15 February 2008

Explanatory Idealizations

Word count: 6,024 words.

Abstract

A signal development in contemporary physics is the widespread use, in explanatory contexts, of highly idealized models. This paper argues that some highly idealized models in physics have genuine explanatory power, and it extends the explanatory role for such idealizations beyond the scope of previous philosophical work. It focuses on idealizations of nonlinear oscillator systems.

1. Introduction

Physicists since Galileo have explained natural phenomena making central use of approximations and abstractions. These explanations involve reasonably accurate models that are quite good (albeit simplified) representations of physical systems, both the explanatory factors and the natural phenomena to be explained. Indeed, it is precisely the representational accuracy of models that is taken to underwrite their explanatory power. However, a signal development in contemporary physics is the widespread use, in explanatory contexts, of highly idealized models which do not seem to fit this "Galilean" approach. Physicists appeal to these sorts of idealizations in their explanations, but we lack an account that makes sense of this practice. The idea that non-approximative idealizations may underwrite bona fide scientific explanation goes against orthodox views of scientific explanation. Ultimately I want to claim that at least some non-approximative, highly idealized models in physics have genuine explanatory power, and I want to extend the explanatory role for such idealizations beyond the scope of previous philosophical work. My focus will be on nonlinear oscillator systems and the explanation of highly idealized models of long-timescale behaviour using highly idealized models of short-timescale behaviour and asymptotic mathematical techniques. I shall propose that this type of explanation is a good one because it meets three criteria: first, the explanation is of a modified Deductive-Nomological form I shall call Deductive-Nomological-Idealization (*D-N-I*); second, it instantiates a general pattern of argumentation successful in a wide range systems, a criterion I shall call *unification*; third, the explanation is characterizable at the base (short-timescale) level, a criterion I shall call *basal derivability* (in Section 4). First, however, it will be helpful to look more closely at other types of explanations involving idealizations (Section 2) and then briefly to investigate how nonlinear oscillators are modelled (Section 3).

2. Galilean and semi-Galilean idealization

Galileo famously developed a range of idealizing techniques aimed at simplifying and explaining natural phenomena, techniques Ernan McMullin calls, following Husserl, "Galilean idealizations" (McMullin 1985). Galilean idealizations make use of simplified mathematical representations, causal decomposition, and "construct idealization" wherein a diagrammatic or theoretical model of the system is constructed incorporating false elements that approximate the real system. Idealizations of a simple pendulum, for example, may include assumptions that the bob is a point mass, the amplitude of oscillation is very small, there is no air resistance or friction, the wire is massless and inelastic, and the support is motionless.[1] As Galileo states explicitly in *De Motu*, he is able to "draw true conclusions even from false assumptions" (quoted in McMullin 1985, 255). All Galilean idealization is characterized by the fact that complementary to idealization are reverse techniques for adding back real-world details and de-idealizing by eliminating simplifying assumptions. Galilean idealizations thus have an intrinsic "self-correcting" feature such that they can (at least in principle) be brought in closer and closer agreement with empirical observations in a non-*ad hoc* way. McMullin calls Galilean idealizations "truth-testifying," because although they are not true, in the sense of perfectly representing the physical system, they admit to self-corrections that approach arbitrarily close to the truth, at least in principle (McMullin 1985, 264). Further, these corrections are not *ad hoc* because each one is independently justified. In the pendulum case, for example, we have good empirical reason to introduce corrections such as assuming a finite mass for the pendulum wire and allowing motion in the support.

Four characteristics of Galilean idealization are relevant to their role in the kind of scientific explanation that is our focus here. First, from a Galilean idealization can be derived results about behaviour of a physical system that approximate the actual behaviour to be explained. Second, this approximation can be improved, in principle, to any degree such that something arbitrarily close to the explanandum can be derived from the idealized model. Third, the self-correcting process improves the approximation by adding back neglected features of the system such that each correction to the Galilean idealization is itself is theoretically justified and thereby fully explained. Fourth, the approximations apply to

[1] These idealizations are only Galilean over a short timescale. Non-Galilean features appear in any damped oscillator over a long timescale (see Section 4).

components (spatial parts) of the system, the explanans, and the behaviour to be explained is determined by, and in some sense a product of, these parts.

Most discussions of the explanatory role of Galilean idealization assume a modified version of the Deductive-Nomological approach to explanation (Hempel 1965). As is well known, the D-N account faces some significant challenges if taken offering formal criteria specifying necessary and sufficient conditions on scientific explanation (see, e.g., Salmon 1984). As we shall see, ultimately the D-N approach will prove inadequate to analyze explanations involving non-Galilean idealizations (Section 4). For now, however, the D-N account provides a perspicuous way of viewing the structure of one type of explanation involving Galilean idealization, explanations of large-scale or structural features of systems in terms of their component parts and the laws of combination and interaction governing those parts. In this type of explanation, the lawlike premises included in the D-N explanans trace causal or structural dependencies in the physical system. They describe how the components combine and interact to produce the explanandum, and this substantive aspect underwrites the genuinely explanatory nature of the D-N derivations.

Galilean idealizations feature in D-N explanations by enabling the derivation of a conclusion, call it E*, that approximates, in the sense of differing negligibly from, E, the actual explanandum (typically the observed values). The characteristics of Galilean idealization just sketched ensure that the differences between E* and E are small and that these differences are themselves fully explained. On what I call the *Deductive-Nomological-Idealization* (D-N-I) account of explanation, Galilean idealizations explain the explanandum by entailing a conclusion that approximates the explanandum, without entailing the explanandum nor even rendering it probable. Our focus is on inter-level explanation of the large-scale behaviour of a system in terms of its component parts, and as with regular D-N explanations, the explanans must include at least one lawlike premise that describes how the components combine and interact to produce the explanandum. But rather than premises being true, as in D-N explanation, these premises are Galilean idealizations about the components of the system.[2]

Galilean idealizations are far more pervasive in many philosophers' *accounts* of science than they are in science itself. Some philosophers have brought attention to characteristics of models in contemporary physics that seem to fit poorly with Galilean idealization (Cartwright 1983; Morgan and Morrison 1999; Batterman 2002). For example, in hydrodynamic models of fluid flow around solid bodies, the Navier-Stokes equations are unsolvable analytically. On the boundary-layer approach, two distinct idealizations are used and the Navier-Stokes equation is solved in each idealized regime. A model of a frictionless, non-viscous fluid is used far from the boundaries, while close to the boundaries the Navier-Stokes equation is solved for a very thin layer of viscous fluid with a no-slip condition. The two models are then matched at their interface, from which dynamical relations can be derived that depend on the configuration of the boundary, the velocity of the fluid and its viscosity. Neither model provides an accurate or approximate representation of the observed fluid flow, and neither model can be incrementally corrected to provide one. However, the models enable the derivation of structural dependencies in systems with fluid flows across solid bodies. As Morrison puts it, the models explain "'how possibly' certain kinds of behaviour take place" (Morrison 1999, 63). Although the models fail accurately to represent many of the details of the observed behaviour, they do enable us to represent these structural

[2] For two developments of D-N explanation along these lines, see (Causey 1977; Elgin and

characteristics of a class of physical systems. “The reason that models are explanatory is that in representing these systems they exhibit certain kinds of structural dependencies” (Morrison 1999, 63). Although the models include non-approximative (non-Galilean) idealizations of components of the system, the end result is an approximate and accurate representation of the large-scale structure of the system. As Morrison emphasizes, “the explanatory role is a function of the representational features of the model” (Morrison 1999, 64).

Other examples of idealizations which fail to fit the Galilean approach have been investigated in detail by Robert Batterman (Batterman 2002; Batterman 2005a; Batterman 2005b). Batterman focuses on physical systems wherein base-level (or “fundamental”) theory breaks down, including statistical mechanical models at criticality, the semi-classical limit of quantum mechanics, the breakdown of the wave theory of light in catastrophe optics, and drop formation in hydrodynamics. These models include idealizations of components of the system that fail to approximate the system itself: an infinite number of molecules in statistical mechanical models at criticality, zero-wavelength solutions in optics, etc. A key feature of these idealizations is their structural stability or stability under perturbations at the base level, a feature physicists call universality. These structural features are revealed through asymptotic analyses in which base-level details are systematically eliminated using asymptotic mathematical techniques. A paradigm case of this approach are renormalization group techniques used to analyze thermodynamic systems at critical points and to derive structural features such as the critical exponent β (Batterman 2002, pp. 37-42). In this way, asymptotic analyses enable idealized models accurately to characterize the underlying

---

Sober 2002).

structural or universal features—without, as in the Galilean case, making use solely of approximate representations of the component parts of the system.

> We can conclude that one reason (and, perhaps, the essential reason) that asymptotic analyses so often provide physical insight is that they illuminate structurally stable aspects of the phenomenon and its governing equations (known or not)…. As a result, we can be confident that our theories adequately mirror or represent the world (Batterman 2002, p. 59).

These asymptotic analyses explain structural features of physical systems to the extent that they enable the derivation of accurate representations of these features using asymptotic techniques. As Batterman puts it, the derived structural stability "is what leads us to take the principle features resulting from asymptotic analyses to be explanatory" (Batterman 2002, p. 57).

Call the sorts of models just discussed *semi-Galilean idealizations*. They are used to understand physical systems in which Galilean idealizations are uninformative or simply break down altogether. As we have seen, semi-Galilean idealizations differ from Galilean ones in that the former centrally involve idealizations that are not approximations of base-level component features of the physical system, and which no amount of incremental reverse-engineering will correct. However, both semi-Galilean and Galilean idealizations provide idealized descriptions of lower-level components in the system, components which combine and interact to produce upper-level structural features. Both aim to provide accurate, approximate representations of these structural features, and both aim to explain these features on the basis of their lower-level component parts. Thus, semi-Galilean idealizations feature in D-N-I explanations as well. Premises in the explanans include idealized descriptions of the components of a system, initial and boundary conditions, and lawlike generalizations. In the semi-Galilean case, some aspects of the idealized descriptions fail to approximate the components, and the lawlike generalizations include asymptotic techniques (such as renormalization group techniques). However, the derived explanandum, E^*, approximates E, the physical behaviour to be explained. In the type of semi-Galilean explanation we are focusing on, what is being represented and ultimately explained is a structural feature of a specific physical system (e.g., its critical exponent β) or the stability of such a feature (e.g., the fact that systems with diverse base-level components have the same critical exponent β). The argument contained in the explanans quite accurately derives facts about the structure of the system and its stability (or universality) based on an idealized account of base-level component parts.[3]

One might worry that in the semi-Galilean case the explanation makes essential use of premises that are not just false, but non-approximatively and incorrigibly so. This worry is unfounded, for whatever the explanatory merits of Galilean idealizations, they are shared by semi-Galilean idealizations. The present paper takes the approach that at least some Galilean idealizations are genuinely explanatory. Thus, one type of scientific explanation provides a deductive argument whose premises include an idealized description of the base-level components, initial and boundary conditions, and one or more lawlike generalizations that trace causal or structural dependencies, and whose conclusion approximates the upper-level physical behaviour or characteristic to be explained. As we have seen, Galilean and semi-

[3] On the approach taken here, Batterman's asymptotic explanations are thus a kind of D-N-I explanation, one in which the explanandum is the universality characteristic of the system (Batterman 2002, pp. 37-60). Batterman's term "asymptotic explanation" is potentially misleading because, as we shall see in Section 4, similar asymptotic mathematical techniques are used in a quite different type of explanation.

Galilean idealizations share all these features and thus, on this approach, are genuinely explanatory on the same grounds. As Morrison and Batterman emphasize, the explanatory merits of their (semi-Galilean) cases rest in part on the accurate and approximate representation of the explanandum, just as in the Galilean case. However, one may take an alternative approach, namely that Galilean idealizations have no explanatory power themselves but rather function provisionally, as markers of progress towards actual (true) explanations based on fundamental, perfect, non-idealized models of the physical system. On this approach, explanations involving Galilean idealizations are at best incomplete or partial, and they are eliminable in favour of true explanations. Again, in the type of explanation we have been looking at semi-Galilean idealizations can play exactly the same role.[4] The more general worry here where to draw the line: should we admit as scientific explanations only D-N explanations with true premises, should one broaden the scope of scientific explanation to include Galilean idealization, or should one go further still and include semi-Galilean idealization? I contend that the line cannot be drawn between explanations involving Galilean and semi-Galilean idealization.

In the Galilean and semi-Galilean idealizations we have looked at, explanatory power followed from a D-N-I account of how base-level components combine and interact to produce upper-level structural features of the system. The remainder of this paper extends an explanatory role for idealizations to physical systems in which this is not the case. It suggests that some highly idealized models in physics have genuine explanatory power that does not

---

[4] For an argument along these lines in the context of Batterman's asymptotic explanation, see (Belot 2005 and Batterman 2005b). For an argument that the "perfect model" approach is misguided and inapplicable to science, see (Teller 2001).

rest on their ability approximately or accurately to represent the effects of base-level component parts. I shall call the idealizations involved in these explanations *non-Galilean*. Non-Galilean idealization plays a key role in understanding nonlinear oscillators, to which we now turn.

3. Nonlinear oscillators

A divide-and-conquer strategy common in physics uses one model to cover base-level characteristics of a system and another model to cover high-level phenomena. For instance, models of statistical mechanics treat fluids as composed of interacting point particles, while, at a larger spatial scale, hydrodynamics treats the same system as composed of a continuous fluid. We have seen several other examples of this strategy already. However, not all divisions between models occur along spatial scales. The modelling technique in the effective field theory program in quantum field theory is to use distinct models for distinct energy scales. In the kind of models we shall be looking at, the base-level models describe short-timescale behaviours while upper-level models describe long-timescale behaviours. Our focus is on nonlinear oscillator systems, and the divide-and-conquer strategy is used successfully to produce idealized linear models of aspects of the larger, unmodellable nonlinear system. It will be useful to start with a brief look at a simple nonlinear oscillator system, the weakly-damped van der Pol oscillator (for a more detailed development of this case, see [self-reference omitted]). As I shall argue, this oscillator is an example of a non-Galilean idealization playing a substantive explanatory role, and, more importantly, is just one instance of a quite general pattern of such explanations in nonlinear oscillator dynamics.

The one-dimensional nonlinear van der Pol oscillator was originally investigated by B. van der Pol as a model of the human heart (van der Pol and van der Mark 1928). It also describes some oscillatory vacuum tube and electronic circuits. The oscillator undergoes near-harmonic oscillations with weak positive and negative damping, resulting in a slow variation of frequency and amplitude over time. A key characteristic of the van der Pol oscillator is its limit cycle behaviour. Whatever initial non-zero value of the amplitude, over time it will converge to a stationary limit-cycle amplitude. For any short-timescale snapshot of the oscillator, of the order of a small number of oscillations, the dominant behaviour will be the near-harmonic oscillations. Indeed, over a short timescale, the behaviour of the van der Pol oscillator is indistinguishable from that of a simple harmonic oscillator with fixed amplitude and frequency. By contrast, over the long timescale, of the order of convergence to the limit cycle, individual oscillations become irrelevant and behaviour is dominated by the rate of change of amplitude. These are two very different sorts of behaviour, and this should motivate the use of two different models to represent the behaviour of this nonlinear system. Figure 1 presents a schematic view of these two sorts of models. Short-timescale (base-level) models are simple harmonic oscillators. They have nothing to say about amplitude variation, limit cycles or long-term behaviour. Rather, they simplify the short-timescale behaviour by treating it linearly. Long-timescale models are obtained by reducing the resolution, or coarse graining, resulting in a linear model of the change in amplitude over time. Long-timescale models have nothing to say about the phase and amplitude of oscillation. The two families of models describe distinct aspects of the same physical system. This process of modelling simple, linear behaviours of interest in complex nonlinear systems is characteristic of what physicists do.

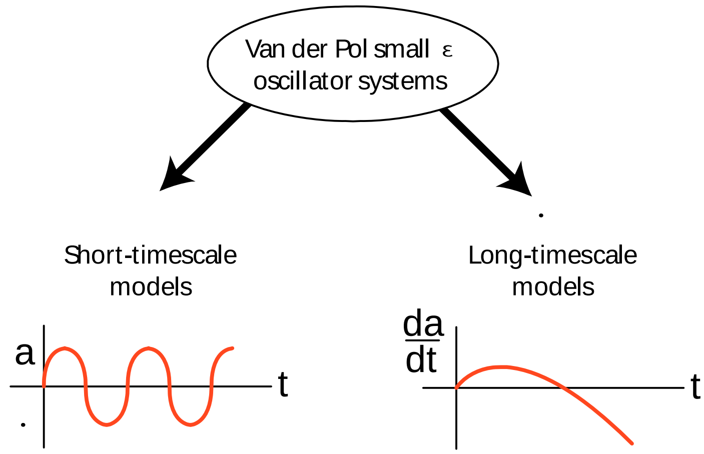

Figure 1. Two kinds of models of van der Pol systems

The short-timescale and long-timescale models are related, of course. One might be inclined to think long-timescale models straightforwardly reduce to short timescale models. However, they do not in the van der Pol case, nor do they in models of nonlinear oscillators more generally. In the van der Pol case, the base-level, short-timescale model, plus an assumption that the amplitude and phase of oscillation vary only slowly with respect to the period of oscillation, plus asymptotic mathematical techniques (involving the replacement of divergent series by an asymptotic cut-off series) are used to derive, to an arbitrary degree of accuracy, the behaviour of the upper-level, long-timescale model. The second-order solution provides a quite accurate prediction of the observed variation in amplitude over time at the upper level and the stability of the limit cycles under perturbation. The van der Pol oscillator is a characteristic example of a self-excited, damped nonlinear oscillator. Self-excited oscillators typically have a number of stationary states of oscillations or limit cycles. These systems are largely insensitive to initial conditions, as we have seen, in the sense that over long timescales ($t \rightarrow \infty$) the behaviour of the systems is determined by intrinsic features of the system. Intertheoretic relations between models of nonlinear oscillators at distinct levels is subtle and interesting, but the details of these relations are beyond the scope of the present paper (see [self-reference omitted]).

Weakly-damped nonlinear oscillators are described by equations of the form

(1) $$y'' - \varepsilon F(y, y') + y = 0 \qquad 0 < \varepsilon << 1$$

where $y(t)$ is the oscillation variable and F is a nonlinear polynomial. This nonlinearity means that the oscillator equation cannot be solved exactly, nor, in general, can it be solved using approximation techniques involving regular limits, such as regular perturbation methods (regular perturbation methods will not work in systems with dissipative forces in which amplitude is time-dependent). Rather, methods used to solve the equation make use of divergent asymptotic series which involve singular limits. These asymptotic methods divide into two main groups, methods of slowly-varying amplitude and phase (such as the Krylov-Bogoliubov-Mitropolsky technique) and multi-timescale expansion methods (Mickens 1981). As we saw in the van der Pol case, asymptotic methods require additional premises and, mathematically, replacement of the divergent series (at the singular limit) with a convergent cut-off series. Thus, solutions to weakly-damped nonlinear oscillator equations follow a common *pattern*, summarized as follows.

1) A nonlinear differential equation describing the system is posited and transformed to dimensionless variables so it has the form of (1).
2) At the base level, short-timescale behaviour is modelled by a simple harmonic oscillator ($\varepsilon = 0$). Phenomena of interest are identified at one or more upper levels (longer timescales).

3) An asymptotic method is used to first order in ε to derive rough approximations of the behaviour at longer timescales (the method of slowly varying amplitude and phase or a multi-timescale expansion).

4) The upper-level model(s) derived in (3) are compared with the phenomenology of the system at long timescales. If more precise description is needed, one or more asymptotic methods are used to calculate higher-order corrections to the long-timescale behaviour.

This argument pattern, which is filled in in specific instances in much detail, works in a broad range of weakly-damped nonlinear oscillator systems.

The approach taken to modelling features of weakly-damped nonlinear oscillators generalizes. Weakly-damped nonlinear oscillators are a subset of nonlinear oscillator systems, and an argument pattern applies to the latter systems that is similar to but somewhat broader than the one sketched above. For example, the short-timescale behaviour of non-weakly damped oscillators will be periodic but not necessarily harmonic, so the argument pattern will make use of a range of harmonic and non-harmonic short-timescale models in step 2. Nonlinear oscillators are themselves a subset of nonlinear systems more generally, including chaotic and other non-periodic systems. These systems typically have physical phenomena of interest at two or more levels, and these levels may be distinguished in a variety of ways, including spatial scale, timescale or energy scale. I suggest that explanations of features of idealized models of nonlinear systems at one level in terms of features at one or more lower levels follow a number of argument patterns that share a family resemblance with the argument pattern sketched above. Each system, for example, is treated in terms of two or more idealized models, as a rule linear ones, in just the way that the van der Pol small-ε oscillators are well represented by a linearized harmonic oscillator at one level and by a linearized model of the change in amplitude and another level. Models of nonlinear systems have similar properties and relations both to other models representing the same system and to other models representing other systems. As well, there will be similar derivational relations, in terms of asymptotic techniques, between model pairs, of the kind we have seen in the van der Pol case. While it is beyond the scope of the present paper to a undertake a detailed survey of all models of self-excited weakly-damped oscillators, weakly-damped oscillators, nonlinear oscillators or nonlinear systems, I hope to have made plausible the idea that the derivations share a set of argument patterns, and that these argument patterns themselves share certain similarities.

4. Non-Galilean explanatory idealization

What explanatory relations, if any, obtain between base-level and upper-level families of models in van der Pol systems? *Prima facie*, it seems that the explanandum, the observed limit cycle behaviour, is explained by the base-level model on the same grounds as Galilean explanation: from the base level it is possible to derive something quite close to the explanandum, and that derivation can be made, in principle at least, to yield a result arbitrarily close to the explanandum (although derivational complexity above second order is a significant practical limitation). However, idealization in the van der Pol case lack two Galilean features that were taken to underwrite the explanatory power of idealizations in the first place. First, the base-level model does not contain an approximation of the dynamics of the actual system. The base-level model is that of a simple harmonic oscillator with fixed amplitude and phase. A key feature of the van der Pol oscillator, and nonlinear systems more generally, is the failure of a regular limit relation.

(2) $$\lim_{\varepsilon \to 0} T_{\varepsilon} \neq T_{e=0} .$$

As we have seen, the dynamics of the van der Pol oscillator (T_ε) is qualitatively different from the dynamics of the ε = 0 simple harmonic oscillator ($T_{\varepsilon=0}$). In the same way, the upper-level model picks out one characteristic of the behaviour of the system, the change in amplitude, but necessarily omits the oscillatory feature of the system. Secondly, no amount of incremental reverse-engineering will transform the base-level or upper-level models into more accurate representations of the system. In fact, the base-level model does not even *aim* to approximate the system; rather, as we have seen, both the base- and upper-level models aim to capture a simple, linear behaviour of interest in a complex nonlinear system.

With respect to the failure of a regular limit relation to obtain, failure to approximate base-level features of the system and their incorrigibility in this regard, the van der Pol case is similar to semi-Galilean idealizations described in Section 2. In another, highly consequential respect, the van der Pol case is quite different. As we have seen, semi-Galilean idealizations are explanatory along D-N-I grounds in a way analogous to Galilean idealizations. In the semi-Galilean case, base-level components are spatial parts of the upper-level description of the system. The key consideration was that the explanans described how the upper-level behaviour to be explained was produced by the base-level component parts and lawlike generalizations. It is here that the van der Pol model differs substantially from the semi-Galilean case. The base-level, short-timescale behaviour is not a spatial part of the upper-level, long-timescale explanandum. And this difference will be crucial for assessing the explanatory merits of van der Pol models and models of nonlinear oscillators more generally. I have called non-Galilean those idealizations that fail approximately or accurately to represent the effects of base-level component parts. In the van der Pol case, as in nonlinear oscillators more generally, there are no base-level spatial parts of the system. The base-level idealization describes short-timescale parts of the system, but this is not analogous to the way spatial components in Galilean and semi-Galilean systems combine and interact to produce upper-level (large-scale) behaviour. In the van der Pol case, there are no structural dependencies, of the sort highlighted in semi-Galilean systems discussed in Section 2. Short-timescale behaviours are not components of the system and do not combine and interact to produce an effect, so the idealizations of the van der Pol case (and cases like it) are non-Galilean.

In the van der Pol case the explanation fits the D-N-I account, so it would seem that non-Galilean idealizations can underwrite scientific explanations. But here we must proceed carefully. We have seen how the explanandum (limit cycle behaviour) is derived based on facts about the base-level model, the van der Pol equation, and asymptotic mathematical techniques. The problem, however, is that the D-N-I account is not a sufficient condition on scientific explanation. We do not want to admit as a *bona fide* scientific explanation just any old idealization that fails even approximately to model the intended physical system and for which derivation of the putative explanandum requires invoking additional premises distinct from the idealization itself. The worry is that if we allow the van der Pol explanation we will have no principled way of debarring all sorts of patently gerrymandered pseudo-explanations. For instance, the base level van der Pol system could be modelled by a periodic non-harmonic function, such as a square wave. Here, the standard asymptotic derivation of limit cycles does not go through. Instead, introduce a premise that the change in amplitude over the long timescale corresponds to the observed behaviour, and the desired result can be derived. This derivation begs the question and obviously fails to have any explanatory merit. D-N-I explanation, as D-N explanation, needs to be supplemented with constraints on the allowable types of derivations and lawlike generalizations that are explanatory. In the Galilean and semi-Galilean cases, D-N-I derivations were explanatory precisely because they traced structural dependencies in the physical system, particularly with respect to the way in which base-level components combined and interacted to produce upper-level structures. Idealizations, disconnected from their Galilean moorings, seem lost in a sea of false premises and *ad hoc* lawlike generalizations, and they seem particularly poor candidates to underwrite scientific explanation.

Nevertheless, I would like to suggest that some non-Galilean idealizations *do* have explanatory power. To see how, we need to broaden our thinking: the explanatory merits of a derivation are not only a function of features internal to that derivation (the truth of premises, the content of lawlike generalizations) but also a global function of how that derivation fits into a larger pattern of derivations in the scientific field. The remainder of this paper sketches and approach to explanatory idealization that provides a perspicuous way to take into account this global aspect of scientific explanation.

The key requirement is *unification*, which ensures that the D-N-I derivation is not *ad hoc* but rather forms part of larger theoretical structure applicable to a broad range of physical systems. On an approach to explanatory unification due to Philip Kitcher (Kitcher 1981, Kitcher 1989), explanations are arguments (deductive derivations) that are appropriately connected to a larger pattern of argumentation in a field of science. One form of unification is derivational parsimony, where one or a small number of argument patterns are used in a broad range of explanations ([self-reference omitted]). These argument patterns share a common structure and a common toolbox of associated models, mathematical techniques and empirical assumptions that underwrite the premises in the derivations. This is clearly the case for nonlinear oscillators. We have seen how a family of asymptotic methods, a common base-level model and a set of common additional assumptions form part of the explanation of long-timescale features of weakly-damped nonlinear oscillators, following a simple four-step structure. My claim is that the derivation of limit cycles and their properties in the van der Pol system forms part of a unified pattern of such derivations for nonlinear oscillators more generally. These arguments succeed in deriving quite accurate accounts of a wide range of long-timescale behaviours while making use of modest explanatory resources within a fixed set of patterns of argumentation. In short, the derivation of the limit cycles and their stability properties in the van der Pol oscillator is explanatory because it is part of a pattern of such derivations that uses few argument patterns to explain and predict a large range of results (for a more detailed development and defence of explanatory unification, see [self-reference omitted]).

A common worry about unification approaches to explanation is the spectre of the obsessive unifier who arbitrarily constructs spurious argument patterns to link otherwise disparate systems. In our case, explanatory unification would seem to allow for the most unified set of argument patterns to be high-level descriptions of common features of nonlinear oscillators, clearly unexplanatory. To preclude this, an additional requirement of *basal derivation* is needed, stipulating that explanations of upper-level features of physical systems must be in base-level terms. As we have seen, non-Galilean idealizations in the van der Pol case play a central role but form only one part of the explanans. The novel upper-level properties of the van der Pol oscillator—the existence of limit cycles, their precise structure and their stability— are derived using asymptotic methods involving a lower-level model of the system, additional empirical assumptions, and mathematical techniques. The asymptotic method includes an assumption of slowly varying amplitude and phase that is itself characterizable in lower-level terms. The key point is that empirical premises in the explanation can be articulated entirely in terms used to describe the lower-level model, even though the premises are not part of the description of the model itself (indeed, the premises are inconsistent with the model in our example). In short, in the van der Pol case, the explanans satisfies the criterion of basal derivation because it contains premises about the lower-level model, about asymptotic mathematical techniques, and additional assumptions that lie outside the scope of the model but are describable in basal terms.

The explanations we have been looking at in nonlinear oscillator systems are of the D-N-I form and satisfy the criteria of unification and basal derivation just described. The idealizations involved in the explanation are non-Galilean yet, I have suggested, provide a excellent cases of idealizations playing substantive explanatory roles.

5. Conclusion

One might object, in the end, that without a description in the explanans of how component parts combine and interact to produce the explanandum (even in the attenuated semi-Galilean sense) we have a very weakened sense of explanation indeed. Limit cycle behaviour in the van der Pol case has been derived, to be sure, but it has not been *explained* because we have no account of how the limit cycle behaviour arises in causal or structural terms. Although there are some superficial similarities between the semi-Galilean and non-Galilean cases we have looked at, especially with respect to the use of asymptotic techniques, features of base-level non-Galilean idealizations simply fail to be explanatorily relevant in the way that they are in semi-Galilean cases.

Nothing said here precludes this kind of objection. However, it raises a worry about the basis of the explanatory merits of semi-Galilean idealizations in the first place. The problem is that semi-Galilean idealizations, which fail even to approximate features of base-level components, seem to be poor candidates for explaining upper-level structure. Recall that in semi-Galilean cases, D-N-I derivations were taken to be explanatory because they trace structural dependencies in the physical system, particularly with respect to the way in which base-level components combine and interacte to produce upper-level structures. The ways in which semi-Galilean idealizations mis-represent features of base-level components leads to scepticism about the extent to which they can be said to truly to trace physical dependencies at all. The explanations discussed in Section 2 involving semi-Galilean idealizations are good ones, the worry goes, but it is implausible that the explanatory relevance of these idealizations follow from their ability to trace causal structure.

The solution, I recommend, is to give serious consideration to an alternative perspective, favoured by empiricists, on which causal relations are derivative from explanatory relations (see Kitcher 1989, pp. 494-500). A derivation does not get its explanatory power by tracing causal structure; rather, a derivation is taken to trace causal structure (or at least an idealization thereof) when it is explanatorily unified. So it may seem that the explanatory merits of semi-Galilean idealizations are underwritten by universality in physical systems while the explanatory merits of non-Galilean idealizations are based on universality in our theoretical structure. In fact, the former is but an aspect of the latter.

References

Batterman, R. W. (2002). The Devil in the Details : Asymptotic Reasoning in Explanation, Reduction, and Emergence. Oxford ; New York, Oxford University Press.

Batterman, R. W. (2005a). "Critical Phenomena and Breaking Drops: Infinite Idealizations in Physics." Studies in History and Philosophy of Modern Physics **36B**(2): 225-244.

Batterman, R. W. (2005b). "Response to Belot's "Whose Devil? Which Details?"" Philosophy of Science **72**(1): 154-163.

Belot, G. (2005). "Whose Devil? Which Details?" Philosophy of Science **72**(1): 128-153.

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

Causey, R. L. (1977). Unity of Science. Dordrecht, Holland D. Reidel Pub. Co.

Elgin, M. and E. Sober (2002). "Cartwright on Explanation and Idealization." Erkenntnis **57**(3): 441-450.

Hempel, C. G. (1965). "Aspects of Scientific Explanation." Aspects of scientific explanation. New York, The Free Press: 331-496.

Kitcher, P. (1981). "Explanatory Unification." Philosophy of Science **48**: 507-531.

Kitcher, P. (1989). "Explanatory Unification and the Causal Structure of the World." Minnesota Studies in the Philosophy of Science, volume XIII. P. Kitcher and W. C. Salmon. Minneapolis, University of Minnesota Press: 410-506.

McMullin, E. (1985). "Galilean Idealization." Studies in History and Philosophy of Science **16**(3): 247-273.

Mickens, R. E. (1981). An Introduction to Nonlinear Oscillations. Cambridge England, Cambridge University Press.

Morgan, M. S. and M. Morrison, Eds. (1999). Models as Mediators: Perspectives on Natural and Social Science. Cambridge, Cambridge University Press.

Morrison, M. (1999). "Models as Autonomous Agents." Models as Mediators: Perspectives on Natural and Social Science. M. S. Morgan and M. Morrison. Cambridge, Cambridge University Press.

Salmon, W. (1984). Scientific Explanation and the Causal Structure of the World. Princeton, Princeton University Press.

Teller, P. (2001). "Twilight of the Perfect Model Model." Erkenntnis **55**: 393-415.

van der Pol, B. and J. van der Mark (1928). "The Heartbeat Considered as a Relaxation Oscillation, and an Electrical Model of the Heart." Phil. Mag. Suppl. **6**: 763-775.