Essay 1:

Physics Avoidance

[Dialecticians] prescribe certain formulae of argument, which lead to a conclusion with such necessity that, if the reason commits itself to their trust, even though it slackens the interest and no longer pays a heedful and close attention to the proposition inferred, it can nonetheless come to a sure conclusion by virtue of the form of the argument alone. Exactly so; the fact is that frequently we notice that often the truth escapes away from these imprisoning bonds, while the people who have used them in order to capture it remain entangled in them. Descartes\(^1\)

(i)

In whimsical moments, Brigham Young would order members of his flock to “travel two hundred miles directly south and plant cotton.” When these exiled settlers encountered some easily skirted ravine or coulee, they would not deviate from their instructions but would painstakingly carve ruts and fastenings into the rock so that their wagon trains could be pulled across the obstacle before them, obedient to their “travel directly south” directives. I fancy that on a blistering Utah afternoon some of these beleaguered souls might have entertained the impious thought, “Mightn’t we have gone around this $%#& canyon?”

And so it is within practical science. The surest and happiest routes to predictive, explanatory and design success do always not lie directly ahead, but employ clever stratagems for evading the computational hazards that render the direct path unpassable. Such evasive approaches succeed through adopting various covert strategies for what I shall call physics avoidance. Most physical treatments one encounters in real life are characterized by some form of physics avoidance or another. Often we should skirt difficult calculations if they aren’t required for the answers we seek and are prone to introduce errors. The resulting reduced treatments will be usually related to one another in rational patterns. We will do better in descriptive philosophy of science if we pay more attention to the sorts of formal inter-relationships that applied mathematicians
understand quite well.

Essentially, our project is to investigate how various strategies for effective “physics avoidance” can alter the background explanatory landscape in which a problem is framed. And we can capture this notion of “explanatory landscape” in (almost) literal terms if we borrow a technique from the mathematicians and view the changes wrought by “physics avoidance” as shifts from one portrait of the physical situation to another or, to adopt the preferred modern jargon, as maps between structurally related manifolds. Let me explain this somewhat abstract way of thinking with a few examples. First consider the vibrations of a violin string. Considered naively, if we pay attention to the current spatial locations $q(r)$ of every particle along the string according to its location $r$ along the central axis, we find that the various $q(r)$’s supply an uncountable number of descriptive parameters (or “degrees of freedom”) that must be fixed before the location of the string as a whole will have been specified. And these $q(r)$ locations both shift rapidly over time and are entangled with one another, resulting in patterns that are very hard to decipher. Abstractly, we can picture (insofar as it is possible to “picture” anything moving through an uncountably infinite space!) the string’s continuously varying $q(r)$ values as the movements of a single representative “phase” point within an ambient space of very high dimensions (which I’ve drawn as a chicken-like thing upon which the “phase point” moves). Viewed in these terms, the representing point will snake through its enclosing manifold along a very convoluted path. But we can make considerable progress if we eschew representing string in terms of its shifting $q(r)$ spatial locations and instead consider the components within its “frequency spectrum.” In other words (to articulate the policy in musical terms), we instead measure how much energy $E_i$ has become allocated to the string’s various harmonic modes of vibration (i.e., the strength of its overtones), as well as keeping track of where each mode currently falls within its own vibratory cycle (this is the current “phase” $\theta_i$ of the $E_i$ mode, employing the word “phase” in a somewhat altered sense). Considered in these $<E_i, \theta_i>$ terms, our string (in the absence of friction) will preserve the same $E_i$ values over its entire temporal career and will simply cycle though the corresponding $\theta_i$ “phases” in completely regular sine wave patterns. Furthermore,
these Eᵢ modes will *operate independently* of one another, for our adjustment in target quantities unravels the unhappy entanglements that complicates our q(r)-based descriptions of the string. Mathematically, our Eᵢ values correspond to curvilinear surfaces that carve up our original “chicken-like” manifold into layers resembling a very wavy form of filo dough. To be sure, it is rare that mechanical manifolds can be sliced up in such a tidy fashion (it can only happen if the system is exceptionally well-behaved) and such “foliations” are often hard to find even when they exist. 

Let us picture this “change in explanatory landscape” as a map from the original q(r) manifold to a “foliated manifold” corresponding to our <Eᵢ, θᵢ> decomposition.

In making this adjustment, we have managed to reduce the number of “degrees of freedom” under consideration to a *countably infinite set* (because the Eᵢ can be easily ordered). But wretched mortals such as ourselves will still experience difficulty in juggling even a “small infinity” of descriptive terms competently. Fortunately, with strings we can reasonably expect that we can adequately capture our string’s “*dominant patterns of behavior*” if we simply ignore all of its high energy modes and concentrate upon the lower part of its energy spectrum (modes E₁ to E₆, say). With this second approximative move, we will have reduced our number of effective descriptive variables from a countable infinity to twelve (6 Eᵢ + 6 θᵢ). Through this focus upon the dominant energy modes, we have made a second adjustment in “explanatory landscape,” winding up with a doughnut-like arrangement where the changes in our string become easy to understand (the “doughnut” is actually a sub-manifold hiding inside the larger chicken-like thing). It is astonishing how many of the great descriptive triumphs of nineteenth century physics represent variations upon this basic strategy for “reducing ones variables” (due mainly to Fourier).

In making adjustments in focal quantities in this manner, we plainly “avoid” the difficulties that attend any attempt to deal with our string in straightforward q(r) terms. However, in this essay, I largely wish to investigate “physics avoiding” adjustments in “explanatory landscape” that implement somewhat more radical strategies than the foregoing. In particular, we want to look at “landscape adjustments” that eschew all *direct consideration of temporal change* whatsoever.
truth, our shift from the string’s q(r) “time domain spectrum” (to employ the standard jargon of Fourier analysis) to its E, “frequency spectrum” incorporates some measure of the “avoid ‘time’ if you can” philosophy we shall study, but I plan to concentrate upon more extreme versions of the technique.

Here’s an example. Suppose that we run a stream of water molecules through a pipe containing a large rectangular block. Obviously, the locations of the molecules will constantly shift as they run into both the walls and one another. Will their movements ever settle into a less chaotic pattern? Certainly not after we first turn on the spigot (when the flow will manifest rapidly shifting transients) but perhaps after the water has been allowed to flow for a long period of time. What we’d like to see is a condition where, as soon as one molecule leaves its present location, a successor stands ready to immediately occupy that position, so that the whole ensemble gives the appearance of remaining in a so-called “steady state” (fresh waters are ever flowing upon us, but the overall condition of the water looks identical from moment to moment). Certainly, from a strict molecular point of view, no “steady state” condition is truly possible, simply because the successor molecules won’t be able to get to their new positions fast enough. But perhaps if we smooth over the molecular stream by some form of “averaging,” a true “steady state” solution might emerge as a “coarse grained” approximation to the molecular flow in the standard mode of continuum fluid dynamics. So our first adjustment in “explanatory landscape” is to shift from molecular dynamics to a smoothed over continuous flow. Under that adjustment, “steady state flows” do indeed become possible as long as the fluid velocity is not too high (otherwise we witness turbulence). From a mathematical point of view, such smoothed-over “laminar” flows are considerably more tractable than our original complex swarm of colliding molecules. Now if a laminar flow also happens to be “steady,” we can affect a second major adjustment in “landscape” that carries an extraordinary degree of “physics avoidance” simplification in its wake and also illustrates the advantages of “purging time.” The trick is this. Consider the shifting flow of our fluid past the block in temporal (or, as I shall usually dub them, evolutionary) terms. I have drawn time slices within such
a flow on the left side of the accompanying diagram. What will happen if we “shine a flashlight,” as it were, through all of these stacked temporal layers? Well, because the flow is in “steady state,” we wind up with a single coherent picture that captures our sense that “the condition of the fluid always stays the same despite the fact that fresh waters are ever flowing upon us.” In more standard jargon, we obtain a the picture on the right hand side of our diagram that registers how the geometry of the fluid streamlines has become altered in the presence of the obstacle. And such a revised portrait highlights the important qualities of the flow in entirely non-temporal terms. Mathematicians commonly portray this second change in “explanatory landscape” as a map from the time-registering evolutionary manifold on the left (comparable to our chicken-like portrait) to the non-temporal “base manifold” on the right where only streamlines are plotted without concern for which particles occupy such positions at what times. Reverting to our flashlight analogy, a “base manifold” captures the shadows that would be left on a photographic plate by our flowing fluid over a long time exposure. Because successor particles in the continuum representation immediately replace their predecessors, the “streamlines” get registered as dark regions to which no beam from our flashlight penetrates. If such a “base manifold” picture of a physical system can be successfully framed, we can usually “avoid” a considerable measure of the very disagreeable physics that we must otherwise track if we continued to reason in a straightforwardly evolutionary mode.

In fact, one of the great challenges of understanding turbulent flow is the need to uncover some valid “base manifold” registration of its hidden structures, presumably as some variety of statistical decomposition.

Under this “landscape shift,” a salient natural question becomes: how will the altered boundaries induced by the block alter the streamlines that we would otherwise witness in an unencumbered pipe? Or how will the flow pattern alter if the averaged velocity of the water far upstream becomes increased? At the same time, we have suppressed “evolutionary” queries such as: through what mechanism does a fluid particle within the flow manage to sense the distant presence of the pipe? Such “adjustments in focal question” are typical of the “shifts in explanatory landscape” we shall study. Alterations of these sort are generally called “shifts in a control parameter” and we will examine their character more closely later on.

All of the adjustments that I have here described as mappings from one “explanatory landscape” to another correspond quite tightly to concrete adjustments in the formulas with which physicists calculate (in the sequel I will provide a number of simple illustrations). But I have initially presented these modifications in
more abstract terms as “mappings between manifolds” because such a
“mathematician’s picture” of the alterations makes it easier to explain the basic
phenomenon we shall canvass in this essay (and the rest of the book). It is simply
this. Key descriptive words like “force” and “cause” are very mutable creatures,
able to acclimate themselves to virtually any “explanatory landscape” into which
they happen to be cast. In normal cases the word “cause” attaches itself to the
processes that affect temporal change within an evolutionary process such as a flow.
But purge “time” from our “explanatory landscape” as we have just done, and
“cause” will come creeping back in (like poison ivy in the old Leiber-Stoller ballad).
In this new environment, the word generally focuses upon how alterations in the
pipe’s boundaries (e.g., inserting a protruding block) will affect the streamline
patterns, in the manner introduced above. Later in the paper I shall argue that many
of the standard disagreements about “cause” that one encounters in the
philosophical literature can be better understood through attending to the word’s
mutable behavior as the background “explanatory landscape” undergoes adjustment.

But to understand these adjustments we first need to appreciate the
underlying rationales for effective “physics avoidance” that underpin these changes
in “landscape.”

(ii)

However, before we embark on this project, we should correct a
methodological misstep that has seriously impeded descriptive philosophy of
science for nearly a century—viz., the inclination to substitute logic-derived notions
for mathematical relationships of a more sharply diagnosed character. I have
elsewhere dubbed this propensity “the ‘Theory T’ Syndrome” in tribute to the
philosophy of science primers of my youth, which trafficked entirely in toy models
of a purely logical character (invariably dubbed “theory T” and “theory T’”). The
most basic problem with such approaches is that the descriptive tools they wield are
too blunt to adequately mark out the varieties of explanatory landscape we must
distinguish if we hope to understand how “physics avoidance” affects real life
scientific practice.

A well-known passage from the philosopher Peter Railton illustrates the
mushy logicist propensities I have in mind. As I see it, Railton is struggling to
isolate a natural and important idea that will be called a “(canonical) evolutionary
mapping” in the sequel (applied mathematicians generally prefer even more
rebarbative terminology such as “well-posed initial-boundary value problem of a
"hyperbolic character"--the underlying idea itself isn’t scary, but the label may be). Railton dubs his own target notion “an ideal D-N-P text” but actually seeks a natural stochastic generalization of the “evolutionary mapping” idea. What he has in mind is this (let us first ignore his “P”-rider and concentrate upon what a simple “ideal D-N text” is supposed to be). Divide the passage of time into a number of stages \( \Delta t_1, \Delta t_2, \Delta t_3, \ldots \) (mathematicians usually call these “time steps”) and then tie the physical conditions assigned to each step together by logical deductions from the data lodged upon earlier stages in conjunction with whatever fundamental “basic laws” apply in the discipline at hand. Such a hypothetical array of logically connected sentences is dubbed “a D-N text” where “D” stands for “deductive” (= the transitions are governed by logic) and “N” for “nomological” (= driven by scientific laws). As such, “D-N texts” comprised the basic ingredient within Peter Hempel’s familiar, logic-based account of explanation. In supplementing such “texts” with “P”s (for “probabilistic”), Railton further considers arrays of sentences connected by probabilistic deduction, in the sense that the “basic laws” now permit a number of following stages governing by some probability measure. A possible history for a target system simply becomes a branch running through such a D-N-P tree as pictured.

Against this backdrop, Railton is interested in a notion of “partial information” that bears some resemblance to the basic notion of “physics avoidance” that concerns us. He writes:

*Consider an ideal D-N-P text for the explanation of a fact p. Now consider any statement S that, were we ignorant of this text but conversant with the language and concepts employed in it and in S, would enable us to answer questions about the text in such a way as to eliminate some degree of uncertainty about what is contained in it.*

That is, statement S provides *partial information* about some fully developed D-N-P text. He writes:

*It is hardly novel to speak of sentences providing information about complete texts in this way: presumably we employ such a notion whenever we speak of a piece of writing .... as a summary, paraphrase, gloss, condensation, or partial description of an actual text such as a novel.*
Unfortunately, I know of no satisfactory account of this familiar and highly general notion ...nor can I begin to provide an account of my own making.

In my opinion, the chief reason that Railton cannot supply a sharper analysis than this is entirely due to the bluntness of the logical tools with which he works. Here’s a simple way to see what’s gone wrong. In the Oz novels, Glinda owns a big book in which all the events in the kingdom get instantaneously described as they happen. But she will surely require a non-denumerable number of pages in her book, because between any two pages separated by a time interval $\Delta t$, there must lie infinitely more pages (some of which might be very important from an explanatory point of view). But if we correct this problem by adding the requisite number of pages, what do we wind up with? Well, in the D-N case, not a “text,” but a temporally connected set of states that mathematicians call a “flow.” And the stages within this flow cannot be coherently linked together by “logical deduction” because no page possesses an immediate predecessor. We have, in fact, left behind any reasonable notion of linguistic connection and are simply looking at a “solution set” for a collection of evolutionary differential equations (we will later see that not all differential equations generate “flows” of the temporal character we seek—phrases like “evolutionary” or “hyperbolic” capture the specific behaviors we require). Such a notion is completely familiar to anyone who works in applied mathematics. Correcting the Glinda problem with respect to Railton’s probabilistic “D-N-P” variants conveys us in a parallel manner to the allied (but tricky) notion of a process governed by stochastic differential equations (we shall not be concerned with such processes in this essay, although they are very interesting constructions). As methodologists of science, we generally want to work directly with the notion of a “flow” generated by evolutionary equations, not with some inadequate surrogate framed in terms of “logical deduction.”

In fact, a variety of lapses trouble Railton’s account beyond our Glinda problem. No plausible notion of “law” should be identified with the full array of demands that come encapsulated within a set of differential equations but it is only the latter that can generate the “flows” that replace Railton’s chains of “logical derivation.” Likewise, Railton and his “Theory T” compatriots usually pay inadequate attention to “boundary conditions” (in the proper, mathematician’s sense, not the muddled notion that philosophers often cite in its stead) and “well-
posed modeling assumptions,” although allied mathematicians have long recognized
that such matters represent vital aspects of an adequately posed explanatory
framework. We should wonder why have so many philosophers of science
persisted in substituting feeble logistical surrogates for the much sharper
taxonomies that can be found in every adequate text in applied mathematics.

To this, a likely retort will be: “Because we philosophers of science seek
greater descriptive scope than your parochial emphases offer. Your differential
equations et al. may comprise the essential instruments of mathematical physics and
a few fields such as that, but we are required, as general methodologists, to speak to
‘explanation’ as it appears across a far wider array of contexts.” Indeed, eschewing
such “technical tools” is a virtue according to David Armstrong:

[Even if philosophers know no actual laws, they] know the forms which
statements of laws take ... It turns out, as a matter of fact, that the sort of
fundamental investigations which we are undertaking can largely proceed
with mere schemata of [the sort "It is a law that Fs are Gs].... Our abstract
formulae may actually exhibit the heart of many philosophical problems
about laws of nature, disentangled from confusing empirical detail. To
every subject, its appropriate level of abstraction.2

But let us turn a cold shoulder upon these apologetics and see if we can’t better
understand some important explanatory patterns encountered within everyday
physics practice if we borrow from the applied mathematician’s better equiped
diagnostic toolkit. If we can gain genuine insight through such means, we should
count ourselves satisfied. In any descriptive endeavor, it’s best to start small, for
the surest route to a miserable defeat is a grandiose assault. The siren call of “wider
generality” has caused more sailors to wreck upon the shoals of empty tautology
than any other single source.

In truth, no moral or intellectual imperative requires us “to speak to
‘explanation’ in wider contexts.” We have no a priori assurance that anything very
interesting is available at that rarified level of attack, in which the contours that
constrain a local “explanatory landscape” in vital ways may have become washed
away in the pursuit of “generality.” So in this essay I shall rather adopt as my own
benchmark of progress the question of whether we can usefully unravel the
behaviors of philosophically puzzling words such as “cause” through employing the
diagnostic taxonomies suggested by the applied mathematicians.

Insofar as I can determine, the most natural notion of “causal process” is that
of a continuously acting evolutionary process, such as the “flows” to which we
already appealed. However capturing this idea in precise terms is not easy and
wasn’t properly accomplished until PDEs were invented circa 1750. In lieu of that development, earlier writers were obliged to discuss such processes in what we now call “finite difference” terms, viz., by dividing time into small step sizes $\Delta t$ and approximating a continuously acting process by rough estimates of how the conditions on time slice $\Delta t_i$ (and before) should affect the conditions on time slice $\Delta t_{i+1}$ (in the general manner of one of Railton’s “ideal D-N texts”). In other words, rather than writing a causal process in infinitesimally focused terms (e.g., $\frac{\partial y}{\partial x} = \frac{\partial y}{\partial t} - 3y$), we express it as an approximation $\left(\frac{\Delta y_{i+1} - \Delta y_i}{\Delta x_{i+1} - \Delta x_i} = \frac{\Delta y_{i+1} - \Delta y_i}{\Delta t_{i+1} - \Delta t_i} - 3y\right)$. In so proceeding, we introduce the familiar temporal divide between “causes” (on $\Delta t_i$) and “effects” (on $\Delta t_{i+1}$). This “unnatural” feature of everyday thinking about causal processes supplied Bertrand Russell with one of the weapons he employed to banish “cause” from the halls of physics on the grounds of anachronism. But Russell’s reaction to this innocent semantic ploy strikes me as excessive, evidence of an inadequate sensitivity on Russell’s part to the difficulties of registering our intuitive conception of continuous process in precise terms, rather than demonstrating anything scientific problematic in the underlying notion of causal process itself. The allied notion of a continuously acting stochastic process provides an excellent comparison point. This notion is also quite intuitive--we naturally conceive of radioactive decay in such terms--but figuring out how to capture the conception of a “stochastic process” in precise terms represented one of the great mathematical accomplishments of the twentieth century (due to Norbert Wiener and many others). Insofar as I can see, Railton’s “ideal D-N-P text” simply represents an attempt to register the same underlying idea employing the “finite difference” tools of naïve commonsense. To be sure, we require the sharper analysis provided by the mathematicians if we hope to sharply diagnose the confusions that commonly come in the wake of employing approximative expressions as a means of capturing sometimes complex underlying ideas. But readers should always remind themselves, as we learn to employ some of the technical vocabulary of the modern theory of partial differential equations in this essay, that we are usually working with a set of commonsensical intuitive notions, albeit not ones that are optimally registered, with respect to precision, in the argot of everyday “causal process” talk. Mathematics does not represent some language of magical powers; it merely registers the data we put into it in sharpened terms. Indeed, I believe that most of the examples of “physics avoidance” we shall examine in this essay represent the implementation of some entirely commonsensical strategy that we frequently apply along other avenues of our thinking (and I shall often explain a mathematical stratagem by citing some rough
Accordingly, in the remainder of this essay I shall assume that Railton’s “ideal D-N text” is simply an attempt to register the natural notion of an “evolutionary flow” in the somewhat ham-handed “finite difference” manner of everyday discourse. By employing the “evolutionary flow” approach, we will find ourselves better prepared to diagnose in a sharper way the sorts of “physics avoiding” strategies that Railton labels as “condensations.” For if we merely analogize “physics avoidance” to motives for making long books shorter, in the mode of the old Reader’s Digest, we will be apt to despair (as Railton does in the quotations cited) that such policies will ever submit to trenchant analysis, for the rationales for “shortening a book” can prove as varied as the books themselves. But if we approach “physics avoidance” in terms of the strategic motives we follow in shifting “explanatory landscapes” in the “mappings between manifolds” manner outlined above, we will be to obtain a significantly richer appreciation of the changes wrought. To be sure, we shouldn’t expect to reach a full and complete taxonomy of “useful techniques for ‘physics avoidance’” (for clever scientists are forever discovering the damnest ways to make their descriptive obligations easier), but we can improve greatly upon where Railton has left matters. And we can eventually employ our new tools to better understand why the word “cause” wanders in the confusing ways that it does.

In thus aligning our endeavors to Railton’s, we should meticulously avoid a rash presumption that underlies Railton’s own thinking (and much of old-fashioned “theory T” thought in general). It is the premise that we somehow know, as a matter of a priori philosophical assurance, that all forms of successful explanation must qualify as “condensations” over richer evolutionary data sets. But we certainly do not know this. As we’ll see in section iv, although such expectations have often proved correct, in other cases, such hopes remain entirely “aspirational” and, in some important historical examples, have proved to be flat out mistaken. So we should take nothing for granted here. We want our investigations of “explanatory pattern” to provide us with accurate descriptive data to which our wider musings about the “nature of science” ought to remain responsible; we should not corrupt that “data” at the outset by making presumptions about the inter-relationships of real life “explanatory landscapes” that are not mathematically credible.

Once Railton’s “ideal D-N text” is precisified to become “flow generated by
evolutionary differential equations,” one immediately sees that most of what passes for “explanation” within physical practice will not fit the pattern suggested. And there are some simple formal tests that can be applied to the associated equations to determine whether the underlying “explanatory landscape” has shifted in some significant way or not. Every modern book on the central equations of mathematical physics (which most commonly come in the form of partial differential equations (PDEs)) instructs one to classify such formula into a variety of formal categories (hyperbolic, elliptic, parabolic, mixed). These “signatures of equational type” in turn often provide vital clues to the varieties of “reduced variable” strategy that are operative in the treatment at hand. Accordingly, in our “block in a pipe” case, the “evolutionary manifold” to “base manifold” portrayal of the fluid flow corresponds directly to a shift in the “equational signatures” of the formulas with which a physicist directly works. Specifically, the “signature” of the original evolutionary modeling equations are “hyperbolic” in character but they get replaced by “elliptic” formulae as the explanatory landscape shifts to a “steady state” centered approach.6

Let me hasten to again assure readers that this fancy technical talk doesn’t signify much beyond the fact that the “time” variable has been purged from the original equations through the invocation of steady state methodology (I shall illustrate concrete examples of these syntactic “purgings” in the next section). Throughout the remainder of this essay I will discuss such alterations largely in the cozy vernacular of methodological commonsense, abandoning to interested parties the easy chore of translating the technical terminologies found in the PDE primers into the user-friendly terms I shall favor.

Every work in mechanics shifts very swiftly from one type of equational format to another with very little explicit commentary provided on the attending “change in explanatory landscape” (which can be quite drastic). To this day, physicists are often quite cavalier about these transitions, learning to execute them through brute imitation rather according to any explicit justificatory rationale. It is mainly the contributions of mathematicians (due, in prime of place, to Jacques Hadamard) who have clarified these issues in the manner we require. In so doing, they have managed to make rational sense out of methodological adjustments that can appear, to the naked eye, as “Rabelaisian” and undisciplined. In truth, the simple “physical stories” one acquires from stock college-level physics primers are rarely ready for real life application until they have been significantly worked over by the applied mathematicians, who must convert the undiagnosed “idealizations” of the textbooks into precisified patterns of modeling that can fit the contours of a complex reality in a better adaptive manner (and not create glitches upon
In fact, another aspect of Russell’s argument that “causal processes represent anachronisms within science” traces precisely to the common physicists’ practice of lumping together all sorts of equational sets indiscriminately. But not all equation sets they study are created equal with respect to their incorporation of “causal processes.” Quite the contrary, “steady state” strategies for “avoiding ‘time’ if we can” gain their simplicity precisely through suppressing any detailed account of how the processes under investigation temporally unfold. When the “explanatory landscape” changes in this way, new sets of equations come into prominence, generally with altered “signatures” that reflect the reductive strategy at work. So we shouldn’t expect to find any clear expression of “causal process” captured within the replacement equations for steady state flow and, indeed, these formula are not of “hyperbolic” type. But this doesn’t mean that causal processes cannot be clearly registered within equational sets of the right “signature.” An excellent survey of these issues can be found in Sheldon Smith’s referenced paper.

If, as Russell in fact did, we search for “causal processes” indiscriminately within the standard “equations of mathematical physics,” we are likely to come up empty-handed, because such processes have been purposively purged from most such sets through wise “physics avoidance.” (two equations extracted from a textbook of differential equations may share no more internal characteristics than two people randomly selected from a telephone directory). So the proper response to Russell is simply: let the applied mathematicians tell you where you should look for the processes you seek.

Before we go further, I should indicate that the best applied mathematical match to Railton’s “ideal deductive text” is a “well-set initial-boundary value problem of a purely evolutionary character.” But let’s wait on “initial and boundary values” until we get a better handle on the notion of an “evolutionary process.”

In this essay, there are two main modes of “physics avoidance” we shall study:

1. appealing to time scaling considerations (e.g., equilibrium or steady state conditions) as a rationale for dropping terms from starting equations in a manner that alters their “hyperbolic signature”

2. adding non-hyperbolic equations to an equation set in the form of higher scale constraints.

We shall look at (1) first.
As I have already indicated, most of the equations commonly employed in everyday physics treatments are not of the “evolutionary” character that best captures the intuitive expectations behind Railton’s “ideal D-N texts.” Extracting the information one wants about a physical system directly from some evolutionary PDE modeling is often nearly impossible and so practitioners have accumulated a large library of evasive tricks for avoiding difficult and untrustworthy computations in very much the same mode as would have improved the lot of our Mormon pioneers. For example, a simple and common evasive strategy of this sort is to concentrate upon predictable episodes of achieved equilibrium, rather than attempting to track every detail of an evolutionary process. We shall study a formal example in a minute, but here’s a homely illustration of the general idea. Jack and Jill have just fallen from the top of a steep hill; can one predict where they will land? “Sure--somewhere in the basin at the bottom.” In thinking thus, one is implicitly assuming that the kinetic energy they acquire in tumbling down the hill will be lost in their encounters with rock and twig and so they will come to equilibrium (= bruised rest) in the potential well at the hill’s base. In reasoning so, we jump immediately ahead to the anticipated equilibrium state and forego any attempt to compute their paths of descent from hilltop to basin. In fact, the schemes one might employ on a computer to track their evolutionary path are notoriously prone to error: such calculations may launch the children into space rather than conveying them correctly down the hill. True, in avoiding such error-prone reasoning, we lose the capacity to predict exactly where around the hill’s circumference they will land, but that detail may prove relatively unimportant. Such an exploitation of equilibrium offers a good example of the kinds of strategies for physics avoidance that we want to study.

In a suitable context, the reply “because the potential energy is at a local minimum” may qualify as an acceptable “explanation” for “why did Jack land there?” (sometimes one will want a fuller account).

Let’s now see how allied adjustments in explanatory frame manifest themselves in more formal terms. I’ll start with a classic example of “physics avoidance,” due originally to Euler (in truth, a fair measure of “avoidance” has already been applied in our original modeling, but I’ll ignore these technical
complications for the time being). Consider a narrow band of
an elastic metal loaded with a weight W at the top with its end
points constrained to two fixed positions (although the weight
W is allowed to fall freely under gravity within its casing). We
can construct a suitable equation for this situation by building
upon the frame of Newton’s “\( F = ma \)” (more exactly, we are
following the “Cauchy recipe” for a simple continuous body
that I outline more fully in the book’s Appendix). Specifically,
we obtain the formula
\[
EI \frac{\partial^2 x}{\partial y^2} + Wx = \rho \frac{\partial x^2}{\partial t^2}
\]
which describes a battle between the efforts of the falling
weight to bend each local section \( y \) of the strut and \( y \)’s efforts to resist such bending
to due its intrinsic elastic response. More exactly, our equation postulates that the
falling weight exerts a torque \( Wx \) upon each section \( y \) according to the degree \( x \) to
which \( y \) fails to lie directly underneath the weight \( W \). At the same time, point \( y \)
resists this turning moment according to its degree of local curvature (by the amount
\( EI \frac{\partial^2 x}{\partial y^2} \) where “E” and “I” capture the strength of the resistance and the
geometry of the strut’s cross-section, respectively). If \( EI \frac{\partial^2 x}{\partial y^2} \) and \( Wx \)
counterbalance each other exactly, point \( y \) will remain inert; if not, it will display a
transverse acceleration \( \frac{\partial x^2}{\partial t^2} \) at each point \( y \) according to the local density \( \rho \). Note
that this equation only captures the physical processes active inside the strut; it
does not describe the rather different factors that keep the endpoints of the strut
fixed. The latter conditions get crudely codified into the problem’s boundary
conditions, of which more later.

Despite the suggestions of D-N model thinking, it is easy to see that our
governing differential equation shouldn’t be identified with “law” in any obvious
sense. The relative strengths of our two critical terms (\( Wx \) and \( EI \frac{\partial^2 x}{\partial y^2} \)) can
eventually be traced back to law-like sources (viz., Newton’s law of gravitation and
the so-called “constitutive equations” for the material in the strut—see Chapter 3),
but this path is not immediate and meanders through a variety of special modeling
assumptions and approximations. For simplicity, we might say that the nature of
these two forces are “determined by basic laws,” but the differential equation itself
also reflects a variety of modeling assumptions as to how these two forces operate
in the present context.
In any case, $EI \, \partial^2 x/\partial y^2 + Wx = \rho \cdot \partial x^2/\partial t^2$ is a classic example of an evolutionary wave equation. If, at some time $t_0$, we happen to know both the displacement $x$ and its velocity $\partial x/\partial t$ at every location $y$ up and down the strut, we will be dealing with a classic initial-boundary value problem that determines, in accordance with the initial conditions of the problem (= the $x$ positioning and $\partial x/\partial t$ velocities at all $y$ for some starting time $t_0$), how waves will move up and down the strut in accordance with our governing equation. At the bottom of the accompanying diagram, I have hammered a little bulge into our band at $t_0$. As time passes, this disturbance will split into two smaller lumps that travel outwards in opposite directions until they hit their respective boundary clamps, at which point they will turn upside down and reverse their direction of travel (as indicated by the little black arrows). These pulses will also alter their shapes somewhat as they travel, due to the fact that the $Wx$ torque is stronger by the clamped vise boundary than near the weight. Note that in this diagram, its “boundary conditions” (= the fact that the strut is held fixed at its two end points at all moments) are registered as vertical lines extending forward in time. Such conditions will not affect the way our pulses move until such time as they bump into the clamps and reflect backwards.

With equations of this ilk, there is an associated collection of “sound cones” (or “characteristic surfaces”) that register how fast disturbances can travel each local section $y$ of the strut—these “top allowable speeds” may differ at different sites along the band, depending upon its local constitution (differing top permissible speeds are indicated by the degree to which the local sound cones open up). In our one-dimensional example, such “characteristic cones” appear as little X’s (possibly curved), but in higher dimensions they assume the form of generalized cones (possibly wavy). As such, “characteristic cones” are familiar from relativity theory, appearing as the celebrated “light cones” in that subject, indicating the top permissible speed for any physical process. In the case at hand, compact initial pulses will travel through the strut at top speed, leaving “vacuities” in their wake, but in more general circumstances pokier disturbances will spread out.
behind the moving wave front.

Thanks to Hadamard, mathematicians now recognize that purely evolutionary equations always carve out little cones like this and that this capacity is closely connected to the formal structure of its differential operator \((\rho \cdot \partial^2 \partial \partial_t - EI \partial^2 \partial \partial_y - Wx)\). Indeed, there is a little polynomial \((\rho \cdot x^2 - EI y^2 - Wx)\) hidden there that captures this cone-making capacity and draws the requisite “characteristic curves” in our space-time diagram. When this happens, mathematicians say that the equation is “of hyperbolic signature” (the name “hyperbolic” traces to the somewhat accidental fact that, for equations of degree two, the associated polynomial happens to be an equation for a hyperbole). But not all differential operators do this—if we render the last two terms positive \((\rho \cdot \partial^2 \partial \partial_t + EI \partial^2 \partial \partial_y + Wx)\), the operator will no longer pass the mathematician’s test for “hyperbolic character” and no (real) sound cones will be drawn.

The salience of these rather fussy observations for our own concerns is this: if the equation set employed within a physical modeling is not of thoroughly evolutionary (= hyperbolic) character, it is a sure symptom that some substantive form of “physics avoidance” strategy is active in the modeling and that we are confronted with the applied mathematics analog of one of Railton’s “condensed texts.” Hadamard’s observations will allow us develop a quite sharp understanding of how such deviations operate and the respective advantages and disadvantages they offer (in contrast to Railton’s inability to provide a “satisfactory account” of his “condensations”). Later on, such distinctions will help us unravel some of the oddities of the word “cause”’s rather eccentric behavior.

Put another way, not all differential equations are created equal! Differences in the “signatures” of their operators supply a secure symptom that such equations are playing rather different explanatory roles within science. Hadamard particularly emphasized the fact that, if modeling equations are not purely evolutionary, they are probably not suited to “initial-boundary value problems” at all. In point of fact, most treatments found in physics books are rarely of true “initial-boundary value” type, although they may superficially appear that way (due largely to “time mimics” of a sort we shall consider later in the essay). Many substantive confusions in philosophy of science trace at root to the error of parsing non-initial-boundary value problems as if they comprised the real article.

In fact, there are few natural questions about struts that can be easily posed within a strict initial-boundary format (my earlier narrative of inserting pulses into struts is hardly the kind of problem that directly concerns civil engineers). Accordingly, let’s examine a celebrated stratagem for very effective “physics
avoidance” with respect to our strut. The basic idea is one of the many glories we owe to Euler, for the ploy neatly suppresses any need to consider a proper “initial-boundary problem” at all. Once we understand its strategic contours from an intuitive point of view, we will be able to appreciate how Hadamard’s distinctions help understand the exact nature of the explanatory “condensation” at hand.

Evolutionary modelings of the sort just considered generate suitable temporal “flows” of exactly the character that Railton anticipates in a “D-N text” (modulo correcting his logic-corrupted intuitions into a “flow”). But trying to augur real life strut behavior using evolutionary models of this ilk often represents a very unreliable affair, requiring huge amounts of computation to predict how influences will pass through the strut with even passable accuracy. So Euler asked, “But do we really need to know all of those details?” If we are interested in how our strut will respond to earthquakes, we probably will but, for most engineering purposes, we only want to know whether our strut can bear the burden of the maximal weight we expect to place upon it or whether it will sag under such loading. As with Jack and Jill’s situation, we are largely interested in a future state of anticipated equilibrium: will the strut sag or remain upright when it finally stops wiggling? And there are three possibilities: A, B or C, where we hope that the answer will be B (the strut will support the weight W we plan to place upon it without sagging). And so Euler proposed that we circumvent the computational obstacles posed by a straightforward evolutionary modeling by simply dropping the time term $\rho \cdot \partial^2 x / \partial t^2$ from the above equation and considering this truncated replacement instead:

$$EI \cdot \partial^2 x / \partial y^2 + Wx = 0.$$ 

(In the sequel, I will dub this formula “Euler’s reduced strut equation”). The rough justification for this maneuver is that, as happens with Jack and Jill, frictional influences will eventually drain all energy of movement from the strut, leaving only the factors $EI \cdot \partial^2 x / \partial y^2$’s and $Wx$’s to balance each other however they can in equilibrium. And Euler observed that, as long as the weight W remains below a certain critical value $W_c$ (determined by $EI$), the strut will be able to bear its load in the admirably upright manner B. But as soon as the weight exceeds $W_c$, the A and
C modes of sustaining a $\text{EI} \partial^2 x/\partial y^2$ to $Wx$ balance become available in which our strut dangerously sags (as also occurs with Jack and Jill, this technique doesn’t allow us to predict which of these two final shapes the strut will adopt, but that doesn’t matter: sagging is bad whichever way it leans). In the usual jargon, the equilibrium states open to the strut undergo a bifurcation once the critical load $W_c$ is passed (additional loading can produce further bifurcations in responsive pattern).

Applying our polynomial test to our newly reduced differential operator $(\text{EI} \partial^2 x/\partial y^2 + Wx)$, we generally find that its “evolutionary character” becomes lost and the associated “sound cones” have become imaginary. Moreover, our new equation now only accepts solutions of a restricted “analytic function” class, a fact of considerable “physics avoidance” salience which we shall consider later.

In essence (I’ll discuss some minor technicalities later), Euler’s “physics avoidance” ploy shifts our problem from being an initial-boundary value problem to a pure boundary value problem. In the case at hand, the “boundary” just consists of the two clamps at each end of the strut, but, with higher dimensional modelings, the term carries its intuitive meaning—e.g., with a bongo drum, its boundary consists in the tacking around its perimeter. Once again, the physical processes responsible for the boundary behavior are not captured within the standard differential equation for the interior membrane (which a two-dimensional wave equation rather similar to that for our strut). Generally, a fair amount of “physics avoidance” is active in the justification of “boundary conditions” for a problem, but we won’t pursue this issue here. Explanatory patterns that invoke pure boundary problems are quite different in their intrinsic character from the (almost) purely evolutionary situations we have been examining. Hadamard underscored this fact in Lectures on Cauchy’s Problem:

"Indeed the two sets of conditions play quite different parts, from the mechanical point of view, and are often and rightly called by different names. We hitherto, according to the geometric formulation of the question, have used..."
the term “boundary conditions” indiscriminately; but if now think of the mechanical meaning of the questions, we shall be lead to give to conditions 18 the name “initial conditions,” the name “boundary conditions” being kept only for conditions 18’ which correspond to the extremities of our string.

A good way of appreciating the differences between these types of problem is through computational technique. Let’s figure out how a computer might figure out how the pulses with move through our strut within the somewhat artificial (but genuine) initial-boundary value problem we considered with the strut. Let’s play our temporal arrangements on a piece of graph paper and note that we start with knowledge of the data (= the facts obtaining there) on three sides of a rectangle, with the extra blessing of knowing the strut’s initial velocities along the bottom “initial data” line. Our computer will then fill in our graph by filling in the squares ahead according to how it has already computed the squares to the left and right and behind it. The technique simply marches up the graph paper like an old Roman phalanx (which is why techniques like this are called “marching methods”). As such, computations for pure evolutionary problems are pretty straightforward. We observe that the main portion of our evolving pulse will follow the “characteristic curves” (= “sound cones”) laid down by our governing equation until it reaches the material’s boundaries, at which point we will need to employ our boundary data to push our calculations forward (note that the interior equation itself doesn’t tell us how to calculate at the boundary--that decision depends upon the physical processes active at the boundary).

But now let us consider a pure boundary value problem, such as the rest state of a drumhead that has been loaded with rocks. Although at first blush, it looks like we have more data on all sides to work with, this isn’t really isn’t true because we don’t know the angles with which the skin will lift away from the boundary (that slope data is the formal equivalent of the initial velocities in the strut.
All our computer can do is *guess* how the skin will lift off the edges and *hope* that its calculations will agree when they meet in the center (this computational situation much resembles the problem that the Western Pacific and the Union Pacific faced in trying to construct a transcontinental railroad working from two sides of the country). In fact, our computer is almost certain to guess wrong at first and so it will need to try again with an improved set of guesses until it finally hones in upon a suitable across-the-drumhead match up. Computational techniques like this are called “iterative methods” or, in a more homey fashion, “shooting methods” because they operate as one does in archery, adjusting the angle of ones bow after seeing how far ones arrows fall from the target.

These differences in computational technique are closely tied to the *removal of time* from the equilibrium problem, although this may not seem obvious at first glance. When we set our strut wriggling by inserting a original pulse, it takes a certain degree of time before that pulse and its descendants can adequately explore every inch of the terrain. But our assumption of equilibrium presumes that enough probing of this nature has occurred taken place during the brief interval of the strut’s relaxation time (more on this in a moment). In other words, enough time has passed that we can assume that processes originating in different parts of the strut have had time to come to accommodation. Hadamard noted a particularly piquant mathematical expression of this “enough time for accommodation.” He noted that equations of “elliptic signature” (e.g., the time-terms-dropped-for-equilibrium version of the drumhead equation) only accept analytic functions as solutions (mathematicians say that such equations “smooth out the data” away from the boundaries). Now analytic functions per se appear to be rather harmless critters (most ready-to-hand expressions are of this type), but they secretly embody a reproducibility property allied to that of a flatworm: from any little piece their behavior everywhere can be constructed by so-called “analytic continuation.” In normal evolutionary circumstances, this is an unwelcome property, for one expects that two folks could hammer on two ends of a beam in completely independent ways, with no blending of results until the pulses from each meet in the center of the beam. But if the wave equation acted like an elliptic equation, the pattern of the lefthand hammerer’s
pounding could be deduced from that of the righthand hammerer, without any signal passing between them. As Hadamard put it:

But such functions [= generic solutions of hyperbolic equations] differ from analytic ones... in lacking one of their classic properties: the extension of a function ... from one part of its domain to a neighboring one is determined. [But a non-analytic] function, being given in (0,1), may be extended to (1,2), for instance, in an infinite number of ways.

This capability of hyperbolic equations is closely tied to the “sound cones’” they inscribe, for they limit the speed at which messages can be sent from one region to another. It is fairly clear that our intuitive understanding of a “causal process” anticipates these limitations on cross-system coordination. In contrast, an proper appreciation of the typical “physics avoidance” lurking in the background of elliptic equations makes their analytic “reproducibility” properties unsurprising because we should only apply such equations to situations where we can safely can assume that our target system has had enough time to bring all of its active influences into the kind of coordinated balance that we call “equilibrium.”

Hadamard employed isolated pulses as a means of exploring how the causal processes embodied within a set of purely evolutionary modeling equations will behave, but he didn’t attempt to break allied generic processes of this sort into the movement of pulses or allied forms of decomposition (this kind of representation is possible in linear circumstances, but not generally). But that, it seems to me, is sort of mistake that philosophers following Reichenbach make in attempting to define “causal process” as the transmission of “marks” (= pulses). No: “causal processes” of the expected “progressive” character are inherently enshrined in the differential operators that are built up in the right way, without the intruding hand of “physics avoidance” and their basic symptomology lies in their formal hyperbolic character. But one shouldn’t attempt to replace the patient itself by her symptoms.

[Add some Remarks on Russell and other “causal process” deniers].

(iv)
Aspirational hopes and confusions. None of these issues just scouted would occasion conceptual problems except for the surprising semantic propensity of words like “time” and cause.” to come sneaking back into our thinking about a problem even when we have resolutely decided, à la Euler, to eschew attention to “time” and “process” in our thinking (like the cat in the old song, they keep returning even though we’ve presumed that they’re banished). There are many ways in which this can happen, included the rather complex reappearance of “time surrogates” in some of the optical situations we shall examine in later chapters. With struts, “time” commonly reenters the equilibrium-dominated scene as follows. It is natural to worry, in planning a building, whether its struts can support the maximal expected weight we plan to place there (as when we install a large record collection in an old apartment). Such considerations lead us to associate our reduced strut equation with a so-called “bifurcation diagram” which I’ve illustrated in both an “artistic” way (on the left) and in the usual, drab textbook mode (on the right). Such charts indicate, as we gradually increase the load W on the upper end of the strut, that a qualitative change (or “bifurcation”) in possible equilibrium shape appears after we pass the critical load \( W_c \), whose A and C branches correspond to bands that bend to the left or right. In fact, if we increase the load W enough, new qualitative changes in strut shape will appear, corresponding to a further critical load \( W_{c'} \), as indicated by the second branching in our right hand diagram.  

The mathematician’s usual name for figures of this sort is control space diagram (the increasing weight on the upper end is the “control variable”). But if we fancy that we are slowly loading the upper boundary with increasing weights up to the moment when the whole affair buckles, we will have tacitly reintroduced a temporal dimension into our considerations, through an identification of the proper control variable \( W \) with passing “time.” But this whole way of thinking sits on top of an Eulerian strategy in which we avoid temporally tracking the strut by looking ahead to its inevitable equilibriums (as we’ll see in a moment, hyperbolic trackings
tend to be quite unreliable predictively and it’s better to evade them if possible). Our discussion plainly shows that this new “time” quantity is quite different than the “time” we have eliminated, for it actually represents a “manipulation time” that captures the driving rate at which we increase the weight on the boundary—it does not represent the natural “autonomous time” that the strut would follow in, e.g., subsiding into equilibrium. As such, our manipulation time scale $\Delta t^*$ must run at a much slower rate than the internal \textit{relaxation time} $\Delta t$ required for the strut to damp out any internal wiggling. In fact, Euler’s basic stratagem will supply reliable predications for our heavy-load-in-an-apartment problem only if we bring in the records rather slowly, for the strut must always be permitted enough “autonomous time” to halt its wriggling before we increase the upper load further (if the load is increased too speedily, we essentially find ourselves in the arena of dynamic loading such as earthquake engineers investigate, in which struts will buckle considerably earlier). To provide a picture of how these two time scales operate in tandem with one another, we should sandwich in insert little pictures of the “fast time” relaxation processes occurring betwixt the “times” marked within our “slow time” bifurcation chart. So “two time scales” are implicit within our control space diagram: the coarse or “slow” time scale is actively reflected within the chart, whereas the relaxation times required to reach equilibrium have become invisible through being squashed into infinitesimal briskness.

If one doesn’t pay attention to these two time scales, one can fall into simple confusions, such as “if beam behavior is deterministic, where does the indeterminancy enter that prevents us from predicting whether the strut will sag to the left or right?” The answer traces to the fact that we will need to open up the suppressed “fast time” behavior to make such a prediction (as well as specifying its boundary region physics in greater detail). Further subtleties involving suppressed time scales are latent in our strut problem which sometimes puzzle beginning students but I’ll delay considering these for a little while.

In any event, as our chief temporal focus shifts to a control or manipulation variable (i.e., $W$) in this fashion, it is not surprising that the word “cause” opportunistically attaches itself to the changes wrought through manipulation of the strut’s upper boundary. In a pure initial-boundary value problem, in contrast, we are largely concerned with the interior causal processes driven autonomously according to unfolding of real time, rather than the “manipulation time” that is determined by the personal whimsies of how swiftly we haul records into our
apartment. The well-observed remarks by Jim Woodward and others with respect to the frequent alignment of “cause and effect talk” with counterfactuals reflecting a system’s responses to potential boundary value manipulations reflects this natural reorientation of the word “cause” in circumstances where “physics avoidance” of an Eulerian stripe is being practiced. Critics have sometimes objected to Woodward’s theories on the grounds that it doesn’t credit “causation” with the intuitive sense of “trackable process” that we often expect from the word. And it is certainly true that the word “causal process” typically fastens upon evolutionary developments of a wave-like character when a true initial-boundary value problem is under consideration. But prudent “physics avoidance” commonly shifts our interest to other forms of explanatory setting, even if we often fail to diagnose these shifts properly (the frequent appearance of deceptive “initial-boundary value mimics” will prove a frequent theme throughout this book, sometimes arising in surprising complex ways). It strikes me as unreasonable to claim that hyperbolic or elliptic settings supply “the real semantics of ‘cause’” because we typically imbibe many flavors of “causal” focus along with the earliest paradigms we acquire of how a satisfactory physical explanation should be constructed (true answer: there isn’t just one pattern). Instead of criticizing Woodward’s suggestions, it is philosophically far more useful to trace the natural adjustments of focus that occur when one abandons paradigmatic initial-boundary value situations in favor of other forms of explanatory matrix according to the stratagems that prudent “physics avoidance” offers.

Although I have glibly appealed to a “relaxation time” for our strut equation, perceptive readers will have already noted that our formula, as it stands, doesn’t support such claims. This is because we failed to incorporate any form of frictional processes within its internal workings. Without some energy dissipation, our strut will wobble forever, passing waves up and down its column, never settling into quiescence. Correcting this lapse in our modeling in a physically plausible manner is not an easy task but let’s consider how the simplest form of frictional repair might work. I have in mind the ploy of adding a simple frictional term to our evolutionary equation that allows it to gradually damp movement in the strut. For example, we might insert the term “-κ ∂x/∂t” to obtain ρ \cdot ∂x^2/∂t^2 = EI ∂^2 x/∂y^2 + Wx - κ ∂x/∂t. When we “turn off the time variable” in Euler’s manner, we still obtain his reduced equation (“-κ ∂x/∂t” now drops out along with the “ρ \cdot ∂x^2/∂t^2”) but with a more plausible hope that the solutions of this new formula will asymptotically approach our A, B and C solutions asymptotically as time \to \infty. In such circumstances, our “Jack and Jill” strategy for “physics avoidance” by focusing directly upon its anticipated equilibrium states will be supported as a limiting
condition of the original evolutionary modeling.\textsuperscript{14}

In point of fact, simply adding a term like $-\kappa \frac{dx}{dt}$ can’t always provide the results we want, because (\textit{inter alia}) the damping provided might prove so strong that the strut will be forced to stop moving before it obtains its expected mode of equilibrium (in other words, the waves moving up and down the strut get eradicated before they’ve had enough time to explore the full length of the band). In point of fact, I don’t know of any plausible evolutionary modeling for our strut that can ratify Euler’s “drop the time terms” strategy in all foreseeable cases (that doesn’t mean that there aren’t such repairs; I simply don’t know of them). Accordingly, our presumption that the successes of Euler’s equilibrium-based strategy can be founded upon an underlying evolutionary modeling of the sort sketched can only be regarded as an \textit{aspirational hope}, not an established verity. In studying physical practice as it really is, one should be prepared to discover many cases where “physics avoidance” of pattern X at first appears to piggyback upon some deeper evolutionary modeling of type Y, but where it is actually impossible to substantiate this claim with any concrete form of Y pattern modeling (see chapter 3 for many examples of this). Indeed, there are many historical examples where the real story of why “physics avoidance” strategy X works turns out to trace to underlying factors Z quite different than initially contemplated (the original Y acts as an “\textit{explanatory mimic}” with respect to Z). To be sure, \textit{some} story of why tactic X works ought to be forthcoming, but it may differ considerably from what one \textit{a priori} anticipates.

Within classical physics, one needs to be especially wary of assuming the existence of a supportive evolutionary story of an exclusively \textit{classical} stripe, because appeals to \textit{quantum mechanical process} are sometimes required to rationalize the “avoidance” strategies it employs. Or, to put the matter more exactly, one shouldn’t always expect to uncover an “underpinning evolutionary model” of a classical stripe, for the real life mechanisms that, \textit{e.g.}, allow matter to remain stable seem to be intrinsically quantum mechanical in their operations.\textsuperscript{15} Allied difficulties arise whenever \textit{thermodynamic processes} such as friction must be invoked in order to persuade a seemingly “\textit{purely mechanical system}” to approach equilibrium or push shock wave structures past their moments of first formation. Such non-trivial difficulties of “physics avoidance” support tend to force a rather peculiar “\textit{theory facade}” structure upon classical mechanics considered as a conceptual system closed unto itself. This topic will prove of central interest in essay three.

In other words, when we consider “physics avoidance” in this book, it is often invoked in applicational circumstances where we lack any clear sense of what
we’re actually avoiding. As just observed, the correct rationalization for the “avoidance” often proves of a significantly different origin than we originally presumed.

These remarks, I think, underscore a serious defect within the “theory T syndrome” approaches we criticized earlier. I have thus far recommended that Railton’s logic-centered “ideal D-N text” be replaced by a “purely evolutionary modeling” in Hadamard’s mode and that his “condensed texts” correspond to the altered equational frameworks we create as suitable stratagems of “physics avoidance” are applied (Railton might unwisely resist this assimilation, for reasons I’ll discuss in a moment). But he seems to presume, upon some nebulously a priori basis, that evolutionary modelings comparable to his “D-N and D-N-P texts” must underlie every successful explanatory pattern within physics (every “condensed text” must have its “ideal expansion”). But this is a rash doctrine, in light of the many historical cases of successful “physics avoidance” whose contours have proved unsupportable in the general fashion of our difficulties in finding plausible frictional underpinnings for Eulerian equilibria. And there have been many wise physicists (Duhem and Bohr, inter alia) who have questioned whether their subject can be adequately founded upon purely evolutionary modelings (there are always tacit issues of suppressed time, length and boundary scales that need to be opened up).

To be sure, there are also physicist who believe quite the opposite—Robert Geroch, for one. And there is no doubt that it would prove intellectually pleasing if every real life puzzle connected with deviant explanatory pattern could be cleared up on purely evolutionary grounds. I think this must prove an empirical matter, dependent upon what Mother Nature decides to place upon our plates with respect to workable explanatory schemes (she doesn’t seem to much care whether we find her workings “intellectually satisfying” or not). We smother many substantive questions about viable methodology into unhelpful invisibility if we wax a priorist in Railton’s vein.

The aspirational presumption tacitly ingrained in Railton’s thinking reminds one of Armstrong’s confident claim that he “knows the forms which statements of laws take,” when, plainly, he knows no such thing. The source of this hubris, it strikes me, traces to the inclination of the “theory T” school to substitute logical deductive relationships for notions of “solution flow” and the like. Once all inferential practice in physics becomes washed out as an indiscriminate “Oh, those guys are just applying Modus Ponens and Universal Instantiation” to their axioms,” we will become inclined to fancy, in Armstrong’s naive way, that “we
“know the basic forms that statements of laws must take.” But this is wrong; as we’ve seen, not all equations of mathematical physics are created alike! But we need to follow Hadamard and pay any attention to the shifts in equational signature that serve as important symptoms of “shift in explanatory landscape.” If we spare ourselves the discipline of struggling with the complexities of explanatory practice as it is found in real life science, attention to logical structure alone is apt to convert aspirational hopes into Firm Verities that underpin all the workings of our universe (and the legions of Lewisian “possible worlds” that allegedly encompass it). If we do this, we’re not learning from Nature, but preaching to her prematurely.

I fancy that a “metaphysician” such as David Armstrong considers it beneath his dignity to fuss as much as we have about the precise formal role that friction plays in bringing in conveying a strut to rest. He will tempted to not identify his “ideal D-N texts” with anything so precise as a purely evolutionary flow, but to generalize the scheme so that the role of “time” etc. in the explanatory proceedings becomes quite nebulous. UNCLEAR However, friction is first cousin to dissipation and then to information loss. And the latter certainly forms a central strand within the tangled docket of issues that trouble quantum theory under the headings of collapse of the wave packet and renormalization. I do not know how to sort those matters out (I doubt that anyone now alive does), but I heartily suspect that some of the untangling will require careful scrutiny of the richly varied patterns in which feasible human computation can relate fruitfully to a rather difficult universe. A patient and unbiased study of how different “explanatory landscapes” might potentially relate to one another strikes me as an important first step in approaching such issues. But the generalist propensities of the “theory T syndrome” school discourage us from investigating such matters in the requisite detail. I doubt we should trust a philosophical dogmatism whose firmness sometimes appears to trace to little more than the fact that mastering elementary logic is considerably easier than learning the modern theory of PDEs.

In fact, the propensity to think that an “analysis of ‘explanation’ or ‘cause’” is inadequate unless it embraces every known use of these words is founded upon philosophical presumptions with respect to concepts and meanings that I have argued (at fulsome length in my WS) are deeply mistaken. Rather than supplying “conceptual analyses” of a generic stripe, it is more vital that philosophy help us unravel the intellectual complexities that often arise along the courses of our ordinary affairs. Within real life explanatory practice, “initial-boundary condition” mimics often arise that superficially look like the real article, but actually operate upon a more subtle, “physics avoidance” basis. Sometimes such “mimics” generate
serious forms of misunderstanding, several of which we shall investigate in some
detail in this book. Washing out the crisp distinctions required to unravel these
matters through a quixotic quest for “generality” can do much harm, as happens, I
think, when philosophers replace the mathematicians’ precise notions of “initial
condition” and “boundary condition” in favor of mushier surrogates.

Before moving on, we should observe that there are several further instances
of compressed time scales implicit in our strut treatment. As already remarked,
most natural problems concerning struts are not true initial-boundary value
problems. The reasons for this divergence traces to some rather subtle forms of
“true time” avoidance at work in our problem. Normally we do not expect our strut
to start wiggling through being plucked in its interior (as required in a true initial
value problem), but as driven somehow by the loading on the upper boundary. So
one expects that our initial boundary manipulation (when we set the weight W there)
must have released some kind of pulse into the upper end of the band. But in the
boundary conditions we assigned to our problem, we modeled that upper boundary
as perfectly fixed at all times of interest. That is, we seem to be claiming that the
upper weight never moves but still conveys movement to the strut at the very same
time! We resolve this contradiction (which is quite typical of the “double think”
implicit in many forms of commonly assigned “boundary conditions”) by observing
that the time scale $\Delta t'$ in which the moving weight affects the string is much shorter
than even the “fast relaxation time” $\Delta t$ required for attaining equilibrium.

In fact, the boundary value/”time scale” anomalies for
our strut run deeper than this, once we recall that it is through
the upper weight’s falling that the torque on the position $y$
becomes generated (as the factor reflected in our equation’s
“$Wx$” term). But position $y$ can lie relatively far down the strut
from the weight W--how does the latter manage to send an
appropriate signal to position $y$? The answer lies in the fact that,
although we have allowed our beam to flex laterally in the $y$
direction, we have seemingly not allowed it to compress in an up and down mode--
in essence, we have treated it as a rigid rod in the $x$ direction (as we’ll see below,
non-compression requirements of this sort should be treated as constraints and
viewed as further departures from a canonical evolutionary framework). This
neglect of x-direction detail is hidden in the stratagem of modeling our strut as a
lower dimensional object, a “physics avoidance” ploy that commonly induces
conceptual anomalies of this sort. A proper excuse for treating these compressive
interactions as rigid can be laid at the feet of the fact that such influences act upon a brisker time scale than is needed for the transverse vibrations.

Here we might observe that it is most likely that a moving strut will shed most of its kinetic energy through minute movements of its upper and lower clamps (rather than by interacting with the surrounding air, in the mode of the $\kappa \frac{\partial v}{\partial t}$ term postulated above). But handling these details in a properly evolutionary mode requires that we model the strut/clamp interaction in a more substantive way than merely as a standard “boundary condition.” It is the suppression of this kind of “fast time” detail that injects the “indeterminancy” into our “slow time,” manipulation-driven treatment above.

(v)

Mixed Level Explanations. Our discussion so far can be viewed as an effort to trace a connective path between the facets of “causal process” that Sheldon Smith has emphasized to those central within the work of Jim Woodward. This “road from Smith to Woodward” is littered with the castoff remnants of forms of evolutionary tracking that can skirted through suitable injections of clever physics avoidance. XXXX of initially paved with the natural deemphasis of evolutionary focus that arises as profitable “physics avoidance” emphasizes the data to be gleaned from manipulation of a target system from the outside, as registered within a “control space” diagram. But we truly arrive in Woodwardian territory when we study a second mode in which a set of modeling equations can depart from a purely evolutionary paradigm. Formally, such mathematical deviation occurs simply when extra equations of non-hyperbolic type are added to an underlying modeling (in contrast to the adjustments inside the equations’ operators as we’ve just studied). Here’s a simple example. The most basic equations for a simple fluid are supplied by the compressible Navier-Stokes equations. But common fluids such as water strongly resist volumetric change, so it is common to add a new equation to our modeling mix that claims that the fluid always moves in a way that conserves volumes (but not, of course, shape). The resulting set comprises the “incompressible Navier-Stokes equations” (which is usually what the unadorned term “N-S equations” signifies). In fact, the extra assumption of incompressibility allows us to simplify our equational mix in quite significant ways.

Supplementary equations of this ilk are called “constraints” and generally reflect higher scale information that we know about our system in advance. That a
bead must stay on a wire or a ball should slide or roll along a table top qualify as
typical constraints. In the simplest cases (the bead on the wire or the sliding ball),
the constraints can be captured within purely algebraic equations (employ no
differential operators at all). For example, we might start with the simple set of
differential equations that report that gravity alone pulls the ball in the y direction
\[ \frac{md^2r}{dt^2} = (0, g, 0) \] but then add the supplementary constraint \( x = y \), indicating that
it slides on a table top slanted by 45° (the official name of such algebraic constraints
is “holonomic”). For our sliding bead, the constraint equations will just be those
that mark out the location of the wire.

Almost always, multiple length scales are tacit when constraints are invoked
in this way, analogous to the “fast” and “slow” time scales that appeared in our
discussion of the strut (we now need to distinguish between “large” and “small” size
scales, which I shall symbolize by “\( \Delta L^* \)” and “\( \Delta L \)”). This is because, at a smallish
scale of resolution \( \Delta L \), our bead will not be able to follow the curve of the wire
perfectly, but will wobble about its locus in a very complicated way (and the ball on
the table will display allied microscopic problems). It is only that the system looks
as if it obeys such constraints viewed at a grander size scale \( \Delta L^* \). Common physics
jargon for this (which I find a useful mode of speaking) claims that “the bead stays
on the wire perfectly when resolved on the macroscopic length scale \( \Delta L^* \).” I’ll
come back to the utility of this way of thinking later on.

Well, our bead may not be able to follow its wire perfectly, but the fact that it
appears to do so upon a length scale \( \Delta L^* \) nonetheless represents very useful
information that we’d like to exploit in classic “physics avoidance” fashion. As we
noted with respect to the hyperbolic-to-elliptic explanatory shifts above, reliably
modeling how Jack and Jill will tumble down their hill will be very difficult,
because if we model any part of it wrongly (e.g., we don’t get the shapes of the
rocks and trees on the hillside quite right), we may calculate trajectories that bear
little resemblance to the children’s actual descents. But our trust that friction will
eventually bring them to rest at the bottom of the hill (that is, within a potential well)
is very strong and well-founded, regardless of our inability to track their
evolutionary progress with any accuracy. As such, we start with some partial,
coarse-grained information with respect to how tumbling children will behave under
gravity upon a macroscopic time scale. Euler’s term-dropping equilibrium
stratagem then allows us to exploit this partial information in a manner that deftly
evades a lot of difficult evolutionary modeling.

We now want to see how we might exploit our \( \Delta L^* \) level constraint
knowledge in an allied way, cutting off a system’s microscopic behavioral
complexities from above, as it were. A classic format for doing so was codified by Lagrange in his “analytical mechanics.” This “mixed level” methodology has vast implications for practice, but let us first examine some simple cases that permit a ready appreciation of the basic mathematical difficulties that attend of the enterprise. In fact, we’ll first a basic constraint of the most primitive sort: the fact that a swarm of metal molecules appear, on a macroscopic scale, to always move together as a rigid rod.

We might roughly divide the contents of a microscopic modeling of a large system into two components: general facts about energetic conservation and momentum balance that can be trusted as fairly secure and the more specific modeling assumptions we make when we explicate in detail how our target system is laid out on a minute scale. In practice, the latter generally prove unhappily speculative and unreliable and much of the time in science we’d like to avoid depending upon such assumptions as best we can. But if one looks at the typical “recipes” one lays down for the basic underlying forms of classical physics (see Essay 3 for details), one finds that it is not easy to see how a “reliable portion” can be readily extracted that can combine wholly harmoniously with the macroscopic data that we codify in the form of “constraints.”

In fact, persuading data extracted from two different scale levels to combine with one another happily often represents a tricky task mathematically. Main thme of Essay 4 To get a preliminary appreciation of the difficulty, we might first survey the classic “combination of observations” problem from astronomy (and statistics in general). We know from baseline theoretical considerations that artificial satellites revolve around the earth in near ellipses and that, over short periods, such motions can be approximated as simple straight lines across the firmament. Such lines are registered in simple linear equations of the form \( y = mx + b \). Although we might be able to compute \( a \) and \( m \) from purely theoretical considerations, we rarely want to do this, preferring to extract values for these constants from direct observations of the satellite’s path across the sky. Here we may easily assemble an abundance of empirical observations \( E \), all of the form \( <x_i, y_i, t_i> \). So our formal task is one of combining a certain portion of “purely theoretical data” (the equational format \( y = mx + b \)) with large packet of observational data \( E \) obtained from terrestrial sightings. Now this second data set will be somewhat corrupted by minor observational errors, but surely, if we have an adequate supply of data work with, we should be able to
predict future satellite movements far more accurately employing considerations of a “mixed data” type than if we tried to augur our satellite’s movements simply through up-from-purely-theoretical-considerations-only considerations.

But here’s the rub. E contains far too much data for all of it to prove mathematically compatible with \( y = mx + b \). We want to estimate m and b from our data set E but we require only two data points to solve for these values. If we try to work naively with more values than this, we are likely to generate inconsistencies because the data in E will not lie on a perfect straight line, due to the errors they include. But it would be unfortunate if we had to discard the rest of the data in E simply because of this consistency problem. What should we do? The standard answer (pioneered by Gauss and Lacroix) holds that we should incorporate unknown error terms \( \varepsilon_i \) into each of our data points (i.e., we consider \( <x_i, y_i + \varepsilon_i, t_i> \) where \( y_i \) is the true value of the satellite’s \( x_i \) ordinate)) and then ask, of all of the possible ways in which the unknown \( \varepsilon_i \) might vary, which policy for “combining observations” might produce the smallest overall error as a rule? The answers can vary according to the assumptions we make about the origins of the errors, but in the simplest situations, we pick the m and b values that would supply the smallest squared error with respect to the data we have gathered.

Formally, we have begun with a theoretical modeling format \( y = mx + b \) that doesn’t contain enough independent parameters to accommodate our rich data set E. More specifically, the collection of equations \( y_1 = mx_1 + b, y_2 = mx_2 + b, \ldots \) is over-constrained in the sense that E places far more demands on m and b than can usually) possess a solution. However, by explicitly including the various \( \varepsilon_i \) in these formulae (i.e., \( y_i = m(x_i + \varepsilon_i) + b \)), we remove the “over-constraint” by providing some extra variables to play with. Indeed, mathematically, we’ve overshot the desired mark and have now wound up with an under-constrained equational system that won’t force unique values upon m and b. To rectify this new problem, we must invoke some further equation (generally of a variational character) into our mix to get the ratio of modeling equations to independent parameters workable (in the same mode as that of getting the baby bear’s porridge “just right”). The demand that the errors satisfy the “least squares” constraint provides us with the crucial extra formula that finally gets our mathematical demands in harmony.

[Here we might parenthetically observe that the task of persuading mixed level data to blend together harmoniously has engendered, as a concomitant side-effect, the necessity for considering a certain flavor of local possibility space, in the sense that we must miminize a certain functorial evaluator (the least squared error
attached to our data) over a certain range of possible circumstances. The role of such “local possibility spaces” will assume great philosophical salience within our chapter ten.]

Our problem of combining higher scale level $\Delta L*$ data derived from constraints with theoretical considerations operating on a microscopic scale displays a mathematical character quite similar to the “combination of observations” problem. Let’s now consider the simple constraint embodied in a rigid rod (which is captured by the algebraic formulas that require that the distances between component molecules remain always the same). But if we attempt to provide a detailed model of the underlying molecular attractions employing conventional lattice binding forces (at size scale $\Delta L$), we will construct models that manage to keep their component particles near the central axis of the rod only approximately. Looking at such components under a powerful microscope, we would find that, at any given moment, the true positions of its molecular masses will lie strewn about a central axis in a pattern very much like the scattered data that we record in the satellite case. But “resolved on a macroscopic scale length” $\Delta L*$, the rod “looks perfectly rigid”--it appears, for all that we can tell at this level, as if the central molecules in our swarm all lie upon a constant straight line. Accordingly, our “constraint” that “the bar is rigidly straight” mathematically serves as a collection of data $E$ drawn from macroscopic observation that situates all of its data points upon a common straight line, despite the fact that no ordinary molecular model can accommodate that demand on a lower scale. So we witness the same mathematical tensions as before, except that the data set we regard as accurate has switched positions with the data set that contain the “errors.”. But we are still faced with the mathematical task of allowing the macroscopic data set $E$ to blend harmoniously with the general theoretical formulae offered by lower scale mechanics. But how can we do this consistently?

The resolution proposed by Lagrange proceeds as follows. The underlying classical physics of molecules works with “forces” (or stresses) in some manner or other--let’s assume (from the possibilities surveyed in Essay 3) that Newtonian physics in its point mass formulation is active at scale level $\Delta L$. The “natural forces” we might lay down in such a modeling will only place our molecules along
the scattered E path at a given time. So let
us now introduce some further “forces of
constraint” into that unassisted molecular
picture in a manner that always supplies
exactly the right extra “oomph” to pull a
wavering molecule back to the central axis.
By tolerating these new “force”
supplements, we will have reached
accommodation with our underlying
mechanical principles, for the new
constraint forces allow our E-level demands upon molecule location to become
perfectly consistent with mechanics’ usual expectations with respect to force-driven
behaviors. Plainly, these new forces remove the “over-constraint” conflicts in our
modeling efforts in exactly the same manner as the \( \varepsilon_i \) factors above. Just as before,
we have overshot our mark and require some additional principle to complete our
equation set (right now we lack enough equations to compute the exact magnitudes
of our new “constraint forces”). The clever answer championed by Lagrange
employs variational reasoning (“least work” + d’Alembert’s principle) to provide
the missing data.

To properly appreciate this scheme, we should consider a slightly more
complicated example. Let’s return to our bead sliding along a wire, subject to some
battery of outside forces (gravity, electromagnetism, etc.) to see how this works.
Lagrange’s ultimate conclusion is that we only need consider the “active part” of the
forces acting upon our bead along the constraint to calculate how it will move. And
we can ascertain the strength of these “active forces” in two ways. First, consider
whatever modeling equations we would naturally supply for the forces acting upon
the bead (the forces that would otherwise pull the bead along the path of
unconstrained fall illustrated). We supplement these with an array of so-called
“Lagrangian multipliers” that correspond directly to our corrective constraint forces.
We then determine their strength by insisting that they always point directly into the
wire wherever the bead happens to be (this is the equivalent of the least squares rule
for the error terms in our astronomical example). But we need to express this
condition in terms of varied approximations to the true path, because we can’t
calculate where the “normal direction into the wire” is until we have a guess as to
where the bead might be at the time in question. And Lagrange chooses amongst all
of these approximations by combination of Bernoulli’s “principle of least work” (to
handle static beads) and d’Alembert’s principle (to handle moving beads). Rules
that work like this are called “variational principles,” even when they don’t assume
the integral form of the “principle of least action” (in fact, the non-integral forms
tend to be more general in their application). When the required optimal
approximation is found, the computed values of its “Lagrange multipliers” will tell
us exactly how strong our “constraint forces” must be.

The upshot of all of this is that we really only need to know how the “active parts” of the applied forces operate to successfully augur our bead’s motions—the
constraint forces simply cancel out those aspects of the applicable forces that
“perform no work” in this setting (i.e., they don’t help move the bead up and down
the wire). In old-fashioned jargon, those parts of the microscopic force situation
become “lost” in the sense that their operations get completely covered over by the
forces of constraint (in a moment we’ll see that these “performing work”
considerations are intimately tied to the system’s capacity to transfer energy across
its contours). Through such “work upon the wire” considerations we manage to
\textit{efface} most of the complexities of a thoroughgoing $\Delta L$ modeling from our $\Delta L^*$
vantage point. Putting the matter a different way, we’ve uncovered a new mode of
“physics avoidance” that allows us to forget (mostly) about how the applicable
forces move our bead with respect to its
Absolutist spacetime background (as the
microscopic force principles will do) and worry
only about what occurs with respect to the
surface of the wire.\textsuperscript{17} And we spare ourselves a
vast number of horrendous modeling headaches
if we engage in this almost indetectable
injection of “physics avoidance.”

So the true mathematical purpose of Lagrange’s variational techniques is to
harmonize the general demands that our underlying mechanical principle place upon
physical systems at a microscopic $\Delta L$ level with the partial knowledge on a larger
scale captured with our $\Delta L^*$ level constraints. And so we have a new explanatory
matrix working in the manner of one of Railton’s
“condensed texts”: invocation of constraints allows
us to omit many microscopic modeling details that
are both irksome and liable to lead us astray in our
predictions in any case. So we can enjoy both our
“physics avoidance” and maintain mathematical
consistency at the very same time.\textsuperscript{18}

In common modeling practice, engineers
often identify the active “forces” operating within a complicated system directly—by either direct inspection or through experimental manipulation (this is the second mode of “ascertaining the strength of the active forces” I mentioned above). Thus in the complicated Rube Goldberg device illustrated (overlooking the parrot and octopus), we witness a complicated mechanical way of transferring a capacity to perform work from one side of the system to another with comparatively little frictional loss (in other words, the energy fed into the system remains largely intact although converted to different forms of application than it originally assumed. And Bernoulli’s “principle of least work” captures the statical relationships at the center of the behaviors we expect to find in setups like this (at least to first approximation, ignoring friction). As such, the doctrine represents a natural generalization of the venerable balancing principles inherent in the lever. But it is easy to identify all of the “active forces” involved here by inspection; we do not need to calculate their strengths from fundamental principles. In other cases, we might experimentally ascertain that the lump pictured can be maintained in equilibrium if we surround it with the array of “forces” pictured. Through such trial-and-error procedures, we can identify how the basic pathways of transfer of work capacity operate within the system without needing to assemble any microscopic mechanism that explains how the blob accomplishes this feat. The “manipulation counterfactuals” central to Woodward’s work on causation often encode data of this $\Delta L^*$ level sort. As such, they provide the vital “mixed scale” information that allows us to bypass many irksome details of microscopic causal process (in the sense we highlight when we consider modelings of a fully evolutionary character).

We should observe in passing that the “forces” and “masses” we consider when working in this higher scale mode won’t always be the same as the “forces” and “masses” we consider in a more fundamental physical modeling, for they often need to be of a “generalized” ilk (i.e., torques rather than linear forces, moments of inertia rather than masses, etc.).

In this section, I’ve sketched the central adjustments in focus that the word “cause” displays as it wanders through the jungle of diverse forms of effective computational strategy. But why, from a linguistic point of view, should we want a word that behaves in such an inconstant fashion? Essays 6 and 9 will suggest some answers.
Aspirational hopes again. However, expectation that the active “forces” considered within practical engineering can be successfully underwritten by detailed molecular models of a classical stripe should generally be viewed as an *aspirational hope*, in the sense of section iv. As I detail at some length in the book’s appendix, such questions are often impossible to resolve simply because the full docket of “special force” laws operating at the $\Delta L$ level rarely get fully specified within a classical setting (why? because Nature works according to quantum principle down there and skillful physics avoidance can allow us to ignore such “filling in” concerns for long periods of time). In fact, efforts to supply lower scale supports for the most common constraints of everyday engineering tend to wobble about considerably in the choice of “fundamental objects” selected: some treatments invoke classical point mass “molecules”, some, rigid bodies and others, flexible blobs (such foundational wobbling supplies “classical physics” with what I call a “theory facade” character). Success in *wallowing off such lower scale modeling concerns* through the skillful exploitation of higher scale data has long allowed working physicists to resolutely evade the alleged necessity of laying down firm “foundations of classical physics” in the canonical mode expected by the “theory T” crowd. Skillful mixed scale physics avoidance is often compatible with a wide variety of supportive underpinnings and many of its most successful $\Delta L^*$ level treatments probably require quantum mechanical mechanisms for their microscopic realizability. After all, most of the “general physics” considerations with respect to “work capacity” that we will wish to blend with our $\Delta L^*$ constraint knowledge are are equally supported within quantum physics as well, so it isn’t surprising that Mother Nature doesn’t appear to pick any specific formalism as her favorite flavor of “fundamental classical physics.”

From a methodological point of view, physics textbook practice is often misleading on such issues. Typically they “prove” Lagrange’s variational principles by simply multiplying basic level Newtonian formulas by arbitrary variations. These manipulations merely show that if a system $S$ can be successfully modeled at level $\Delta L$ in a wholly Newtonian mode, then Lagrange’s principles will be (approximately) true of the system when it is resolved upon a scale length $\Delta L^*$. But there are many systems where Lagrange’s principles genuinely hold true “resolved at scale length $\Delta L^*$” but plausible classical underpinnings at a lower size scale cannot be provided. For such reasons, Lagrange’s principles are often *stronger* in their physical import than the classical principles than our books employ to “prove” them (a situation that
only the best books on mechanics make clear\textsuperscript{19}). From a logical point of view, you can’t legitimately \textit{strengthen} a formalism via deductive manipulations that only serve to \textit{weaken} its import (as “multiplying by variations” properly does).

All of this supplies our $\Delta L^*$ level engineering statements with an \textit{autonomy in truth value} that makes it preferable to maintain that they should be viewed as “(simply) true resolved at the scale length $\Delta L^*$” rather than holding that they are “approximately true at scale length $\Delta L^*$, when the truth-values of the $\Delta L$ claims to which they are supposed to “approximate” are far less secure in their truth-values (and even existence) than their $\Delta L^*$ scale counterparts. Writers like Bob Batterman have sometimes employed metaphysically loaded terms such as “emergence” to capture this improved reliability and autonomy within the $\Delta L^*$ scale modelings, but I’d prefer to speak directly about the “reliability of truth-values” with respect to a linguistic format.

There are many classical techniques that strive to eschew speculative lower scaling modeling through the skillful exploitation of higher level behavior. Foremost of these is the “top down” approach to continuum mechanics pioneered by Cauchy, Green and Stokes, which exploits the fact that many materials tend to respond to pushes and pulls in the same way over a tremendous number of size scales (such behavior is, appropriately enough called “scaling”). This provides another way in which general lower level mechanical assumptions about stress and strain can be fruitfully combined with higher scale experimental information to produce a trustworthy mechanics that successfully sidesteps the tribulations of honest, totally bottom up modeling (“honest” in the same way that there are certain “totally honest” people in the world that one should avoid like the plague).

In truth, Cauchy-style continuum mechanics goes a bit too far in its scaling assumptions, presuming that they carry validly all the way down to size 0. But at various smallish size scales at the nanophysics level, we witness a range of different plateaus of behavior involving material grain, dislocations and so forth, where the dominant physics operating at that level tends to be rather different than at other scale sizes. In such cases, rather than bringing the data of just \textit{two} size scales ($\Delta L$ and $\Delta L^*$) into cooperative coordination (as occurs in both classical continuum mechanics and Lagrangian analytical mechanics), we would like to do the same for the data pertinent to a much larger range of concerns $\Delta L$, $\Delta L^*$, $\Delta L^{**}$, $\Delta L^{***}$, etc. AS one might guess from our study of Lagrangian technique, effecting
such coordination is not mathematically easy (it is sometimes called “the tyranny of scales problem” presumably in analogy to a continent of little duchies who would do better if they pooled their respective resources but are prevented from doing so by uncooperative local autocrats jealous of their special authority over the principalities in which they rule). Much of Bob Batterman’s recent work is inspired by some recent successes in overcoming this traditional problem.

Here I should emphasize--since the point is frequently misunderstood by commentators--that the overall objective of these techniques is not to “derive” the higher scale behaviors from wholly bottom-up considerations, but to successfully blend the secure bits of knowledge we possess with respect to a wider range of scales than just two. This observation is simply an elaboration on the fact that the strength of analytical mechanics within applications is much greater than its apparent “derived strength.”

Once again, if the departure from a purely evolutionary setting that is induced by the invocation of constraints is not noticed, substantial conceptual confusion can result, for we are working within a substantially altered explanatory matrix whenever constraints are involved. Here’s a simple, but instructive, example. As noted above, fluids are often assumed to be incompressible (= maintain the same volume throughout the flow) in practice. That demand represents a higher scale constraint upon the fluid’s potential motion, albeit not a demand as restrictive as that of the full rigidity we attribute to a steel rod. Suppose that such a liquid suddenly encounters a sharp constriction as it moves through a pipe. In order to maintain its density, each little packet of fluid must adjust its shape so that it can move swiftly through the narrow end of the pipe without an altered volume. But how can the back end of a blob of fluid first entering the constriction manage to “see” what its front end is doing after it crosses the bounding line into the constriction? Proper answer: a “fast time” process of pressure equalization must occur within the blob over a short interval of adjustment time, the details of which get completely suppressed (= “avoided”) under the constraint of perfect incompressibility. That is, a real fluid blob must have time to shuttle minute compressive waves rapidly across its interior in order to restore its varying densities to locally uniform values (these “fast time” requirements are similar to the communication processes that allow a position y on our strut to “see” the falling weight above). If we suppress these swift mechanisms through the imposition of “incompressibility,” we ipso facto force the physical
significance of “pressure” to shift in a subtle, but very real, way (indeed, “pressure” loses its “absolutist” significance as a true stress, as every advanced book on fluid mechanics ably explains). Processes within “incompressible fluids” often become “hard to understand” for precisely this reason, although no one would wish to forgo the remarkable computational simplifications that “incompressibility” brings in its train.

So we once again witness a subtle form of “initial value imposter” whose “physics avoidance” departures from normal evolutionary expectations can be hard to unravel (the setting is crudely “evolutionary” overall, but contain a subtle seasoning of component processes that are not approached in that fashion). I’ll have to leave to some of the other essays in this book the task of demonstrating just how trouble that allied puzzles of “physics avoidance” harmonization have occasioned across the entire field of philosophy, even within arenas that seem nominally far removed from physical concern altogether.

The delicacy of these departures underscores the real harm that theory T thinking has brought to our diagnostic sensibilities. A writer like Peter Railton is apt to view the insertion of a large block of “constraint” information into an explanatory matrix as a rather simple affair, tantamount to simply dropping the unnecessary chapters from an “ideal text” novel delineating our heroine’s boring developmental history between June and December (“Reader, I married him--everything else is too tedious to relate”). But this simple conception of “text condensation” doesn’t suit the real life nature of constraints, for here the “important events” are continuously conditioned by their ongoing interactions with the “boring events” that maintain the constraints. Our “narrative abridgement” situation more closely resembles a Hollywood filming of Othello where Iago has been dropped from the plot line for fear of an excess of characters. But without his conniving presence, it becomes hard to understand why the remaining characters sometimes act as they do, for we’ve eliminated all communications between the Prince and his evil counselor. And the same phenomenon is at work in our incompressible fluid example: we’ve suppressed the compression waves that allow a distorting blob of fluid to even out its densities.

In later essays in this book, we shall often attempt to unravel some stock controversies within the lore of science employing our Hadamard-inspired diagnostic tools. Thus Essay 8 will be largely concerned with unmasking the “initial value mimics” that occasion so much trouble within classical optics.
Since we have already devoted a fair amount of attention to the manner in which the word “cause” drifts with the background adjustments in explanatory strategy, it is worth briskly remarking on a situation where this drift is carried to extremes that have historically occasioned large gobs of significant philosophical misunderstanding.

I have in mind the explanatory oddities that attach to the notion of mechanism. We have already noted that the algebraic equations that capture (holonomic) constraints have a remarkable to efface the ΔL* description of a system from its ΔL level underpinnings. Now in the typical situations in which analytical mechanics is applied, this effacement is only partial and Langrange’s multipliers provide a methodology that allows each level to contribute its wisdom in a fairly equitable manner. A situation like the Rube Goldberg device pictured involves a lot of rigid parts but they don’t completely overwhelm the character of the physical system in which they appear. But suppose that we assemble a collection of rigid parts into a closed kinematic chain like that pictured, where the pattern in which the parts are pinned together permits only a single internal degree of freedom to the ensemble (turning the crank 2 on the left completely determines the position of every other part in the device). Mathematically, we confront a set of purely algebraic equations (expressing the constraints that tie the rigid pieces together) that completely override any differential equations that might be operating on a lower scale level (in such cases, we’ve really avoided the underlying physics). This overpowering of the lower level physics creates a situation where our “closed kinematic chains” operate according to ΔL* level principles that are quite unique in character. But “closed kinematic chains” are not creatures unfamiliar to us; they, in fact, provide the modern explication of what, mathematically, makes a mechanism in the ordinary sense qualify as such (so standard courses on mechanism are largely devoted to the special properties of closed kinematic chains). And the founder of the modern theory of machines (Franz Reuleaux) recognized this extreme effacement from ΔL level principle in very pungent terms:

[T]he sense of the reality of this separation [of the theory of mechanism from
general Newton-style mechanics] has been felt not only by engineers or others actually engaged in machine design, but also by those theoretical writers who have had any practical knowledge of machinery, in spite of the increasing tendency in the treatment of mechanical science to thin away machine-problems into those of pure mechanics.\(^2\)

Very substantial philosophical controversies trace, at root, to an inclination to mistake the usual explanatory matrix we apply to machinery for a true initial value framework (those who fancy that proper Newtonian physics tells us of a “clockwork universe” have been dramatically misled by a substantive initial value mimic, for the notions we typically apply to machinery are nothing of the sort). The great historical attraction of “relationalism” with respect to space traces in large part to this source, I think, for the internal mathematical closure provided within a mechanism’s algebraic constraints make it nearly impossible for outside influences (such as positioning with respect to a spacetime backdrop) to play any role in their dynamic behavior (mechanists call these identical internal movements “inversions of the mechanism”). But I’ll explicate these issues more fully in the essay on Descartes.

But for present purposes, I’d merely like to comment on how the word “cause”‘s behaves within this much effaced setting. Let’s look at a more complicated sewing machine device that converts a circular motion on the right hand crank into a back-and-forth motion suitable for stitching at the top. Looking at the picture, I find it hard to anticipate with certainty how it will move as one turns the crank leftward (I’m sure that experts with a better “feeling” for mechanisms will have less trouble with this task). In particular, I’d like to know whether the triangular piece on the left will rock clockwise or not. And I might naturally express my worries thus: “I see two causal pathways leading from the crank to the triangle and I can’t determine which one will dominate?”

But what exactly can this question signify? I can’t be asking about the real causal processes involving compression waves that actually carry the energy inputted at the crank through the molecular matrix to the triangle, for those undoubtedly travel in patterns bearing no discernible relationship to the question I’m asking. Nor does the case appear much like the manipulation dependencies that Woodward highlights in his work. Insofar as I can see, the question I’ve asked is entirely procedural in character. If we are given a set of entangled equations to solve, we would sometimes like to know whether one of these variables (z, say) will
turn out positive or negative, for that information might help us hone in on a correct answer to the whole set through educated guesswork more swiftly (if we guess this sign wrongly, we might waste a lot of time pursuing unworkable answers). And I think my question about “which way is the triangle caused to move?” is of this same character: how can I best begin the process of “solving” how this device will behave?

And I think it is quite natural for words to wander like this, for there is usually some question of “what should I think about first?” involved in all of the situations that we have surveyed in this essay. The case of mechanism is unusual in that the word “cause” has almost totally severed itself from the question of wave transport with which we started. The correct methodology, I think, is not to seek some diaphanous “meaning” for the word that embraces of these situations, that in situations but to trace the background shifts in explanatory setting wrought by various effective policies of “physics avoidance.”

I do find it somewhat ironic that certain philosophers have selected mechanisms as their chief paradigms for “causal behavior,” when it strikes me that the word’s behaviors within that refined setting proves particularly anomalous, at least in terms of the underlying structural factors that the word highlights within this context.

**Summation:** Eschewing purely logical classifications in favor of the sharper diagnostic tools available in modern mathematics can help us better appreciate the wide variety of explanatory settings according to various strategies of physics avoidance that underwrite their utilities. Through better appreciating the many patterns in which the various “condensed books” of physics practice relate to one another, we can better understand the complex drifts to which promiscuous word like “cause” are liable. In these respects, real life scientific explanation adopts the frequent shortcuts and lifts of a game like “Chutes and Ladders” more often that it dutifully follows the direct pathways of a “Candy Land.”
Appendix: A troublesome “initial-boundary condition problem” mimic.

Let’s consider a seemingly simple optical situation where light enters some transparent medium such as glass from the air. It is common for textbooks to loosely announce, “let’s consider what happens when oscillatory data of a fixed frequency $\omega$ enters a crystal” (i.e., a beam of a pure spectral color). This looks like a standard initial-boundary value problem but it’s really not. Why? We have instead been invited to consider a steady state situation: what happens to continuously streaming light as it strikes our glass? We are not asking, “what transient light patterns will form after the leading edge of the incoming wave packet enters the crystal?” (that is a much harder question involving transient response of a considerably more complex character\textsuperscript{22}). As noted above, steady state problems typically represent a projection of an evolutionary development onto a “base manifold” in the manner of our fluid passing over a block. In terms of the relevant equations, we will have factored off the true time dependency through appeal to “separation of variables,” a manipulation which converts our governing equations from hyperbolic to elliptic. This loss of hyperbolic character is not surprising because we evoke “steady state conditions” precisely as an effective “physics avoidance” dodge to get the complexities of a detailed temporal development (including its opening transient patterns) out of our hair. But, as we saw before, “time” has a manner of sneaking back into an explanatory setting from which it has been officially banned. If the rate of flow is slow enough, it is easy to paint “pretend time” arrows on the pattern, indicating that the water is flowing in from the lefthand side of the diagram.\textsuperscript{23} But such a situation, with its “pretend time” reintroduction, should not be mistaken for a true initial-boundary value problem.

A similar situation arises when we ask how monochromatic light scatters from a perfectly reflecting, infinitely long razor blade. To talk of “monochromatic light” in such a context implicitly requires that we are
considering an evenly spaced wave pattern coming in from spatial infinity—i.e., another steady state circumstance, in which we again factor away the time component. In terms of equations, we move from a Maxwellian wave equation containing a time variable to the elliptic “Helmholtz equation” appropriate to steady-state conditions that lacks any mention of time. The usual recipe for allocating a “pretend time” to this situation becomes a little more complicated because the incoming light is now thoroughly mixed with the outgoing light everywhere in the diagram (so we can’t identify the former as the “flow witnessed to the left of the block”). However, Sommerfeld found an exact solution for the razor problem where the “incoming light” correlates with a major term in his solution whereas the “outgoing light” is represented by a mixture of other terms. This decompositional fact allows us to introduce suitable “pretend time” arrows into our (strictly speaking) “timeless” steady-state solution. But, again, we are not addressing the problem in a true “initial-boundary value” mold.

Things quickly become confusing because Sommerfeld’s starting point allows us to build up localized “wave packets” from his basic patterns through interference between the monochromatic solutions for a spectrum of wavelengths. These new structures will naturally move through our diagram as coherent entities, restoring a robust sense of “temporal development” to the proceedings (in essence, we can generate a “true time” development for the packet from the differences in the “pretend times” arrows that we have assigned to its decompositional steady state patterns). As such, we obtain a description of a temporally adjusting field that can also accept a coherent formation as a true initial-boundary value problem (in contrast to the “oscillatory data” from which the “packet” was constructed). In other words, although we now have an intelligible initial-boundary value problem before us, we have addressed it through an odd detour through problems of a different equational signature. Failing to recognize the strategic logic behind this roundabout procedure has given rise to many conceptual puzzles within optics (some of which will emerge, with considerable teeth, in essay xxx).

An additional remark is pertinent here. Since we are dealing with an intervening elliptic equation, we anticipate that its solutions will possess the flatworm-like reproducibility properties indicated earlier. This remains true in
Helmholtz-based scattering, but we shouldn’t assume that such continued solutions will be content to remain single-valued everywhere around the razor blade (Hadamard emphasizes this fact in his Lectures). In fact, the true analytic continuation of Sommerfeld’s solution climbs onto a so-called “Riemann surface” above the scattering plane and it is only by projecting single-valued pieces of this oddity back into the plane that Sommerfeld constructs his “real light wave” scattering proposal, where the “boundary condition” associated with the razor blade disguises the slits where the analytic function’s unfolding upon its Riemann surface have been cut up (the diagram provides a somewhat symbolic “picture” of this relationship). Considered with respect to its “regular space” projection alone, we find ourselves in a situation where “connection formula” still interrelate the three main sectors of our light scattering diagram, but they no longer represent continuations of a true “analytic function” type.

In other words, the mathematical notion of “analytic communication between sectors of a domain” does not always serve as a perfect match for the physical notion of “equilibrium or steady-state communication between sectors of a domain.” Wary readers should be advised that other departures from fully analytic continuation frequently appear in physics, although the logic of why “analyticity” fails can prove quite different than the considerations we have highlighted under our discussion of hyperbolic behavior.

Let’s reconsider our monochromatic light entering glass problem in more detail. Superficially, it looks much like the razor blade scattering case, except that we now allow waves of some sort to form inside the glass. But what formulas should govern how the electromagnetic field behaves inside the medium (in the previous example we only needed to know how light moves within a vacuum, for we modeled the razor blade itself as a “boundary condition” demanding that all incoming energy is immediately reflected back into the surrounding void). But our glass case requires equations to govern light’s movements inside the new medium as well as plausible interfacial conditions that monitor how the crossover between media occurs (we’ll assume that no reflection occurs).

An important historical oddity complicates the picture. Originally, Maxwell proposed a family of equations for a range of different media, differing only in their “coefficients of permittivity,” and did not credit the form he assigned to the vacuum (i.e., his “aether”) with any special virtues beyond the fact that light travels fastest there. Nowadays, his formulas for materials other than the vacuum are called his “bulk equations,” wherein $\text{ME}_\varepsilon$ might represent the special form appropriate to ordinary glass. But there are many materials (calcite, sulfur) that behave in more
complicated ways than Maxwell’s models allow and it was eventually decided that “Maxwell’s vacuum equations” (ME\textsubscript{V}) enjoy a special privilege as the ones that capture how the electromagnetic field behaves everywhere, the behavioral differences within optical media being entirely due to the ways in which surrounding electromagnetic fields are affected by the charges of the atoms they encounter. From this vantage point, Lorentz proposed rough models to explain how lattice arrangements of this type might give rise to Maxwell’s old set of “bulk equations.” Accordingly, when a modern textbook speaks of “Maxwell’s theory,” they properly intend the ME\textsubscript{V} equations alone (although the traditional “bulk equations” often sneak back on stage without appropriate fanfare).

But there are two problems with this standard approach. (1) In the revised Maxwellian framework, we are not, in fact, properly supplied with a full equational story for how charged matter should interact with the vacuum field (crucial pieces such as the rules needed to govern the “back action” of the field upon the particles are left missing). (2) Lorenz’ modeling assumptions often prove incorrect for many common forms of transparent material. This unfortunate situation invites us to look for some “physics avoidance” ploy to help us work around these foundational lacunae. A clever stratagem now enters the scene, based upon little more than the innocuous fact that most optical materials can support steady state dispersive solutions— that is, one can shine a flashlight on them and, after a short while, they will settle down to continuously allowing the light to pass through in possibly complex, yet steady, manner. That such unruffled transmission is possible at all must be regarded as a purely empirical matter; for other materials can trap the incoming light in such a way that they eventually disintegrate or blacken after some lapse of time (viz., in this refusal to tolerate a steady state, they resemble the struts that never settle into equilibrium). But how can we securely exploit this rather innocuous form of “steady state” behavior to glean important data about the behavior of light in exotic media in the absence of any plausible Lorenz-like story for how their molecular constitution interacts with its surrounding light field? In particular, there are two questions we would like to know. (a) To what degree will the medium cause monochromatic waves of different wavelengths to travel at different speeds (or “disperse”) relative to one another? (b) To what degree will the medium cause monochromatic waves of different wavelengths to lose energy (or “attenuate”) relative to one another? At first blush, it would seem as if these two sorts of behavior should prove unrelated.
A little further thought, however, suggests that, no, the two responses must cooperate in some complex manner for steady state propagation to prove possible over a long term, for otherwise internal changes in the medium are likely to appear. And it turns out that, under remarkably weak assumptions, that steady state behavior is truly possible only if the extinction rate $\chi_2$ of a signal at frequency $\omega$ is related to its susceptibility $\chi_1$ through the Kramers-Kronig relationship:

$$\chi_2(\omega) = -\frac{2\omega}{\pi} \text{P} \int_0^\infty \frac{\chi_1(\omega')}{i(\omega'^2 - \omega^2)} \, d\omega'$$

where we integrate over the various susceptibilities of all possible frequencies (“$P$” indicates “Cauchy’s principal part”). Unfortunately, typical presentations of this reasoning in textbooks make its workings appear rather mysterious, largely because they describe their techniques in fake “initial-boundary value” terms. But such presentations are misleading, for, strategically, we are following Euler’s advice and skirting the unknown details of how our material might asymptotically evolve into steady-state conditions by simply looking ahead to the postulated final condition directly (as we’ll see, the dispersion relations will supply us with a lot of information about what the missing (in an “aspirational” manner) evolutionary equations for the medium will need to look like). Here’s an outline of how the core “physics avoidance” reasoning operates. For any given frequency $\omega$, we first consider what a pure steady state response should look like when a wave pattern arrives at the glass with a frequency $\omega$. How will the differences between the vacuum and glass affect its pattern once the interface is transversed? It is useful to compare this situation to one that electrical engineers ask about complicated electrical circuits: what will the steady-state output of the circuit look like if it is fed an input current oscillating at a pure frequency $\omega$? And the answer comes in the form of a transfer function that redistributes the incoming energy into other frequencies $\omega'$ with some attenuation. And we likewise anticipate that the differences in response between ordinary optical materials will operate in this basic “transfer function” fashion, except that, in contrast to the interior workings of our circuit, we presently know little about the precise manner in which the atoms within our glass will accomplish this feat. To avoid unnecessary speculation about small scale modeling issues, we exploit the fact that steady state solutions (through their elliptic character) must be analytic in their functional character, indicating that the extinction rate $\chi_2(\omega)$ must depend in some
analytic manner upon what happens to the energy that transfers to other vibratory modes. Standard arguments employing contour integration then establish that, under our steady state response assumptions, the Kramers-Konig relationship must hold, for the only salient pole lies upon the real axis at $\omega$ itself. This is a classic “calculus of analytic functions” computation that reappears in many forms of physical guise. Physically, the “analyticity” merely captures the degree of across-all-frequencies cooperation that a material must display in order to support a steady-state response.

The next natural step is to presume that the two traits we have now linked together in steady state conditions (viz., susceptibility $\chi_1$ and extinction $\chi_2$) will maintain those same relationships under all circumstances of illumination, steady state or not. But how could it be otherwise?—the glass isn’t smart enough to plot, “Oh, I’m going to disperse incoming light by different rules according to how long the experimenter plans to leave her light on.” In consequence, we find that we can exploit the empirical assumption that steady-state responses are possible within our glass to derive some very valuable restrictions as to how waves move across different mediums (dispersion properties are much easier to measure than the internal refractive indices central to Maxwell’s bulk equations). And we can draw these conclusions before we have gathered enough information about our system to frame initial-boundary value problems for our vacuum + medium system in a proper manner.

In short, standard approaches to the dispersion relationships supply a valuable lesson in how missing or dubious foundational assumptions (“how do the molecules of glass respond to an ambient field?”) can be skirted with some clever strategizing and a modicum of higher scale information (“glass can support a steady state response to an incoming light stream”). As noted above, physicists (as opposed to mathematicians) often present the salient manipulations in the guise of an ersatz “initial-boundary value problem” relying upon the “pretend times” we can associate with the spatial wavelengths corresponding to the frequencies $\omega$. Sometimes philosophy of science has fallen into considerable confusion through failing to unmask these presentations as “initial-boundary value problem” mimics. Here’s an example. In attempting to rebut Russell’s faulty claim that modern science eschews any robust notion of “causation,” many writers of Nancy Cartwright’s school have argued that the notion needs to be brought into scientific model construction as a supplement (and even overlord) to the canonical “laws of physics” upon which Russell constructs his case. Under the present point of view, this response is
unnecessary; Russell simply didn’t look in the right places in the right sorts of “fundamental equation.” In Cartwright’s vein, Mathias Frisch cites the invocation of “causation” in standard derivations of the dispersion relations and he appears to regard this intervention as a new constraint (or “weeding out principle”) applied to the “solutions” that the normal equations of electrodynamics would otherwise accept. More specifically, he views this supplementary “constraint” as operating in the following manner. Consider the two forms of “fundamental solution” that one might consider with respect to concentric wave motion: (i) “retarded motion,” where the waves travel outwards from a source and (ii), “advanced motion,” where they travel inward towards a focus. In nature--and ripples on a pond are commonly cited--, we almost invariably witness events of type (i), whereas those of type (ii) appear unphysical (and are only “seen” in films that are run backwards). It is sometimes urged that we should therefore invoke “causality” as a supplementary condition to standard electrodynamics as a means of discarding “unphysical solutions” such as (ii).

In truth, patterns of type (ii) are not quite as rare in nature as all of that, for lenses or curved barriers can produce ingoing waves of the required sort, although they are, admittedly, less frequent. So an outright ban on such solutions is not a viable point of view.

In any event, Frisch and most of his anti-causation critics have mistakenly assimilated the explanatory logic of the dispersion problem to that pertinent to a true initial-boundary value problem, over which they argue whether such “causality supplements” are needed or not. In proceeding thus, they overlook the rather significant datum that we commonly confront situations where half of the equations required (the ME\(_G\) equivalents for the medium in question) for a well-set initial-boundary problem are either missing or regarded as untrustworthy. Properly speaking, the textbook invocations of “causality” that Frisch cites merely codify the empirical observation that steady-state conditions are possible for such media (in the guise that we can “see” the pure frequency \(\omega\) entering the glass as a “cause” in the same sense as we “see” the incoming “causal” flow in our base manifold projection of fluid behavior”). But the purity of the pretend time “incoming flow” we witness in both setups is special to the steady-state circumstances under consideration; it does not represent the general circumstances of either light or water. Nonetheless, the restrictive considerations we can extract from the feasibility of steady state behavior (which clearly represents data of a higher scale character) allow us to place important restrictions upon light behavior in all circumstances, without reliance upon a full-blown set of underlying evolutionary equations based upon
untrustworthy lower scale modelings. So the invocation of “causality” is doing real work for us in such “missing equation” situations, but it does not assume the “extra constraint upon solutions” character that Frisch imagines.

**Endnotes:**


2. David Armstrong *What is a Law of Nature?* pp. 6-7. I can scarcely imagine a pithier encapsulation of the sangfroid that often makes philosophical speculation about science embarrassing. We have a plain duty to capture the methodological gambits of real life practice in *descriptively accurate terms* before we launch into self-aggrandizing fantasies of what that “methodology” should look like. Essay 6 visits these depressing themes in some detail.

3. Codifying complex mathematical relationships in what we now regard as “approximation space terms” represents a very common developmental ploy within the history of mechanics and we shall study an intriguing examplar in Essay 7. We can often appreciate the descriptive strategy behind an explanatory format better if we know how the routine plays out in numerical terms and I shall employ this expository device often in this book of essays.


6. Classifications become more delicate when algebraic constraints enter the picture but we’ll not worry about such issues here.

7. In truth, a fair amount of “avoidance” has already been applied in our original equational modeling for the strut, especially with respect to its drop in dimensions and the manner in which its upper end loading is treated. In proper strictness, we
are not considering a true “initial-boundary value” problem here, because we tacitly assume that any initial velocity in the band will be the result of the velocity with which the weight is loaded on the top of the strut at time $t_0$. Properly characterized, such assumptions indicate that we contemplate circumstances where an impulsive side condition has been assigned to the upper boundary: the endpoint remains fixed at every instant except $t_0$, at which time a delta function “displacement” is allowed to enter the top of the strut through the instantaneous loading. “Side conditions” of this ilk are very common in electricity and are often loosely characterized as “initial conditions.” See Essay 8 for more on this, where such delicacies become truly critical.

An expository problem I face is that it is rather hard that provide simple illustrations of hyperbolic PDEs that don’t tacitly incorporate some degree of “physics avoidance” in their background derivations--indeed, this automatically occurs anytime we reduce the dimensions of our problem lower than three (our strut is one-dimensional). Some trickier issues to be discussed in later essays (particularly in Essay 7) require that we attend carefully to such technicalities, so I need to now acknowledge a certain looseness in my current characterization of the strut. Nothing material in the present discussion will turn upon these subtleties.

8. In a non-linear medium, shock waves can sometimes exceed these restrictions, but we’ll not worry about these exceptions here.

9. Special relativity requires that the “sound cones” of any basic physical process must fall inside the light cones carved by the characteristics of Maxwell’s vacuum equations.

10. Technically speaking, the B state still qualifies as a possible equilibrium even when the strut is loaded beyond $W_c$, but it is unlikely to be witnessed in practice since it is unstable (= easily prompted into movement by any slight disturbance). Unstable equilibra often appear in standard bifurcation diagrams as dotted lines, but we shall ignore them altogether.

11. Generally speaking the reduced operator becomes elliptic but, in the one-dimensional setting considered here, we obtain a simple ODE.

13. Note that in our earlier considerations, we treated W as fixed and not variable at all.

14. Usually the equilibrium only appears asymptotically (as $t \to \infty$). In this modeling, we tacitly assume that the frictional drag is entirely occasioned by the ambient air. In reality, a good deal of frictional leakage will occur at the endpoints, but such boundary conditions are hard to handle.


16. The common policy of encouraging philosophy students to regard all of mathematics as nothing more than “applied logic” abets this unfortunate transmogrification, to such an extent that the analytical metaphysicists of Essay 5 have convinced themselves that it is utterly inconceivable that the universe might behave in any other manner with respect to its “fundamental ontology and its metaphysical groundings.”

17. Residual Coriolis effects etc. need to be considered.

18. Standard textbook practice, we might parenthetically observe, is frequently misleading on these issues. Such books frequently claim to “prove” virtual work from Newton’s laws but these arguments can’t properly support the validity of variational reasoning as it is actually applied within real life engineering. For more detailed commentary on these issues, see my “What is ‘Classical Mechanics’ Anyway?” in Robert Batterman, ed, *The Oxford Handbook of the Philosophy of Physics* (Oxford: Oxford University Press, 2013).


20. In these abridgements, some of Iago’s personality traits sometimes get transferred to Othello himself.


22. Mathematicians are often much clearer about such issues, whereas physics presentations are frequently confusing. Jean Leray (“The Meaning of Maslov’s Asymptotic Method”) comments upon the proper nature of steady state asymptotics as follows:
But let us assume that [our target equation] is a continuous mechanical system: therefore the physicists cannot impose its initial condition and velocity: they have no more interest in Cauchy’s problem. What they are interested in are the waves.

23. This maneuver resembles the shift from normal Cauchy-style determinism to what I call “here-versus-there determinism” in WS.

24. This “pretend time” dodge makes a frequent appearance within so-called “Hamiltonian ray optics,” in a manner discussed in Essay 8.

25. Mathematically, this means that we should look at Fourier or Laplace transforms of the spatial response, which is usually the way the detailed argument is handled in the textbooks.