

1 February 2010

Expanding the Scope of Explanatory Idealization

Andrew Wayne

Department of Philosophy, University of Guelph

Draft – please do not cite

Abstract

Many explanations in physics rely on idealized models of physical systems. These explanations fail to satisfy the conditions of standard normative accounts of explanation. Recently, some philosophers have claimed that idealizations can be used to underwrite explanation nonetheless, but only when they are what have variously been called representational, Galilean, controllable or harmless idealizations. This paper argues that such a half-measure is untenable and that idealizations not of this sort can have explanatory capacities.

1. Introduction

Knowing *why* is a singular achievement, distinct from other scientific accomplishments. Science aims at describing and representing nature, predicting and controlling it; but science also aims at explanation. One standard approach to scientific explanation holds explanations to be deductive arguments with true premises. Not all such arguments are explanations, though. Characteristics that make some deductive arguments explanatory include that the statements in the *explanans*—the premises in the argument—include a relevant scientific law

and background conditions (Hempel 1965), or that they fall under widely-used patterns of argumentation (Kitcher 1989). Another standard approach holds that they give a true causal or counterfactual story relevant to the occurrence of the *explanandum*—the phenomenon (behaviour, regularity, structure) to be explained (Salmon 1984). One thing philosophers do generally agree on, however, is that statements in the explanans be true, whether they are about some feature of the system, about a relevant natural law, or about a causal or counterfactual relation doing explanatory work. It seems a reasonable requirement that the statements adduced to explain should describe laws, regularities, causal relations, properties, structures, and so on, that obtain in the physical system exhibiting the explanandum.

The present paper focuses on explanation in physics. It limits discussion to deductivist approaches to explanation, on the assumption that these approaches are more likely to be relevant across a broad range of fields of physics. In many fields, there is no causal story to be told, or at least such a story is not part of standard physics. This and other considerations to follow aim to motivate a deductivist approach, but it is beyond the scope of the present paper to defend it.

Our starting point is an observation about explanation in physics that may come as a surprise: virtually all cases of what physicists take to be *bona fide* scientific explanation fail to satisfy even the basic requirements just articulated. Explanation in physics relies essentially on *idealizations* (idealized models) of physical systems, and the explanations themselves contain false statements about both the explanatorily relevant features of the physical system and the phenomenon to be explained. Some philosophers of physics have recognized this, and have responded in various ways that our notion of scientific explanation should be broadened somewhat to accommodate these practices. More specifically, they have

claimed that idealizations can be used to underwrite explanation, but *only* where the premises in the explanans are approximately true of the target system and are fully corrigible, at least in principle. This paper argues that such a half-measure yields an account of explanation that is untenable. As we shall see, many putative explanations in physics are based on idealizations that fail to meet these conditions. These are explanations in which the idealized model doing explanatory work does not successfully represent the physical system. The paper suggests that the close link hitherto assumed between successful representation and explanation should be loosened.

2. Explanation via Galilean Idealization

An idealized model is known not to represent accurately some elements of the target system. Abstractions leave elements of the physical system out of the model; approximations simplify and misrepresent features of the system in various ways; other sorts of idealization may posit structures in the model not present in the system, or worse, even physically impossible. Galileo famously developed a range of idealizing techniques aimed at predicting and explaining natural phenomena. Galileo's "idealized construct," as he called it, of a simple pendulum includes the assumptions that the pendulum is not subject to air resistance, that the wire is massless and inelastic, and that there are no other material hindrances or imperfections. He hypothesized that an ideal pendulum would continue to oscillate indefinitely with the same amplitude and period and that it would obey his pendulum law: "As to the ratio of times of oscillation of bodies hanging from strings of different lengths, those times are as the square roots of the string lengths" (Galilei 1638/1989, 97). Now, Galileo well knew that this failed to describe and predict accurately the behaviour of any of the real pendulums he used in his extended and painstaking experimental work. The

oscillations of real pendulums get smaller and smaller over time and they are not isochronous, as his pendulum law requires. But Galileo, and generations of physicists since, have taken the pendulum law to be part of the explanation of the behaviour of physical pendulums. And therein lies a problem: none of the standard philosophical accounts of explanation canvassed at the outset makes sense of this sort of explanatory practice.

There is a response to this problem in the literature, first articulated by Ernan McMullin (McMullin 1985). McMullin proposes that a handful of characteristics pick out idealized models that can underpin scientific explanation, models he dubs *Galilean idealizations*. Galilean idealization is characterized by the fact that the idealized model approximates the target system and, more importantly, 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 ever closer agreement with empirical observations in a theoretically justified, non-*ad hoc* way. Laurence Sklar and Robert Batterman make a similar distinction between what they call controllable and uncontrollable idealizations. “An idealization is controllable means that it is possible, via appeal to theory, to compensate in some way for the idealization,” whereas uncontrollable idealizations typically involve singular limits and preclude explanation (Batterman 2005, 235). Mehmet Elgin and Elliott Sober sketch an account of explanation based on this sort of distinction, calling the types of idealizations that can underwrite explanation “harmless idealizations” (Elgin and Sober 2002).

Two posits seem to underlie the idea that Galilean (controllable, harmless) idealizations have explanatory power. First is the notion that all idealizations involved in these sorts of

cases are approximative, so explanations given, while strictly speaking they apply only to the idealized model, are not too far off when applied to the physical system of interest. Second, it is possible systematically to refine the idealization to bring it closer and closer to the target system such that the statements in the explanation become, in the limit, true of the physical system as well.

I believe that motivating and supporting these assumptions is a deeper intuition about the connection between representation and explanation. The reason Galilean idealizations are taken to support scientific explanation is that these idealizations achieve a kind of common-sense representational success. Philosophers of science are inclined to say that some scientific models are *about* bits of the world, and that models can do a more or less successful job of representing those bits of the world. What exactly constitutes a more successful representation is a thorny question, with considerations of similarity (Giere 1988) or partial isomorphism (da Costa and French 2003) between elements of the model and elements of the physical system playing leading roles. Ideas such as these surely reflect what physicists consider successful representations of target systems in Galilean cases. In what follows I shall use the term “successful representation” to refer, somewhat imprecisely, to these sorts of widely-held ideas. Galileo’s simple pendulum models are successful representations, in this sense, because they are similar to physical pendulums in what are taken to be obviously relevant ways. Thus, while statements in the explanans may not exactly describe the physical system, and while the explanans may not entail or raise the probability of statements about the actual explanandum (the behaviour, regularity or structure to be explained), certain features of the idealized model are alike enough to the real system to underwrite the explanatory capacities of the idealization. As Margaret Morrison puts it, “the explanatory role is a function of the representational features of the model” (Morrison 1999,

64). This has been a central assumption of much philosophical work on modeling (Morgan and Morrison 1999).

Let us call this strategy *explanation via Galilean idealization* (EvGI). EvGI maintains that there is a significant difference between Galilean (controllable, harmless) idealizations and other idealizations, and only the former underwrite scientific explanation. It will be helpful to make this last claim a bit more precise in the context of the covering law (deductive-nomological) approach to explanation. Galilean idealizations feature in covering-law explanations by enabling the derivation of a conclusion that approximates, in the sense of differing negligibly from, the actual explanandum-statement. We can sum up the characteristics of Galilean idealization just sketched in terms of the following conditions on the explanans and explanandum.

Explanans condition. The premises in the explanans are true of the idealization and approximately true of the target system, and they are fully corrigible, at least in principle.

Explanandum condition. Differences between the conclusion derived from the explanans and the actual explanandum-statement are small and are fully corrigible, at least in principle.

Both conditions are based on the intuition, noted above, that statements figuring in successful explanations describe elements of an idealization that one would normally be inclined to say “successfully represent” or “are about” relevant features of the physical system. The EvGI strategy maintains that explanation as a normative goal of science can only be achieved in the context of Galilean idealization, for only Galilean idealization ensures that these conditions are satisfied.

3. The Challenge from Non-Galilean Idealization

The trouble is, Galilean idealizations are far more pervasive in philosophical *accounts* of physics than they are in physics itself. Physicists offer explanations of phenomena based on idealizations which are not Galilean. These are idealizations that cannot plausibly be said successfully to represent a physical system, in the sense in which we are using the term here. They are explanations for which the explanans condition fails to obtain. In short, a large part of the explanatory practice in physics simply does not fit the EvGI strategy.

Some philosophers have brought attention to characteristics of models in contemporary physics that are taken to underwrite explanations by physicists yet fail to meet the standards of Galilean idealization. In these sorts of models, which I shall call *non-Galilean idealizations*, certain, putatively explanatorily relevant, elements of the model cannot plausibly be regarded as being *about* the physical system, in the sense of successful representation articulated above. Statements about these elements, which figure in the explanans, are not even approximately true of the physical system. Moreover, we have reason to believe in these cases that this situation cannot be ameliorated, even in principle; there are good arguments to the effect that it is not possible systematically to refine the idealization to bring these elements in closer and closer agreement with the target system. As we shall see, in these sorts of cases the explanans condition is violated.

Examples of what physicists take to be explanatory idealizations that fail to fit the Galilean approach have been investigated in detail by Robert Batterman (Batterman 2002; Batterman 2005a; Batterman 2005b; Batterman 2009). Batterman focuses on physical systems wherein base-level (or “fundamental”) theory breaks down, including statistical mechanical models at criticality, 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. In statistical mechanics cases, explanations of critical behaviour, for example in phase transitions from solid to liquid or liquid to gas, are based on idealizations in which the number of molecules and correlation length go to infinity. These and many other features of the idealizations do not approximate the physical system, nor can they be incrementally eliminated to enable the idealization to represent more successfully the physical system. Finally, physicists take these non-Galilean features to be essential to the explanations proffered of observed phase-transition behaviour.

Other examples of non-Galilean idealizations physicists take to have explanatory capacities might include semi-classical and quantum chaos models (Bokulich 2008), computer-generated simulations of physical systems, as for example in climate modelling (Parker 2006), and nonlinear oscillator systems ([self-reference omitted]).

It will be instructive to consider Galileo's pendulum model in somewhat more detail.¹ Based on his work on freely-falling objects, Galileo developed a model of air resistance in which resistance increased linearly with the velocity of the object. We now know that for small oscillations and with linear damping (as Galileo modeled it), the equation of motion for the pendulum is of the form

$$(1) \quad \ddot{\theta} + 2\alpha\dot{\theta} + \omega_0^2 = 0,$$

¹ The remainder of the section develops an example of an idealization that is non-Galilean yet putatively explanatory. Readers not interested in the details of the physics should skip directly to the start of Section 4 on p. 13.

where θ is the angular displacement of the pendulum from vertical, ω_0 is the initial frequency ($2\pi/\text{period}$), and α is a parameter that is typically set to optimize fit with the observed phenomena. With appropriate initial conditions this equation can be solved exactly to yield expressions both for the decrease in the amplitude of oscillation and the increase in its period due to air resistance (Nelson and Olsson 1986, 115). In this way, we can compensate for the discrepancies between the idealization and the physical system in a theoretically justified and non-ad hoc way. This seems a paradigm case of a Galilean idealization.

But it is not. This is because for real pendulums of the sort Galileo used, the damping effect of air resistance on the pendulum bob is not linear. It can be shown that air resistance produces a force that has a quadratic component, based on a calculation of the Reynolds number for the pendulum bob in air at maximum velocity. The equation of motion of a pendulum subject to damping that is quadratic in velocity is of the form

$$(2) \quad \ddot{\theta} - \varepsilon \dot{\theta}^2 + \omega_0^2 \theta = 0,$$

where ε is a small dimensionless parameter related to the strength of the damping. In fact, two equations of motion are required, one for each half-period, with the sign of the epsilon term reversed in the second equation as the angular velocity changes sign in the second half-period. We do not here consider the second equation or matching conditions, as we are interested in the correction to the period of oscillation, which is the same in both half-periods. As we shall see, the idealization in this case is non-Galilean and it seems to underwrite a full explanation of the quadratic damping behaviour.

Eq. (2) is of the form

$$(3) \quad \ddot{\theta} - \varepsilon F(\theta, \dot{\theta}) + k\theta = 0,$$

where θ is the oscillation variable, ε is a small parameter related to the strength of the damping, 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. Over a long timescale, regular perturbation methods break down, yielding results with secular terms (terms that grow without bounds) that diverge from the observed behaviour over time. Nonetheless, the behaviour of pendulum over a long timescale *can* be derived to an arbitrary degree of accuracy using one of several singular perturbation techniques.

The Krylov-Bogoliubov-Mitropolsky (KBM) method, for instance, can be used to determine a solution of any differential equation of form (3) to an arbitrary degree of approximation. The method involves one substantive assumption about the solution: the amplitude and phase of the solution vary slowly, if at all, with respect to the period of oscillation. In other words, the solution is periodic and there are no secular terms. Thus, the KBM method assumes a solution of the form

$$(4) \quad \theta = a \cos \psi + \varepsilon u_1(a, \psi) + \varepsilon^2 u_2(a, \psi) + \dots,$$

where the u_i terms are periodic functions of ψ with period 2π , a is the amplitude of the first fundamental harmonic (the dominant oscillation) as a function of time, and ψ is the frequency of the first fundamental harmonic as a function of time. The time derivative of a and ψ can be written

$$(5) \quad \dot{a} = \varepsilon A_1(a) + \varepsilon^2 A_2(a) + \dots$$

and

$$(6) \quad \dot{\psi} = 1 + \varepsilon B_1(a) + \varepsilon^2 B_2(a) + \dots .$$

The functions u_i , A_i and B_i are chosen so that when the expressions for a and ψ obtained from integrating eqs. (5) and (6) are substituted into eq. (4), it becomes a solution of eq. (3). So the problem of finding a solution to eq. (3) reduces to the more tractable problem of solving eqs. (5) and (6). In practice, eqs. (4), (5) and (6) are limited to a finite number of terms, and in many applications, including the pendulum with quadratic damping, consideration of terms up to order ε^2 is sufficient. Surprisingly, the success of the KBM method does not depend on whether the infinite series eqs. (4), (5) and (6) converge, but rather on their asymptotic properties for a fixed finite order in the limit $\varepsilon \rightarrow 0$. The KBM method provides a general mathematical procedure to determine the functions u_i , A_i and B_i to any finite order for any $F(\theta, \dot{\theta})$ in eq. (3) because the finite analogues of the series (4), (5) and (6) are increasingly accurate approximations as $\varepsilon \rightarrow 0$ and as the number of terms in the series increase. For a table-top pendulum of the sort used by Galileo, the result is a measurable effect, one that is just slightly less than the effect due to linear air resistance and one comparable in magnitude to other known corrections such as the flexibility in the wire. (Nelson and Olsson 1986, 120).

As is typical of singular problems, in the quadratic damping case the idealization is non-Galilean (uncontrollable, harmful), and no Galilean idealization is available to underwrite the explanation. For the equation of motion (2) with small ε , a simple harmonic oscillator solution with fixed amplitude and period works well over a short timescales. Over long timescales, however, qualitatively new behaviours arise that need to be incorporated into the idealization, such as the increase in oscillation period and decrease in amplitude (going to zero). For any non-zero ε , the long-timescale behaviour remains qualitatively different from the $\varepsilon = 0$ behaviour. This well-known characteristic of singular perturbation problems is sometimes represented as a failure of a regular limit relation,

$$(7) \quad \lim_{\varepsilon \rightarrow 0} T_\varepsilon \neq T_{\varepsilon=0},$$

where T schematically stands for the theory in question (Nickles 1973). To put the point another way, there is no smooth deformation from the phase space trajectory of the simple harmonic oscillator to the phase space portrait of the quadratically damped oscillator.

The key point is that the quadratic oscillator model cannot be reverse-engineered or corrected in a theoretically principled way. Asymptotic methods, such as the KBM method, are not physically justified in the way that a regular limiting procedure is. For one thing, one may wonder about the mathematical propriety of throwing away terms in an infinite series that does not converge (the same sorts of worries one might have about renormalization techniques). More salient is that fact that the method includes ampliative steps, that is, steps in which additional substantive information is required beyond the basic equation of motion and initial conditions. These assumptions are *ad hoc*, in the sense that they are based on empirical observations about the behaviour of the system, not on any theoretical understanding contained in the equation of motion. In the quadratic damping case, we introduced the assumption of a periodic solution with certain properties in order to make it possible for us to derive such a solution. In sum, the idealization in the quadratic damping cases is non-Galilean because it is not fully corrigible in a theoretically justified way, and the explanans condition is violated.

If this is right, we have the remarkable result that one of McMullin's paradigm cases of Galilean idealization, that of Galileo's own pendulums, ends up not being Galilean at all. Recall that Galilean idealization includes the claim that the idealization can be refined, in principle, to an arbitrary degree while satisfying the Galilean conditions. This turns out not to

be the case with Galileo's pendulum models once anything more complex than linear air resistance is incorporated into the model.

4. Explanation without Successful Representation

Based on the foregoing considerations, one could easily suppose that something has gone horribly wrong in the explanatory practice of contemporary physics. Physicists put forward putative explanations that fail to meet even the most basic and reasonable criteria of normative philosophical accounts of explanation, including the extensions contained in the EvGI strategy. This supposition should be resisted! As we shall see, it is based on an assumption about the connection between explanation and successful representation, an assumption that should be questioned and ultimately rejected.

I have already reviewed the broad role non-Galilean idealization plays in contemporary physics, and the range of physicists' explanatory claims it underwrites. Proponents of the EvGI strategy will want to resist these as cases of *bona fide* scientific explanation, and there are two obvious ways to do so. One way is to show, on a case by case basis, that *each* explanation based on non-Galilean idealization should be rejected. One effective method to do this would be to demonstrate that there is a more basic or fundamental route to the same explanandum, one that makes use of the EvGI strategy, so that the non-Galilean explanation has only derivative status. This has been the response of some philosophers to Batterman's claims about explanations of criticality in statistical mechanics (Redhead 2004, Belot 2005, Batterman 2005). However, given the fact that in many cases even simple predictions require appeal to non-Galilean idealizations, it seems unlikely that further work on the details of Batterman's (and other) examples will demonstrate that explanatory appeal to non-Galilean idealizations can be eliminated in each and every case.

Another way proponents could defend the EvGI strategy would be to declare, wholesale, that physicists are simply mistaken in their claims to have achieved explanations of phenomena of interest. What would be needed here is a principled argument for the EvGI tenet that explanation as a normative goal of science can only be achieved in the context of Galilean idealization. Proponents of the EvGI strategy have not put forward such an argument. But they may not have felt the need to do so, because the EvGI strategy is founded on an intuition so basic and powerful it may not seem in need of articulation and defence. The problem with explanations in non-Galilean cases is that they are based on statements about a model that fail (and are known to fail) successfully to represent the explanatorily relevant features of the physical system. The EvGI strategy is motivated by an intuition about a deep connection between explanation and successful representation. This connection is broken in the non-Galilean cases, and this might seem enough for the wholesale rejection of non-Galilean explanation.

Let us briefly examine the connection between successful representation and explanation. Successful representation is commonly assumed to go hand-in-hand with successful prediction and successful explanation, and indeed that is true in much of science. As we have seen, proponents of the EvGI strategy are motivated by what they take to be a constitutive connection between a common-sense view about successful representation on the one hand and scientific explanation on the other. I think this is misguided. All we really have is a pervasive but contingent connection between the two, and one that does not seem to obtain in some fields of physics. Granted that in many contexts, both scientific and practical, predictions and explanations are better where statements in the prediction or explanans, making up essential parts of the prediction or explanation, are about features of the idealized model that approximate or are similar to features of the physical system. However, it is

widely appreciated, correctly I think, that judgments of similarity are made relative to a context and relative to a determinate set of practical ends (Goodman 1970; Teller 2001). If this is right, whether an idealized model counts as a successful representation, in the sense we are using the term here, depends on the use to which it is put. By contrast, the explanatory relation is understood by proponents of the EvGI strategy, again rightly, *not* to be a pragmatic matter, or at least not entirely; rather, it is determined by the basic explanatory goal of science. I hope to have drawn attention to the possibility that in some contexts the goal of explanation may not be best accomplished by what we have here been calling successful representation. If the physicists are even partially right, this is indeed the case with non-Galilean idealizations of apparently high explanatory value.

5. Conclusion

Explanations are put forward in a range of fields of physics in which idealized models doing putative explanatory work do not successfully represent. These are explanations that fail to meet the requirements of the EvGI strategy, and particularly its explanans condition.

Proponents of the EvGI strategy have not shown, nor is it likely that they can show, that each such putative explanation is underwritten by an explanation meeting the EvGI conditions.

This paper has argued that the EvGI assumption—that explanation as a normative goal of science can only be achieved in the context of Galilean idealization—should be rejected.

A reactionary option remains open, that of resisting *any* role for idealization in scientific explanation. But this option is unlikely to succeed; for one thing, the arguments developed against proponents of the EvGI approach (Section 4) tell equally strongly against such a move. Rather, the moral of this paper is that the role for explanatory idealizations should be expanded. Doing so would require more detailed work on the cases in physics in which these

putative explanations are developed. It would also require a normative account that makes sense of explanation in the context of non-Galilean idealization in physics, and indeed explanation without successful representation more generally.

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. (2005). "Critical Phenomena and Breaking Drops: Infinite Idealizations in Physics." Studies in History and Philosophy of Modern Physics **36B**(2): 225-244.
- Batterman, R. W. (2005). "Response to Belot's "Whose Devil? Which Details?"". Philosophy of Science **72**(1): 154-163.
- Batterman, R. W. (2009). "Idealization and Modeling." Synthese.
- Belot, G. (2005). "Whose Devil? Which Details?" Philosophy of Science **72**(1): 128-153.
- Bokulich, A. (2008). Beyond Reductionism and Pluralism: Reexamining the Relationship between Classical and Quantum Mechanics Cambridge.
- da Costa, N. C. A. and S. French (2003). Science and Partial Truth : A Unitary Approach to Models and Scientific Reasoning. New York, Oxford University Press.
- Elgin, M. and E. Sober (2002). "Cartwright on Explanation and Idealization." Erkenntnis **57**(3): 441-450.
- Galilei, G. (1638/1989). Two New Sciences, Including Centers of Gravity and Force of Percussion. Toronto, Wall & Thompson.
- Giere, R. (1988). Explaining Science: A Cognitive Approach. Chicago, IL, University of Chicago Press.
- Goodman, N. (1970). "Seven Strictures on Similarity." L. Foster and L. J. Swanson. Boston, University of Massachusetts Press.

- Hempel, C. G. (1965). "Aspects of Scientific Explanation." Aspects of scientific explanation. New York, The Free Press: 331-496.
- 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.
- 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.
- Nelson, R. and M. G. Olsson (1986). "The Pendulum - Rich Physics from a Simple System." American Journal of Physics **54**: 112-121.
- Nickles, T. (1973). "Two Concepts of Inter-Theoretic Reduction." Journal of Philosophy **70**(7): 181-201.
- Parker, W. S. (2006). "Understanding Pluralism in Climate Modeling." Foundations of Science **11**(4): 349-368.
- Redhead, M. (2004). "Asymptotic Reasoning." Studies in History and Philosophy of Modern Physics, vol. 35, pt. B, no. 3, pp: 527-530.
- 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.