on a macroscopic scale and treated its energetic losses in the form of heat as a sub-dominant process, to which we only attend when temperature effects become important. Alternatively, we can add a frictional term to our elastic modeling and address simple heat loss on all fours with elastic response in our dominant behavior rules for upper-scale events. Historically, purist top-down modelers such as Pierre Duhem tried to develop all-encompassing macroscopic rules of this sort, which could account for all effects of lower-scale complexity in a single-leveled manner. They preferred this alternative because bottom-up treatments operating on a molecular level are highly prone to significant modeling error and are immediately confronted with tyranny-of-scales intractability. Multiscale techniques avoid this traditional top-down-versus-bottom-up dichotomy by pursuing hybrid policies that incorporate interactive modeling ingredients drawn from a wide range of intermediate RVE scales.²⁰

Metallurgists possess excellent, individualized models for how dominant behaviors operate upon each of the RVE scales encountered within a complex material like steel (in the example ahead, we shall only worry about a few of these). Their chief task, in constructing an overall multiscale scheme, is to fuse these individual sub-models into some form of mutual cooperation. But this is exactly where our greediness-of-scales concerns raise amalgamationist difficulties. Why? Each RVE-focused sub-model requires differential equations for its proper mathematical formulation, but the descriptive ranges of such equations must ipso facto reach to the infinitesimal level. Our greediness-of-scale considerations then indicate that these descriptive demands will typically clash on intervening scales. But then the local conclusions that scientists reach within each localized sub-model can’t be straightforwardly mixed with the results reached on other RVE scales, for fear of lower-scale inconsistency. So how do multiscale modelers profitably combine their various pools of sub-model conclusion without courting inferential catastrophe? Answer: they devise clever strategies for imitating the energetic hierarchies of dominant behavior relationship found in nature. These mathematical techniques are called homogenization methods.

¹⁹ Just because a specific form of behavior proves merely dominant upon a size scale doesn’t render it impotent or unreal. An ongoing truck can kill you just as effectively if its oncoming path displays slight perturbations about a straight-line central trajectory.

²⁰ In this fashion, they have constructed better models of hysteresis effects, whose treatment had proved a particular challenge to Duhem’s single-level endeavors. Later I shall sketch the architecture of one of these models as a typical example of multiscale policy. Further commentary on these dichotomies appears in Essay 4.
But why does any of this present us with philosophical worries? At first glance, we might presume that our greediness problem can be rectified by simply replacing our differential equation models with truncated versions that eschew all of the excessive lower-scale details within each $\Delta L$-centered modeling. The result would be a discretized collection of data posed at the lowest-length scale retaining descriptive validity on a $\Delta L$ basis. Why, after all, do we need to express our $\Delta L$-centered knowledge in the infinitesimal guise of a differential equation? Why not cut off these unwanted extensions and confine our claims to what happens in little blocks at the $\Delta L$ level and above? Mathematically, this proposal recommends that the continuous solutions to our differential equation should be replaced by broken line plots which more accurately encapsulate the discrete data that we actually possess about our material on the relevant RVE level.

A complete response to this query involves mathematical complications that we will delay until Essay 8, but a baseline response can be supplied. The consequences of such a cutoff policy would prove inferentially disastrous, because profitable reasoning pathways must travel through these infinitesimal levels of description in order to reach worthy conclusions. This is a familiar aspect of viable mathematical reasoning—the inferential pathway that connects A to B often runs through unexpected intermediate territory C (a methodological observation that periodically appears within these pages). Let’s go back to our standing waves. The standard eigenfunction technique locates these important entrapment modes $\phi_i$ via a Fourier transform based upon an underlying differential equation modeling and then decomposes our target system’s superimposed complexities into the independent sub-behaviors associated with the $\phi_i$ spectrum. But if we attempt to apply these same transformation techniques to our cutoff data structure,

---

21 Or, more generally, through the Sturm-Liouville maps of Essay 9.
we obtain what is technically known as a *God Awful Mess*. More exactly, rather than obtaining descriptive simplicity, our Fourier technique spews forth the torrent of spectral terms appearing in the lower right-hand corner of the diagram.

Why does this happen? By truncating our data in a cutoff manner, we inadvertently introduce a lot of little corners into what was formerly a smooth curve. Even inserting a single little blip into an erstwhile smooth curve generates great complexities of this sort; a Fourier transform that formerly split into three simple $\phi_i$ components now requires thousands of high-frequency terms to get that little blip to show up correctly within its final summations (we’ll see why in a moment). A fortiori, if we replace a smooth curve by a broken line surrogate, its Fourier transform will go *crazy* simply because of the many little blips we have introduced. In consequence, we find that the perfectly smooth solutions of our over-extended differential equation modeling supply us with an inferential bridge we should cross in finding our way to the most effective policy for describing a steel beam with fixed endpoints. Moral: Throw out the differential equations in your modeling efforts, and you’ll have probably thrown out your most valuable physical conclusions in the process.

As it happens, this particular difficulty can be alleviated through recourse to more sophisticated decompositional techniques such as the Fast Fourier Transform, but the general point remains—the most suitable landscapes for evaluating our descriptive opportunities include out-of-country extension elements as landmark intermediaries. In the situation before us, the infinitesimally positioned differential equations operate in this out-of-country role, a point of view commonly stressed within the applied mathematical literature. Accordingly, popular policies of codifying modeling considerations as differential equations aren’t directly motivated by the ambition of describing the real life behavior of, e.g. a steel beam upon a truly microscopic scale. Instead, the equations primarily serve as the inferential way stations we should visit on route to important conclusions pertaining to much

---

22 “Analytic data” (= information codified in formula-like terms) offer many inferential advantages over the discretized information supplied by a numerical method. Any thorough book on the latter stresses the fact that the analytic conclusions extracted from a differential equation modeling are vital to understanding the computational landscape behind any form of numerical technique. My Fourier transform example is intended as an illustration of the general utilities of analytic data. The “perfect volcanos” of Essay 6 make this same methodological point in a different way. The suggested picture of how differential equations operate is very much Leibniz’s—cf. Essay 3.
The Greediness of Scales

higher scale levels. An important, if insufficiently acknowledged, fact lies at the root of these difficulties, which we might call “the stupidity of unassisted mathematics.” Reviewing our concerns closely, we realize that it is we who are interested in the dominant patterns appearing in our curve, whereas, left to its own devices, mathematics cannot understand what the phrase “dominant behavior” means. For it loves all of its little creatures equally and, to mathematics, any tiny blip is just as important as any other aspect of a curve. With respect to issues of dominant behavior, unassisted mathematics cannot see the forest for the trees. We are forced to trick the subject into supplying us with the central behavioral answers we require in practical science.

How do we get around this problem? Let’s consult the wily electrical engineer Oliver Heaviside for sage advice:

*It is said that every bane has its antidote and some amateur biologists have declared that the antidote is to be found near its bane.*

Our greediness problems have been generated by the fact that our various differential equation models hog all available scale lengths in order to operate properly. So perhaps we can “cure” the ensuing descriptive exaggeration through a second dose of a counteracting over-extension. Here’s an example of how to do this from Heaviside himself. The behavior of a typical electric circuit is best understood as comprised of an initial transient response mixed in with a long-term, steady-state response. The first represents the dominant initial-time behavior and the latter, the dominant long-term behavior. But in real time, the transient response never completely dies away, and the pure steady pattern never completely appears. Heaviside tricks our circuit equations into divulging their long- and short-term secrets by “turning them on” with the help of different “impossibly compressed” forcing conditions.

There are other exaggerated infinity cures that are more familiar. The simplifying capacities of infinite collections are familiar from statistics. Keeping track of a large set of independently acting gamblers ruining their lives within a large network of casinos requires a large number of descriptive variables. But as more gamesters are included in our remit, a

---

24 Respectively, a delta function and a Heaviside unit function convoluted with the signal.
simple dominating pattern emerges from their collective follies that is called a Gaussian (or “bell-shaped curve”) behavior, characterized by two numbers, the mean and variance. For reasons similar to our blip-on-a-smooth-function considerations, these simple numbers don’t mathematically coalesce until we reach an infinite population of gamblers. The asymptotic maneuver of blowing up a gambling population to infinite size acts as a filter that brings out the “good part” (= dominant behavior) of our accumulated data.

Multiscale modeling efforts employ a hierarchy of sub-models linked to an RVE scale that need to exchange information with one another without resorting to naïve data amalgamation. Typically, applied mathematicians obtain these communication links by repeatedly blowing up their sub-models to an infinite population size, keeping the proportions of internal details intact as they go. Such techniques are generically called homogenizations, although they differ considerably in their precise techniques, which need to be tailored to the problem at hand.\textsuperscript{25} For example, the laminate structure of a typical granular composite requires a high level of mathematical sophistication in how we go about this task (it’s not simple “averaging”).

In the discussion to follow, it is important to bear in mind that the central purpose of these peculiar mathematical manipulations is that of imitating, as best we can utilizing mathematics’ overly simple parameters emerge in an infinite population.

\textsuperscript{25} To this day, the underlying operations of some of our most effective asymptotic techniques remain rather mysterious. Why divergent series expansions often supply better information than convergent ones fall into this class, at least for me. Somehow the relevant extraction techniques scramble the underlying information in a more beneficial way, in the manner of a helpful Fourier transform.
punctilious tools, the dominant behavior interactions that we empirically witness within the natural hierarchies of structure found in a complicated solid material. Roughly speaking, a suitable homogenization technique should overcome mathematics’ baseline stupidity by filtering its lower-scale modeling results in a manner that mirrors the physical manner in which one characteristic scale level “sees” the events arising within its lower-scale companions. Accordingly, a homogenization policy is well chosen if it ably apes the physical manner in which relatively simple forms of dominating behavior, characterized by a limited set of descriptive parameters, emerge at higher scales from their large, lower-scale underpinnings. Physicists dub the general phenomenon of complex behaviors that blur into higher-scale simplicity, universality. The emergence of Gaussian parameters from large populations of gamblers is frequently supplied as a paragon exemplar. Citing Gaussian circumstances, Terence Tao explicates this “universality” as follows:

The inability to perform feasible computations on a system with many interacting components is known as the curse of dimensionality. Despite this curse, a remarkable phenomenon often occurs once the number of components becomes large enough: . . . the aggregate properties of the system can mysteriously become predictable again, governed by simple laws of nature. Even more surprising, these macroscopic laws for the overall system are often largely independent of their microscopic counterparts that govern individual components of the system.26

In the simplest circumstances, randomization alone serves as the chief engine of higher-scale simplification. As noted in section (i), perfect crystal lattices are inherently anisotropic (= they respond by different rules to identical pushes and pulls according to their orientation). However, most minerals cool from an original melt through a process in which millions of little crystals gradually enlarge from independent nucleating centers. Such growth patterns get eventually arrested as neighboring fragments begin to press against one another. Because the nucleating centers were originally scattered in an independent manner, the congealed mass forms into an RVE conglomerate built from little chunks of randomly oriented crystal. Under reasonably small stresses, each tiny crystal fragment will retain its ability to return to its original relaxed state and so, after a simple homogenizing limit, the dominating behaviors within a largish RVE sampling of these crystal grains emerge as simple Hookean elasticity, requiring only two descriptive parameters (the moduli E and μ).

But simple randomization may not capture the proper smearing out of relationships that interconnect RVE levels within a complex material like steel because its lower-scale RVE units are often either laminates or “frozen disorder” crystal arrays.27

27 The igneous rock examples of Essay 1 supply excellent illustrations of the variations I have in mind. I should also note that most homogenization techniques involve further forms of subtle stochastic considerations that I’ve largely ignored here, but these appear as significant factors within most plausible forms of multi-scale modeling.
instance, a suitable homogenization technique must attend to both layer orientation and the manner in which their interfaces are affected under distortion. In the second case, dislocation movement and recrystallization become significant factors. Indeed, simple randomized composites of the sort sketched in the previous paragraph prove rather rare in nature, because the interstices between our little jammed crystals serve as sites of chemical change that complicate our homogenization policies further.

Once again, it is a profound mistake to view these scale-based dependencies as grounded in “mere epistemology.” The homogenizations at the core of multiscalar technique attempt to mirror the natural energetic cascades that allow a solid material to retain, and to eventually lose, its capacities for performing coherent work upon its neighbors. At the same time, in assembling an effective computational architecture, we must battle against mathematics’ inability to supply us with the dominant behavior conclusions that we really require.

In section (iv) we shall find that these asymptotically extended policies of homogenization allow the various RVE-centered sub-models within an encompassing multiscalar architecture to “talk to one another” without falling victim to amalgamationist contradictions. Here and elsewhere, asymptotic relationships supply some of the essential stitching that holds the fabric of present-day science together.

Due to these additional considerations, proper homogenizations often deliver limitations on the range of values that macroscopic parameters such as $E$ and $\mu$ may assume, rather than providing specific numerical values. This limitation doesn’t impair the significant utilities of the homogenizations, for we are often mainly interested in obtaining lower-scale messages that warn of potential structural damage. But I shall not enlarge upon these interesting complications here.

28 In their classic Treatise on Natural Philosophy (Vol. 2, §444 (New York: Dover, 1962) reprinted as Principles of Mechanics and Dynamics), William Thomson (= Lord Kelvin) and P. G. Tait prophetically write:

Enough, however, has been said to show, first, our utter ignorance as to the true and complete solution of any physical question by the only perfect method, that of the consideration of the circumstances which affect the motion of every portion, separately, of each body concerned; and, second, the practically sufficient manner in which practical questions may be attacked by limiting their generality, the limitations introduced being themselves deduced from experience, and being therefore Nature’s own solution (to a less or greater degree of accuracy) of the infinite additional number of equations by which we should otherwise have been encumbered.

In fact, they are discussing the physics avoidance advantages of exploiting upper-scale constraints like rigidity in Lagrange’s virtual work manner. In Essay 7, I argue that such reasoning techniques should be viewed as early progenitors of modern multi-scale methods.

29 Since Ernest Nagel, philosophers have construed the problems of relating higher-scale behaviors to lower-scale behaviors as “intertheoretic relationships” between two independently formulated theories, $T$ and $T^*$. The manner in which multiscale models effectively combine reliable data extracted from a variety of observational scales shows that this picture is too simplistic. In effect, Nagel and his followers have attempted to replicate the workings of sophisticated collections of homogenization arrangements employing simple logistical resources only. This “logification” of non-logical relationships is characteristic of the Theory $T$ thinking often criticized in these essays. These observations are similar to those advanced by Robert Batterman in his Devil in the Details (New York: Oxford University Press, 2001), a book of great influence upon my own thinking. At root, Nagelian policies appear to stem from a naïve mischaracterization of the proper relationships between statistical mechanics and thermodynamics. For more on the complications that solid materials contribute to this picture, see Essay 4.
Based upon these remarks, let's now construct a simple multiscalar modeling for steel that can consistently link RVE models on three size scales together through a clever policy of communication through homogenization. Here's a typical illustration of the difficulties we face. Suppose we run a very hefty object (a 4-6-4 Hudson locomotive) back and forth across a victimized bar of steel. These high-scale locomotive events eventually display trickle-down effects at the minute length scale of the dislocations, after passing through an elaborate hierarchy of higher-scale structures (that we shall largely ignore in our toy model, but may require direct attention for better results). As we noted, from a molecular perspective, dislocations represent slight imperfections in the bonding lattice, but they cluster together as lines and other groupings when viewed at a coarser resolution scale. As the locomotive stresses the bar on a macroscopic level, the little dislocations far below move across their encompassing lattices, shielding the individual molecular bonds from excessive stresses. Through these mitigating interventions, the bar becomes tougher than it would otherwise be.

But all good things must come to an end, and, after repeated migrations, the dislocations pile up along the cementite walls that hem them in. When this happens, the molecular bonds become less protected, and our bar becomes brittle. As a gradually accumulating trickle-down effect, running a locomotive repeatedly back and forth across a piece of railroad steel eventually makes the material crack more easily due to the microscopic migrations of its dislocations. As noted previously, hysteresis is the official term for history-dependent effects of this type. They are notoriously hard to model through conventional, single-level descriptive methods.

Clearly, the range of size scales that we must tie together in this scenario of lower-scale damage inflicted by upper-scale punishment is tremendous, giving rise to the tyranny-of-scales issues raised by Oden et al. above. In practical terms, this tyranny tells us that we cannot develop an adequate account of rail hysteresis working upwards from the molecular scale in a naive manner. Multiscalar models evade these computational barriers by enforcing a cooperative division of descriptive labor amongst a hierarchy of RVE-centered sub-models, each of which is asked to only worry about the dominant behaviors arising within its purview. This allocation of localized endeavor captures the
tactics that the physicist Weinan E has in view when he writes of “assuming that nothing interesting is happening at the larger scales”:

"Traditional approaches to modeling focus on one scale. If our interest is the macroscale behavior of a system in an engineering application, we model the effect of the smaller scales by some constitutive [modeling] equations. If our interest is in the detailed microscopic mechanism of a process, we assume that there is nothing interesting happening at the larger scales, for example, that the process is homogeneous at larger scales."

In other words, each local RVE sub-model directly responds only to its local environment, rather than to events that arise within distant sectors or upon alternative size scales. After these parochial assessments are finished, each RVE sub-model sends appropriate messages to its companion sub-models that report upon the local conclusions it has reached. Here the mathematical filtering of message through homogenization plays a crucial role, as we shall see in detail later. We then readjust local parameters within each of our RVE units until our flurry of interscalar reports is rendered self-consistent. The essential secret to multiscalar reasoning compression lies in the required self-consistency checks.

Here’s how the computations might unfold in a simple implementation of this scheme. We begin by modeling the locomotive’s weight as a macroscopic boundary condition applied to a standard Hookean continuum model of the rail itself, relying upon the bulk parameters (E and μ) appropriate to normal, isotropic steel, along with appropriate assessments of its toughness (e.g. fracture strength). On this upper-scale basis, we map out how stresses should distribute themselves across the entire bar, employing some conventional numerical method.

---

such as the finite element techniques of Essay 8. Once these initial, top-down computations have been completed, we survey the results looking for particularly high levels of shearing upon a RVE sub-level localized to 50 μm, which is where the grain structure of the steel first becomes evident. We then shift our attention to a replacement modeling that describes our formerly smooth 50 μm cube as a laminate composed of disordered layers of ferrite and cementite. We insert this new sub-model into a stress environment that pushes and pulls upon the cube in a manner determined by the local traction force strengths that we obtained within our initial upper-scale modeling. Employing our 50 μm sub-modeling, we compute how a laminated cube will shear within the posited stress environment. The upshot may indicate that the assigned stress levels will damage the cube, reorienting its grain into anisotropy.

However, let’s assume that no damage of this sort occurs; that the composite grains shear in a simple manner that harmonizes fully with our upper-scale computations. But we still need to worry about hysteresis damage arising at the RVE-level of the dislocations. So we must extend our modeling hierarchy with one (or more) layers of sub-model. In particular, we should now consider a new structure comprised of a minute block of pearlite matrix to which an unbiased statistical distribution of dislocation lines has been assigned, in conjunction with fixed geometrical barriers, internal structures that represent the hard cementite walls meandering throughout the interior of the pearlite. To this new sub-model we once again attach a uniform stress environment extracted from our higher-scale laminate modeling. Under a directional stress of the sort that heavy locomotives supply, the trickled-down stress environment will make the dislocation lines shift in a plastic manner—they don’t return.

---

32 If our beam is sufficiently symmetric, we might have employed our earlier Fourier analysis (rather than finite elements) to perform this initial analysis, generally at a lower computational cost. However, finite elements are better suited to the lower-scale corrective messages we will soon utilize, so it is customary to employ the latter in our assessments of upper-scale stress distribution.

33 In formal terms, the macroscopic stress distribution within our rail usually varies substantially across its full breadth, but these wider scale complexities are treated as uniform tractions at the laminate level, on the grounds that each small RVE sub-model directly responds only to its local environment. The latter will not vary much at this scale level. For simplicity’s sake, I have neglected temperature effects in our model, which would be a mistake if we were dealing with a material prone to temperature-driven phase changes.

34 This is an issue of prime concern within the granite/gneiss example of Essay 1.
move too close to the cementite walls and entangle themselves there. In that condition, they lose their capacities for protecting the molecular bonds from the locomotive’s mighty shearing stresses. At this point, a corrective message must be sent to higher submodels reporting on this damage, demanding local corrections in the key parameters of these models. Thus, our lower-scale pileups may force a stiffening of the original $E$ and $\mu$ values assigned to the steel, as well as lowering its estimated fracture strength. Because a large number of repeated locomotive poundings are required before the dislocation drifts affect our rail’s higher-scale behaviors in an appreciable manner, most of the time we must merely retain a record of how much closer to the cementite walls the dislocations have statistically drifted. However, once this accumulated history exceeds a critical threshold, an urgent message of parameter adjustment must be sent to the higher sub-model—“your fracture strength has just decreased in an alarming fashion!” In this manner, the computational architecture of a multiscalar modeling scheme can successfully ape the physical manner in which hysteresis effects slowly corrupt the properties of a macroscopic body.\footnote{Utilizing a metaphor from Essay 1, a dominant-level accumulation process operating at the level of the dislocations grows into an-invader-from-below effect that overturns the usual upper-scale malleability of steel and makes it more brittle.}

Once parameters in our higher RVE sub-models have been adjusted to reflect these damage reports from below, we must execute our original finite element stress calculations over again, to ascertain what effects local patches of damaged material might have on the rail as a whole. To this purpose, most multivalued schemes employ successive approximation iterations in the general manner of many of the explanatory landscapes examined in this book. The telltale signal of these iterated correctives is the large feedback arrow on the right of our completed computational flow chart. Commonly,

\footnote{To estimate these shifts, practitioners often simplify their task by calculating how a single test inclusion is likely to shift within the posited stress environment, following the procedures pioneered by J. D. Eshelby. On this basis, they can derive a statistical estimate of how a larger population of dislocation lines adjusts. But these details will not concern us here.}
our adjusted parameter corrections must be repeated many times over until a stable, self-consistent answer is reached, at which point we can declare that our individual RVE sub-models have reached descriptive harmony with one another. Such searches for self-consistency amongst dominant behavior pitched at a variety of scale sizes is a general hallmark of multiscalar technique.37

For this essay’s purposes, the most intriguing aspect of these procedures lies in the manner in which each sub-model “sends corrective messages” to its comrades-in-modeling pitched at other RVE scales. Scrutinizing our three specimen forms of sub-model closely, we find that they all employ exactly the same language—classical continuum physics—in incompatible ways. That is, they each employ infinitesimal equations characterizing their internal material as some form of continuous, flexible matter.38 But they each do this in incompatible ways, and, accordingly, pose substantive greediness-of-scales problems. If we blithely combine each sub-model’s output in Quine’s amalgamationist manner,

37 With respect to the explanatory landscape issues of Essay 2, note that all of the sub-model calculations within our scheme operate upon a “find a new equilibrium” basis; no explicit representation of passing time in an evolutionary manner occurs anywhere. But hysteresis is plainly a time-dependent affair: where does time show its handiwork within our scheme? Answer: its “clock” is codified within the altering records we keep of dislocation movements after each cycle of load/unload patterns.

38 Many authors comment upon the striking fact that very complicated material effects can be treated utilizing classical sub-modeling techniques alone, with little need to descend into the realms of quantum modeling. But appeals to quantum theory are sometimes required, and classical/quantum crossovers can arise in intriguing combinations. A case in point: the enclosing walls of nanotubes often demand a quantum mechanical treatment whereas a classical description suffices for the fluid within. Interesting “handshake” techniques have been developed to allow communication between these two modes of RVE-level description. Cf. John C. Berg, An Introduction to Interfaces and Colloids: The Bridge to Nanoscience (Singapore: World-Scientific, 2009).
serious forms of syntactic inconsistency will ensue. Plainly we must keep the parochial conclusions of our various sub-models disjoint from one another. On the other hand, we must find some method for allowing our sub-models to communicate their findings with one another. How should we compose the interscalar messages required? The remedy utilized within standard multiscalar methods processes each sub-model's parochial conclusions through one of the homogenization filters outlined above (viz. enlarging the local model to an infinite population and extracting the desired parameter corrections through classic “mean and variance” asymptotics). And so the corrective messages between scales consist in management imperatives of the form:

*Your original computations are based upon a modeling that utilizes the parameters A, B, C, . . . within region R. However, in regions of problematic stress, my local calculations inform me that small portions of the material at position R will behave differently within the stress environment you have assigned to R. Please alter your A, B, C, . . . parameters assigned to R and try again.*

In this manner, acceptable communication between scales operates in the manner of a telegraph, in which homogenizations serve as the equivalent of Morse code.

Why does this tactic so often succeed, when single-level modeling attempts often fail? Physically, breakdowns in the behavioral patterns normally prevalent at size scale \( \Delta L^* \) commence their destructive careers as normal smaller-scale \( \Delta L^* \) events that eventually grow into substantive effects with serious repercussions for \( \Delta L^* \). Our computational policies imitate these relationships as closely as we can in mathematical terms.

In adopting a trickle-down architecture of this type, we follow a policy that physicists commonly describe as “opening up frozen degrees of freedom only as needed.” By also accepting corrective adjustments from other sub-models when they are genuinely required, we find ourselves able to correct for the exaggerations inevitable within any individual modeling policy that attends only to dominant behaviors. In this manner, multiscalar techniques represent a deft compromise between the purist top-down and bottom-up methodologies that
once divided nineteenth-century philosophy of science into warring philosophical camps.39

The outputs of multiscalar modelings supply mixed-level explanations in the sense that their descriptive architectures generally stem from direct empirical observation of the manner in which various RVE scales causally affect one another within a complicated material. We don’t pretend to have “derived” these empirical hierarchies from molecular fundamentals; we instead exploit our direct knowledge of physical layering to better computational advantage.40

It is plain that these techniques can significantly curtail the computational explosions characteristic of tyranny-of-scales circumstance without ruining our capacities to profit from the rich modeling knowledge we possess of a material on its various RVE scales. But notice: each RVE sub-model communicates with the other RVE units through homogenized messages. \textit{Pace} Quine’s amalgamationist asseverations, they do not attempt to reach a univocal account of “what’s going on” in the material. Despite this lapse, we should still feel that we understand the material pretty well through such means, a consideration to which we’ll later return.

(v)

Let us now consider several formal observations that may strike some readers as picayune, but which will bear important methodological fruit later on. In many parts of this book, I complain about philosophical propensities for lumping together significantly different forms of explanatory architecture. In this manner, standard Theory T thinking loosely classifies the foregoing computational scheme as an “initial value problem,” but this characterization is quite inaccurate. Properly speaking, our multiscalar modeling consists of a sequence of linked equilibrium problems,41 in which our

39 See Essay 4 for references on the debate between “rari- and multi-constant” approaches to elasticity.

40 This observation is allied to the misconstruals of renormalization group explanation discussed in Essay 2. We might also observe that our homogenization techniques nicely illustrate Heaviside’s recommendation that descriptive exaggeration of a greediness-of-scales character should be corrected by a contrary dose of infinitary extension, in this case the employment of infinite population asymptotics.

41 There is no barrier to modeling our problem in a truly evolutionary manner at the highest size scale—we might wish to model our rail with hyperbolic modeling equations. The ensuing dynamic complexities are probably not important in our railroad circumstances, although they might become so if inertial effects play a larger role in the developing hysteresis. In the treatment provided, the only “clock” appearing in our modeling is one determined by the alternating schedule of loading and unloading that we considered. I mention these considerations only to underscore a chief recommendation from Essay 2. To correctly identify the strategic structure behind a successful bit of scientific explanation, we should pay careful attention to how considerations of temporal change are reflected (or avoided) within the modeling effort. As I argue on several occasions within these essays, significant conceptual confusions can be clarified through stricter diligence on these formal scores, whereas Theory T propensities for recasting every form of explanatory gambit into initial value terms is not helpful at all. In an allied manner, we should also attend to background assumptions as to how rapidly various relaxation times transpire. In the situation at hand, we have tacitly presumed that the characteristic relaxation times of our various microscopic RVE units are much faster than the timescale upon which macroscopic level events unfold.
steel rail is subjected to a schedule of alternating upper-scale boundary value loadings (with locomotive and without), wherein we keep track of accumulating lower-scale damage. The only “initial values” utilized involve a statistical estimate of the original distribution of dislocations at the appropriate RVE level. Furthermore, the upper-scale boundary conditions originally assigned to our problem—e.g. a continuously distributed load of five tons placed upon the rail between spots A and B—are macroscopically described and ipso facto unsuited to the lower-scale sub-models utilized within our scheme. Indeed, most of these little RVE volumes lie at strongly shielded distances from the locomotive itself and, accordingly, require their own specifically adapted forms of side conditions. These we derive indirectly from the finite element calculations of stress distribution that we carry out on the highest, completely macroscopic scale; we might call these trickle-down boundary requirements. In this indirect manner, our RVE-localized boundary conditions reflect the locomotive loading originally specified in our problem, but these RVE-tailored boundary strictures must become more sharply detailed as we reach ever smaller physical units. Through these trickle-down dependencies, assignments of a localized stress environment at the RVE level remain computationally hostage to our overall requirements of cross-scalar consistency. We may need to recalculate suitable boundary conditions for a particular RVE unit many times before our estimates stabilize upon a mutually accepted value. For such reasons, multiscalar approaches follow neither purist top-down nor purist bottom-up tactics in their approaches to modeling, but qualify as intriguing hybrids.

Reflecting upon these formal matters, we can recognize why purist bottom-up techniques are likely to fail with respect to solid materials that contain significant quantities of higher level, frozen-in-place structure. The dislocation lines provide a characteristic illustration. Consider the problem of correctly estimating the local stress environment around a small RVE block of steel, call it A. The actual binding strengths within A’s interior will be significantly affected by the intensity of its environmental confinement. Furthermore, we are chiefly interested in what happens to A when we run a locomotive over its enclosing rail on a scale size far above. As we saw, the coherence and intensity of any energetic pulse traveling from this macroscale down to A will be greatly affected by the energetic cascade lying in between. Our multiscalar modeling policies take direct advantage of our empirical knowledge of how these between-RVE-level transitions characteristically operate, for this structural knowledge directly informs the computational architecture we assign to the problem. In contrast, a bottom-up approach must estimate A’s local stress environment using facts available at a molecular level only. Rarely can we do this adequately (we’ll consider a few

42 Typically, we smooth out these results to provide a completely homogeneous stress environment around our RVE interior. The physical presumption behind this tactic is that a small piece of the steel will “feel the locomotive” only through the averaged pressures it receives from its next door neighbors. On our locomotive’s size scale, the loading doesn’t need to be very precisely specified, but at the level of the dislocations the trickle-down boundary conditions must be.
exceptions in a moment). This is because A’s local stress environment generally carries structural markings of its prior history, especially with respect to the manner in which intervening structures have modified the wave coherence within the locomotive’s original poundings. The more brittle our steel, the stronger the directive coherence of shear remains at the trickled-down level of A’s local stress environment. But upon this lowly scale, the environmental adjustments produced will be extremely delicate.

A glance at microscopic images of such materials indicates the problem. On the left, we witness a scanning tunneling photograph of a tiny block of molecular lattice, in which some barely detectible slight irregularities appear. Viewing the same material under lower magnification, we recognize that these same tiny irregularities have collected themselves together into the line-like patterns that we call dislocations. We have also noted that the mobility of these lines greatly affects the energetic cascades that connect A with our locomotive, by diluting an incoming wave’s energetic coherence significantly. This degradation occurs because much of the energy within our arriving pulse becomes diverted into moving the dislocations across their lattice background, rather than coherently attacking A’s lattice bonds at full strength.

Once dislocation lines become locked into a material, removing them can be quite difficult. When we attempt to straighten out a small block of slightly irregular lattice, the dislocations simply slide into the block next door, as cockroaches move from one apartment to another when the exterminator arrives. But how did these little distortions get into our material in the first place? Like most of the distinguishing features of an everyday solid, they are there because their frozen disorder was locked in place long ago.
by some form of higher-scale process, ranging from the rapid cooling of a molten state to the mutilations inflicted by the village smithy. 43

As acknowledged above, purist bottom-up modeling techniques can sometimes supply good estimates of material parameters like toughness, if the material in question is simply structured. Indeed, the official demands of textbook “equilibrium thermodynamics” require such simplicity, for they expect that all frozen-order irregularities will have shaken themselves out of a target material before thermodynamic concepts like temperature and entropy are regarded as applicable. By these lofty standards, we will need to wait a very long time before the required conditions overtake most familiar solids—there are a lot of out-of-equilibrium rocks that have been with us for eons. Officially, then, we’re not supposed to talk of the “temperatures” of rocks, steel bars, or locomotives (I’ll come back to the plausibility of these stern strictures in a moment). The doughty reluctance that most solids display with respect to the demands of conventional thermodynamics trace to the fact that, once in place, most significant forms of higher-level structure are hard to dislodge, in the general manner of the dislocations. In terms of official “equilibrium thermodynamics” demands, the conditions laid down for a steel rail at room temperature require that the material has sorted out its chemical components into a motley of perfected crystalline whiskers like the specimens sometimes grown within a pristine, gravity-free environment. Applied to these very special materials, conventional methods of statistical averaging can supply reasonable estimates of localized stress environment following purist bottom-up tactics. 44 But these successes very much trade on the level-free simplicity of the perfected lattice bonding involved. Much the same can be accomplished with respect to materials that are completely disordered, such as a close-to-ideal gas.

In the great fullness of time, our steel rail should eventually wiggle its way to its thermodynamically preferred, perfect lattice configurations, at which point these bottom-up estimates of local stress environment will prove successful. But these are not the solid materials with which we are concerned in the here and now. Workers in materials science now recognize that the hierarchies of trapped, albeit “thermodynamically unstable,” structures encountered in most solids stymie modeling attempts that proceed on a conventional statistical mechanics basis. In this vein, Ellad Tadmor and Ronald Miller complain about the errors in applying standard statistical mechanics techniques to solids:

43 The igneous rock example of Essay 6 illustrates these morals vividly. The structural factors that distinguish one igneous rock from another reflect gross differences in their prior histories—how they were blown about or crushed by volcanos, squeezed by massive hot mountains, or cooled rapidly through exposure to the air. The full energetic environment of a minute piece of molecular granite reflects the full accumulation of these higher-scale, frozen disorders in a very delicate manner that does not submit readily to direct descriptive attack.

44 Doing so usually requires extending our little whiskers into infinite lattice monstrosities, to counteract mathematics’ stupidity in refusing to overlook finite boundary effects. But such considerations don’t reveal the more important forms of coherence filtering characteristic of our multiscalar homogenizations.
The Greediness of Scales

[A] standard statistical mechanics phase average integrates over all configurations accessible to the system regardless of the amount of time that one would have to wait to reach such a configuration. The resulting phase average will therefore provide a smeared view of the liquid-like behavior that a solid will exhibit over infinite time.\(^\text{45}\)

Allied observations apply, mutatis mutandis, for attempts that attempt to directly apply the group theoretic techniques that work ably for perfected crystal lattices.

In an allied vein, Gene Mazenko comments upon the advantages of the new multi-scalar techniques (which he calls, somewhat misleadingly, “coarse graining”):

The key idea in developing a coarse-graining approach is to average over inconsequential microscopic details but leave unaveraged the important degrees of freedom that control the long-wavelength behavior of the system. The price we pay in many cases is to inherit a theory characterized by phenomenological parameters . . . It is difficult to determine the parameters characterizing these interactions from first principles. Although this coarse-graining approach has not traditionally been favored by those advocating a rigorous approach to statistical mechanics, its influence has increased with time.\(^\text{46}\)

This is a subdued means of highlighting the advantages of assembling computational architectures that blend together reliable data extracted from empirically substantiated models centered along a variety of characteristic scale lengths.

Conventionally educated philosophers will frequently dismiss these mixed-level computational architectures as merely useful approximation techniques, possessing no philosophical salience whatsoever. On what grounds? Most likely, they will reply, “In principle we can always compute the tensile behaviors of our steel bar in a bottom-up manner working from the initial conditions of its molecular components.” In the light of the tremendous tyranny-of-scales obstacles that arrest such endeavors, we should mistrust the unverifiable character of these “in principle” appeals (in the next section, we’ll discuss why airy dismissals of this character may overlook important clues to the mysteries of quantum mechanics). For the moment, however, let us merely note a characteristic fallacy that often lies behind these glib “in principle” assurances. Confronted with the severe tyranny-of-scales obstacles to bottom-up technique that solids present, philosophers often appeal to the notable successes of what Mazenko calls “traditional statistical mechanics,” viz. standard approaches to ideal gases and perfect crystal arrays. But single victories do not insure the success of an extended campaign. We shouldn’t generalize from the real but limited successes of “traditional statistical

\(^{45}\) Ellad Tadmor and Ronald Miller, Modeling Materials (Cambridge: Cambridge University Press, 2011), p. 553. These authors are thinking of materials like rubber, whose room temperature equilibrium should be a soup of detached polymer strands, rather than a perfected lattice whisker.

\(^{46}\) Gene Mazenko, Fluctuations, Order and Defects (New York: Wiley, 2003), pp. 57–8. Articulating our discussion in alternative terms, traditional attempts to estimate tensile strength through purist bottom-up techniques such as mean field estimation fail miserably, due to the energetic meddling of hard-to-see, higher-level structures like dislocations. Philosophers who blithely appeal to “what we can calculate from the governing Hamiltonian” should reflect on the fact that such formulas (if they don’t pretend to govern the entire universe) invariably contain residual terms of the position-dependent form \(V(x)\). Within these little \(V(x)\) supplements, the shaping subtleties of higher-level coherence or incoherence secretly lurk. Just try to write down a concrete \(V(x)\) that performs the required trick!
mechanics” to the conclusion that the same pristine methods will work “in principle” everywhere. That supposition overlooks the fact that these bottom-up success stories rely heavily upon the absence of higher-scale irregularities within a gas or perfect crystal.

At this point our discussion connects with the tricky conceptual issues concerning the term “thermodynamics” canvassed within Essay 4. In truth, we can’t coherently understand the utility of notions like temperature and entropy unless they can be meaningfully applied to normal solids and the like, despite the fact that such applications must, perforce, belong to the theoretically fraught dimensions of so-called “non-equilibrium thermodynamics.” As the Duhemian analysis of that essay makes clear, the prissy demands of orthodox equilibrium thermodynamics reflect the subtle conceptual hierarchies required to raise “temperature” and “entropy” to reasonable applicational breadth. These issues in turn trade upon the difficulties we confront in outfitting notions like “loss of informational coherence” within effective forms of mathematical clothing. I’ll return to this issue in the next section.

Imprecise philosophy of science diagnosis plays a significant role in encouraging these misapprehensions as well. As remarked earlier, Theory T proclivities characterize our “what happens to a steel rail when a locomotive pounds upon it?” concerns as a simple initial value problem, which can be usually articulated in a completely bottom-up manner. However, this categorization is incorrect; we are instead confronted with a complex computational architecture that has been assembled from localized forms of constrained equilibrium sub-models, linked together through homogenized relationships that have been selected to imitate the cross-scalar dependencies found within the target material. A proper identification of these structural ingredients makes it evident that these modeling practices do not conform to purist bottom-up expectations, but directly incorporate significant doses of higher-level empirical observation within their explanatory architectures. In the next section, we’ll find that non-Theory T structural arrangements of a multiscalar cast may provide important harbingers of future science; I do not harbor dogmatic opinions on this score. But we shouldn’t dismiss such alternatives out of hand, based upon dubious claims such as “in principle problems like yours can be recast as initial value problems.” At present, we know no such thing, even if coarse Theory T categorizations encourage us to presume otherwise.

Essay 9 briskly reviews the fact that the early pioneers of mathematical physics did not presume that nature’s activities can always be captured in mathematical terms—we must instead search for those special opportunities in nature where our computational procedures manage to match physical behaviors in an effective way. But many of those original reservations derived from the fact that differential equations of a sufficiently broad class had not yet been invented. Sometime in the 1700s a contrary mathematical optimism began to fall in place, in which it was first hoped that all of nature’s behaviors might submit to
mathematical description, roughly following the patterns of Newton’s ODE successes within celestial mechanics. But for this new faith to become viable, the traditional dominions of mathematics needed to be greatly liberalized, as Essay 9 also indicates.

Many contemporary philosophers regard this descriptive optimism as established dogma, but closer attention to the working architecture of actual science suggests that greater fragmentation appears within applied mathematical practice than they recognize. Indeed, the present-day fabric of working science is sewn together with liberal doses of scale-focused patchwork and asymptotic stitching. Modern multiscale methods illustrate these formal features in the beautiful manner in which homogenization relationships (a form of asymptotics) link together distinct patches of dominant behavior modeling. The practical viability of these divisions of descriptive labor policies suggest that compartmentalized modes of scale-focused description might comprise non-eradicable elements within any future science. Such speculations have certainly occurred to others, such as Graeme W. Milton:

> [W]hat we learn from the field of composites could have far-reaching implications in many fields of science. Significant progress in improving our understanding of how microscopic behavior influences macroscopic behavior could impact our understanding of turbulence, of phase transitions involving many length scales, of how quantum behavior influences behavior on classical length scales, or, at the more extreme level, of how behavior on the Planck length scale influences behavior on the atomic scale. While that may seem unlikely, it is hard to deny the impact that our understanding of classical physics had on the development of quantum physics. Therefore it is conceivable that a better understanding of classical questions involving multiple length scales could have large reverberations.

Perhaps science will never obtain the purist syntactic contours beloved of Theory T enthusiasts, purged of all forms of interscalar cooperation and unequal data registration. “We cover the world in protectorates” might serve as the watchword of such an eventuality.

It is important to recognize that these concerns have little to do with any issues of “scientific realism,” but instead focus centrally upon the representational capacities of mathematics. Are the descriptive tools it naturally provides too “stiff” to track evolving natural processes ably over long stretches of time or distance? It is undeniable that we can’t always trust differential equation models to the representational extent tacitly

---

47 See Essay 6 for examples.
49 We should add “or scale” to this list as well, for it often happens that a relativistic model may supply a useful account of cosmological evolution on a coarse-grained scale, but must be patched to a finer-grained description of the resident matter through some asymptotic matching technique. Such formats of “protectorate stitching” merit closer formal attention by philosophers.
presumed by naïve optimists. Our greediness-of-scales dilemmas illustrate some of the basic concerns. To serve as proper engines of effective inference, the application range of our interior modeling formulas must often be extended far beyond their proper range of descriptive applicability (down to an infinitesimal scale, in fact). To compensate for the resulting descriptive distortions, multiscale methods impose various corrective “anti-dotes,” involving sundry forms of homogenization filtering and the like. Only through these compensatory tactics can we persuade our various patches of RVE modeling to cooperate with one another profitably. Nonetheless, we can still feel that we have captured the natural situation fairly ably within such an integrated framework—as remarked before, we don’t seem to be neglecting anything of importance. But we can no longer regard each individual sentence within the encompassing scheme as directly “capturing the entire truth” about nature’s relevant activities. This is a general feature of distributed labor policies; no worker in an assembly line builds an entire automobile either. On such grounds, canny mathematicians recommend that the descriptive merits of a modeling should never be laid entirely at the door of its interior equations alone, but should be distributed over the entire family of side condition ingredients that operate as a cooperative family in implementing a well-designed explanatory strategy.50

In a related vein, naïve optimists have a difficult time explaining why the applicational locales to which conventional boundary conditions are attached generally represent physical regions in which far more complex processes transpire than arise inside the interiors they surround. The seemingly placid surface of a steel bar conceals a riotous array of chemical carryings on, far more complex than any that occur within the interior steel. Why should science register its representational data in this grossly unequal fashion, wherein interior regions are accorded more detailed attention than their boundaries? A formal response mixes the mathematical with the pragmatic. Mathematical, because differential equations are quite picky about the formal aspects of the side conditions with which they will inferentially cooperate, and pragmatic, because nature often allows us to get away with cruder rules of thumb within the narrow locales that we characterize as “boundaries” and “interfaces.” How can we get away with this inequity? Mainly, because the relevant regions are comparatively small and we can always open up suppressed details as necessary. For such reasons, applied mathematicians frequently warn against presuming that the interior differential equations utilized within a

50 For examples, see Essay 8. Present reflections, in fact, comprise the chief theme of that essay, there analogized to the manner in which Donald Duck regularly steals credit that justly belongs to his cohort of helpful nephews.
modeling are ipso facto “more accurate” than their endpoint characterizations. Indeed, as pieces of descriptive machinery, the boundary region specifications enjoy their own descriptive prerogatives and the imperious demands of interior equations must often be recalibrated to suit the cooperative harmony requirements of the wider ensemble to which they belong. Essay 8 illustrates in detail how this adjustment of interior demand sometimes proceeds.

To these cooperative family considerations, we can add the further methodological datum that the functions that lie most readily to hand within mathematics often display internal characteristics alien to the physical processes they allegedly represent. At present, these discrepancies are often corrected by relatively crude remedies, such as inserting a brute singularity to prevent analytic continuation for installing an unphysical form of dependency within a modeling function. See Essay 9 for more on this.

Finally, capturing the crucial notions of information loss and energetic degradation in satisfactory mathematical terms has proved an elusive goal within present-day physics. Within the schemes studied here, these factors are encoded within the homogenization averagings that connect one RVE sub-model with another. As Essay 4 discusses, modeling decisions as to when a coherent pressure has degraded sufficiently to become an incoherent heat appear to be handled in allied ways. Perhaps these formal considerations suggest that interscalar averagings serve as some flavor of Heavisidean antidote for counteracting the descriptive exaggerations of a dominant behavior sub-modeling. I don’t know, but find the suggestion intriguing.

(vii)

However, deciding that no adjusted form of mathematical optimism is viable is premature, for only the naïve differential equation expectations tacitly favored by Theory T thinkers should be summarily rejected. After all, conceptual advances within modern mathematics have displayed a remarkable capacity for generating new forms of mathematical object that remain in better accord with recalcitrant target phenomena. Sometimes the very fact that familiar modes of numerical reasoning break down over longer stretches of space and time provides the essential clue that clever mathematicians follow in assembling structures that map onto external behaviors in a more accurate manner. Indeed, this observation provided the essential guidance that led Riemann, Weyl, and others to introduce the notion of “manifold” (in its sundry guises) into modern mathematics. For consider how conventional inference goes astray when we are confronted with a curved manifold such as the earth.

Suppose that we have constructed a simple differential equation model for a goose’s flight path. On this basis, we can plot out our goose’s progress on a piece of graph paper. However, after the bird has traveled a certain distance across the globe, strange

51 Mathematical entities can sometimes be defined in terms of what they’re not.
anomalies will appear in our plottings, no matter what coordinates we adopt within our graph paper charts. Why? Because the earth’s curvature doesn’t allow it to be smoothly mapped to a two-dimensional flat map. These topological discrepancies force us to replace any starting map with an overlapping replacement chart once our goose flies a sufficient distance. The differential equations employed within our initial plottings stop making descriptive sense once these computational horizons have been breached. Accordingly, our original formulas need to be replaced by refurbished equational sets tailored to the new piece of graph paper in which we plot our fowl’s continuing flight.\(^5\) Plainly some underlying pattern lies behind these inferential adjustments, which can be regarded as a telltale symptom that “something funny” has occurred within our target manifold (viz., the topology of the earth differs from that of a coordinate chart). Such considerations led mathematicians to the conclusion that unfamiliar structures (such as differentiable manifolds of high dimension) could be introduced into mathematics through appeal to the patterns in which coordinate charts should be stitched together to overcome their internal descriptive failures.\(^5\) Once we have these new manifolds in hand, we can revisit our old notions of “differential equations” and recast their demands in manifold-friendly terms, where we can now freely appeal to vectors, one-forms, and the like, which tacitly incorporate the necessities of shifting charts within their internal operations.

The abstractionist “philosophy” inherent in this approach rather resembles the old parable of the blind men and the elephant:

> A number of blind men came to an elephant. Somebody told them that it was an elephant. The blind men asked, “What is the elephant like?” and they began to touch its body. One of them said: “It is like a pillar.” This blind man had only touched its leg. Another man said, “The elephant is like a husking basket.” This person had only touched its ears. Similarly, he who touched its trunk or its belly talked of it differently.\(^4\)

\(^5\) The analogy to the analytic continuation of complex-valued functions is evident and, undoubtedly, inspired Riemann and others in their thinking.

\(^4\) A longer discussion of this familiar construction appears in Essay 9. A related bootstrapping of the notion of “differential equation” comprises the primary topic of Essay 8.

\(^4\) Credited in an anonymous Wikipedia article to Ramakrishna Paramahamsa (http://en.wikipedia.org/wiki/Blind_men_and_an_elephant; 10/20/2014).
Have the observers managed to “understand the elephant” adequately if they pool their local information effectively with one another? The modern mathematical response to this old philosophical conundrum is assuredly “yes”; the local bits of information correspond to the covering charts in the standard treatment of a manifold. The line of thought behind all of this is deeply pragmatic in the philosophical sense—to “understand” an unfamiliar object is to know how to reason properly (i.e. calculate) in its presence.

Proposal: the sub-models within a multiscalar modeling scheme may “cover an unfamiliar earth” in more or less the same atlas-of-charts manner as the standard treatment of manifolds.

Perhaps these suggest manners in which we might remain mathematical optimists without succumbing to the Theory T dogmas criticized here. In truth, I don’t know where a proper dividing line between “optimism” and “opportunism” should be placed in view of mathematics’ remarkable abilities for expanding its conceptual horizons. But I do know that current practice within metaphysical inquiry ignores many difficult issues with respect to both mathematics and descriptive practice that ought to remain closer to the center of its concerns.55

(viii)

If our division-of-linguistic-labor considerations force a substantive rethinking of how ably the world’s behaviors can be directly captured in single sentence terms, Quine’s criterion for “ontological commitment” will become a first casualty of that re-evaluation. How to remedy this fault, I don’t know. Perhaps the “absoluteness” with which many philosophers approach the customary categories of “object” and “property” requires reconsideration as well. In many areas of physics, small-scale “objects” (e.g. phonons or solitons) only emerge with clarity within specially prepared environments, just as the

55 This is the central theme of Essay 6.
overtone spectrum “objects” lurking within a confined violin string become readily apparent only in the context of tightly pinned end conditions. Niels Bohr, in a long series of celebrated, but murky, writings, insists that modern descriptive science requires a “complementarity” between various modes of description. Unfortunately, he employed this term in two, possibly disconnected, senses—to capture the fact that non-commuting quantum variables cannot be employed at the same time and to claim that quantum-style descriptions must be articulated within a wider descriptive frame that relies upon distinctions of a “classical” character.56 It is only with this second, unequal-registration-of-data thesis that we are here concerned. Notoriously, Bohr was prone to articulate its contours in terms of human observation and understanding:

[The notion of complementarity refers directly to our position as observers in a domain of experience where unambiguous application of the concepts used in the description of phenomena depends essentially on the conditions of observation.57]

Insofar as I can ascertain, the intended sentiment is better captured in the terminology of unequal data registration, so that the gist of Bohr’s remark can be re-expressed as:

[The notion of complementarity refers directly to our algorithmic capacities within domains of experience where the obtainable range of trustworthy calculation within a specific mathematical modeling of the target phenomena depends essentially upon the degree of unequal data registration employed within the side conditions attached to the overall modeling effort.]

Likewise, when Bohr criticizes Laplace’s vision of a wholly predictable science based upon initial conditions:

In this whole conception which has, as is well-known, played an important role in philosophical discussion, due attention is not paid to the presuppositions for the applicability of the concepts indispensable for communication of experience,

the intended moral can be reconfigured as:

In the canonical expectations that science will eventually produce a wholly hyperbolic set of cosmological equations (which has, as is well-known, played an important role in philosophical discussion), due attention has not been paid to the unequal data representations required in effective mathematical modeling and the homogenization techniques required to fit these local patches together into a more effective global coverage.58

56 Perhaps the symbiotic relationships between interior equations and their enclosing boundaries provide a suitable formal analogy to what Bohr has in mind.


58 See Essay 2 for an explication of the formal classifications employed here.
Possibly, then, Bohr’s emphasis on the complementary importance of “classical” description should be reparsed as an insistence upon the policies that we have grouped together under the heading of a “cooperative distribution of linguistic labor between unequal modes of data registration.”

I am not prepared to wholeheartedly endorse any of these Bohr-like musings, for, at this moment in time, no one should confidently augur the syntactic future of science (even if some philosophers pretend that they can). At present, successful scientific description relies heavily upon elaborate assemblies of divided labor stitched together through non-amalgamationist tactics such as homogenization, found within reasoning architectures that roughly imitate the ways in which nature itself connects together its various forms of dominant and recessive behavior. Will future science continue to be characterized by these cooperative labor patterns, or might all of this partitioned complexity someday evaporate in favor of a simple amalgamated scheme of the sort that Quine and the Theory T crowd anticipate? I can’t answer this question, but I firmly feel that, in the here and now, philosophy should cultivate a warmer appreciation of the many varied patterns of reasoning that a useful descriptive practice might adopt.

To these ends, multiscalar arrangements of the sort surveyed in this essay provide excellent prototypes for the potential structuring of future science. Any definitive resolution of these questions must turn upon a finer estimation of our computational capacities—how far can local stretches of reasoning be extended before corrective remedies from other forms of data registration are wanted? The resolution of such questions lies in the hands of nature, not ourselves, and we’re not yet in a position to untangle all of its riddles. As we await its pleasures, we should cultivate philosophical open-mindedness with respect to mathematics’ descriptive capacities.

In many ways, I believe that this recommendation echoes the great metaphysicians of the past, whose grandest philosophical themes were partially inspired by close attention to the nitty-gritty oddities of effective mathematical modeling. These authors did not pretend (as many modern philosophers do) that with complete accuracy we can readily catch nature’s empirical behaviors within the descriptive framework of a differential equation modeling; they realized that more complicated arrangements of compensating ingredients might be required instead. As Essay 3 documents, Leibniz, in particular, anticipated many of the worries that we have collected together under the heading of “the greediness of scales.”

Such attitudes strongly contrast with those that prevail within philosophical circles today, in which considerations of “metaphysics” are pursued without concern for the issues of mathematical capacity and linguistic representation raised here. Insofar as I can determine, these separatist attitudes stem largely from the presumption that future science will assume the simple and fully amalgamated Theory T contours that Quine’s standard of ontological commitment requires. But from what crystal ball have they extracted these prophesies? Surely, not by induction upon current practice. These writers may someday prove correct in their futurist hunches, but none of us will live to witness that jubilant moment of Theory T ratification. In the meantime, we should expand our appreciation of the alternative manners in which blocks of reasoning can
usefully fit together, rather than fastening prematurely upon a Quinean architecture simply because it makes philosophy easier.

To me, the most striking characteristic of the recent metaphysics literature lies in its lack of engagement with the troubling oddities of successful mathematical description, declaring “That’s not my style” in the haughty manner of Casey at the Bat. They confidently presume that, even if present-day physics happens to be burdened with significant amounts of multiscalar patchwork, such blemishes will surely be removed within some hypothetical “fundamental physics” lying in our future.

But, once again, how do they know this? In my diagnosis, such presumptions trace to a certain vein of wishful thinking:

Wouldn’t it be lovely if, on the final day of epistemological judgment, scientific practice wound up conforming to the simple syntactic contours outlined by Carl Hempel in *Philosophy of Natural Science*? And: Wouldn’t it likewise be peachy if informational content and inferential warrant in a scientific language always fit together in the tidy patterns supplied by Alfred Tarski in his writings upon semantics?

Perhaps, but wishing doesn’t make it so. George Eliot offers sage advice on this very matter:

*Nature has her language, and she is not unveracious; but we don’t know all the intricacies of her syntax just yet, and in a hasty reading we may happen to extract the very opposite of her real meaning.*

---