Marc Lange on Minimal Model Explanations: A Defence of Batterman and Rice

Abstract

Marc Lange (2015) has recently raised three objections to the account of *minimal model explanations* and their import offered by Batterman and Rice (2014). In this paper, I suggest that these objections are misguided. In particular, I suggest that the objections raised by Lange stem from a misunderstanding of the what it is that minimal model explanations seek to explain. This misunderstanding, I argue, consists in Lange's seeing minimal model explanations as relating special types of models to *particular* target systems rather than, as Batterman and Rice urge, seeing minimal model explanations as looking to explain robust patterns of behaviour that are exhibited by a variety of physically diverse systems.

1 Introduction

Marc Lange (2015) has recently raised some objections to the account of *minimal model explanations* and their import offered by Bob Batterman and Collin Rice (2014, hereafter "B&R"). For B&R, a minimal model is one that, as Nigel Goldenfeld puts it, "most economically caricatures the essential physics" (Goldenfeld, 1992, 33). A minimal model explanation, then, is aimed at explaining what principled reasons we have for employing these minimal model caricatures in our investigations of disparate real systems. All in all, Lange puts forward three objections to B&R's approach: that it cannot account for the explanatory asymmetry that obtains between the minimal model and the target system, that it ultimately threatens to collapse into the kind of 'common features account' that B&R reject, and that B&R are wrong to suggest that an explanation appealing to a 'common feature' would have to explain the additional fact of the common feature's own prevalence.

These objections are, it seems to me, misguided. In particular, Lange's objections appear to stem from a misunderstanding of what it is that B&R suggest that minimal models explain. In phrasing his objections, Lange focuses on explanatory relationships and common features that obtain between the minimal model and the target system in question, whereas B&R are interested in minimal models explanations of patterns of behaviour that are exhibited by systems which differ in micro-scale physical detail. That is to say that Lange's objections appear to retain the idea that the aim of a scientific, model-based explanation is to stand in some explanatory relation to a *particular* target system, and it is precisely this idea that B&R seek to challenge on the basis of their examples of minimal model explanation.

In §2, I will sketch the approach to minimal model explanations proposed by B&R, briefly introducing the notion of a universality class that they employ and outlining their key example. In §3, I will outline each of Lange's three objections in turn. In §4, I will argue that Lange's first and third objections turn on a misunderstanding of what it is that B&R's minimal model explanations are seeking to explain. In §5, I will argue that Rice and Batterman's account does not collapse into a common features account as per Lange's second objection.

2 BATTERMAN AND RICE ON MINIMAL MODEL EXPLANATION

It is commonplace, B&R point out, for accounts of explanatory models to claim that a model is explanatory because it shares certain common features with the target system. Examples of these so-called "common features accounts" include those that require that the model accurately describe the causal mechanisms operating in the target system, those that require that the model accurately represent the relevant causal relationships or processes within the target system, and those that require that the model accurately specify the mapping relations that obtain between the model and the target system.

Overall, B&R claim that these accounts will respond similarly to two questions that they take to be importantly different. The two questions are:

1. What accuracy conditions are required for the model to explain?
2. In virtue of what does the model explain?

B&R suggest that common features accounts respond to the second of those questions "by citing that the accuracy conditions noted in response to the first are met" (Batterman and Rice, 2014, 351). However, B&R suggest that there are "classes of explanatory tasks in science for which the question "what makes a model explanatory?" is not properly answered by appeal to even minimal representational accuracy conditions" (Batterman and Rice, 2014, 356). These kinds of explanations employ *minimal models* and are focused on explaining patterns of behaviour that are exhibited by physically diverse systems. Used in this way, minimal models are explanatory "in virtue of there being a story about why large classes of features are irrelevant to the explanandum phenomenon" (Batterman and Rice, 2014, 356).

But what *is* a minimal model? On this point, B&R quote Nigel Goldenfeld, who contrasts approaches to modeling that involve including as many of the fine details of a system as possible with a view to faithful representation, and those that involve trying to subsume all of the microscopic physics into as few constants as possible. In other words, this second approach involves studying the patterns of behaviour exhibited across a range of systems, ignoring as much of the realistic, fine-grained detail as possible. In this, "it is only important to start with the correct **minimal model**, *i.e.* that model which most economically caricatures the essential physics" (Goldenfeld, 1992, 33).

Used in this way, the ability of such stripped back minimal models to reliably mimic the macrobehaviour of a diverse range of systems forms the target of what B&R call "minimal model explanations". In presenting a picture of how these explanations work, B&R use the language of *universality classes*. In essence, the picture is as follows. We consider a space of possible systems, and outline some kind of mathematical technique or transformation designed to eliminate irrelevant features and perform it on all of the systems in our abstract space. We are interested, then, in the fixed points of this transformations. Those systems that are taken by the transformation to the same fixed point are said to be in the same universality class – B&R say that the universality class is thereby "delimited." The mathematical delimitation of the universality class, B&R suggest, provides us with an explanation of why the systems that exhibit the same macro-level behaviour have the particular common features that they do. In other words, the mathematical techniques that we use to delimit our universality class show which factors are irrelevant to the patterns of behaviour that we see in diverse systems, and in doing so reveal which ones are the relevant ones. B&R thus conclude that "we can employ complete caricatures – minimal models that look nothing like the actual systems – in explanatory contexts because we have been able to demonstrate that these caricatures are in the relevant universality class" (Batterman and Rice, 2014, 364).

It is not, however, the *mere* fact that the minimal model and the other systems in the universality class share certain common features that accounts for the explanatory work such minimal models perform. Rather, the mathematical delimitation of the universality class demonstrates that "these minimal common features are relevant in the sense that they are necessary for the phenomenon of interest ... *to occur*" (Batterman and Rice, 2014, 361). B&R suggest that minimal model explanations so conceived are able to answer three questions that a mere common features approach to minimal models cannot. These are:

Q1. Why are these common features necessary for the phenomenon to occur?

Q2. Why are the remaining heterogeneous details (those left out of or misrepresented by the model) irrelevant for the occurrence of the phenomenon?

Q3. Why do very different systems have certain features in common?

It is precisely these three questions that we need to answer, B&R suggest, if we are to have an account of why minimal models are explanatory. The mathematical delimitation of the universality class allows us to answer these questions with reference to the fact that the minimal model is in the same universality class as the systems exhibiting the behaviour in which we are interested.

In illustrating their notion of minimal model explanations, B&R primarily refer to the following example. A range of behaviours in fluid mechanics are governed by the Navier-Stokes equations, along with some continuity equations. The quantities involved in the continuum Navier-Stokes equations make no reference to the small-scale details of the fluid they purport to describe. These equations are incredibly "safe" in that one can employ them successfully and reliably in engineering contexts of all kinds. The fact that this "safety" arises in conjunction with a virtual inattention to microscale detail is referred to by physicists as *universality*. It is "an expression of the fact that many different systems exhibit the same patterns of behaviour at much higher scales" (Batterman and Rice, 2014, 357).

We can exploit universality in order to develop models that capture the behaviour of the system at a macro scale. One such minimal model is the *Lattice Gas Automaton* (LGA).[1] We construct an updating algorithm based on the movement of particles on a hexagonal lattice, and if we perform some coarse-grained averaging we can yield macroscopic fields like number and momentum densities. The result accurately reproduces many of the macroscopic features of fluid flow. In particular, the LGA model accurately reproduces the

[1]Naturally, my discussion of this example parallels that of B&R in their original paper, which the reader may consult for a more extended treatment of the technical detail. A very accessible outline of the actual block-spin transformation can also be found in Goldenfeld and Kadanoff (1999).

parabolic profile of momentum density that is characteristic of incompressible laminar flow through a pipe (Kadanoff et al., 1989). The question is: given that it caricatures the physical structure of the fluids involved, why and how is this minimal model explanatory?

As outlined above, B&R suggest that such questions call for explanations in terms of the delimitation of a universality class. Fortunately, in the LGA case we are able to make this process precise and less metaphorical by employing a renormalization group strategy. Roughly, the idea is as follows. We induce on our space of systems a transformation called the Kadanoff block spin transformation, which involves replacing "a collection from within the lattice by an average or block spin that captures, in some way, the interaction among spins in the original system" (Batterman and Rice, 2014, 362). The result is that the transformation "takes the original system to a new (possibly nonactual) system/model in the space of systems that exhibits continuum scale behavior similar to the system one started with. This provides a (renormalization group) transformation on all systems in the abstract space" (Batterman and Rice, 2014, 362).

That is to say that the renormalization group transformation allows us to eliminate details or degrees of freedom that are irrelevant for the continuum scale behaviour in which we are interested. When we consider such a transformation as delimiting a universality class on our original space of possible and actual systems, we can see that it provides us with a mathematically precise way of accounting for the similarities in macrobehaviour that we see in the systems that are mapped to the same fixed point. That is to say that we are now in a position to answer Q2 above, since the renormalization group transformation has the effect of eliminating degrees of freedom from our input system that not relevant for the production of the macrobehaviour of interest. Examining the fixed points of the transformation (and thus delimiting the universality class) allows us to identify what features are shared by the systems that are in our universality class – in this case locality, conservation and symmetry. As B&R suggest, "we get an explanation of why these are the common features as a by-product of the mathematical delimitation of the universality class" (Batterman and Rice, 2014, 363). In other words, we are now in a position to answer Q1 and Q3 as well.

3 LANGE'S OBJECTIONS

Lange presents several objections to the approach advocated by B&R. To begin with, Lange presents what he calls an objection from *explanatory asymmetry*. The target system and the minimal model, Lange suggests, are simply two systems in the same universality class. Why is one explanatorily privileged over the other? Lange suggests that if shared membership in the universality class is what allows us to conclude that the minimal model explains the macrobehaviour of the target system, then this would also allow us to conclude that the target system explains the minimal model. This would be an unfortunate conclusion for B&R to license, Lange suggests, since "scientific practice does not include cases where the macrobehavior of some austere, minimal model is explained partly by the macrobehavior of some messy, real-world system" (Lange, 2015, 296-7). According to this objection, any account of the explanatory power of certain models that does not preserve the explanatory asymmetry displayed by explanations that employ scientific models does not properly capture the phenomenon at hand.

Lange's second objection is that the account offered by B&R itself collapses into a kind of common features account. Lange suggests that on B&R's account, "the responsibility for the target system's behavior would seemingly lie solely with the target system's membership in the given universality class" (Lange, 2015, 299). Yet is the fact that both the minimal model and the target system are brought to the same fixed point in the state space by the relevant transformation not a 'common feature'? Lange suggests in fact that since this property (of being brought to the same fixed point) is common to all members of the universality class, it constitutes precisely the kind of common feature to which B&R wish to deny explanatory importance.

Finally, Lange's raises an objection to the criticisms that B&R level at common features accounts. Considering the three questions that B&R insist that minimal model explanations must answer, Lange's simply asks: why must they? Of course one might be interested separately in answering these questions, but Lange insists that "a proposed explanation of fluid behavior by some symmetry is not rendered void or incomplete just because it provokes but fails to answer the question of why the symmetry obtains" (Lange, 2015, 304). In other words, minimal model explanations may well *provoke* these questions in us as a psychological matter, but this does not mean that they must settle these further questions in order to count as an explanation of the phenomenon in the fullest sense.

4 PATTERNS AS EXPLANANDA

Both Lange's first and third objection are the result of a misunderstanding of what it is that B&R suggest are the typical aims of minimal model explanations. In particular, Lange considers minimal model explanations as though they aim to provide explanations of *particular target systems* as opposed to robust patterns of behaviour. The objection from explanatory asymmetry that Lange puts forward asks why it is that if the minimal model explains the target system by virtue of shared membership in the universality class, the reverse is not true. Yet it is not the mere behaviour of the target system that is the explanandum in the case of minimal model explanations.[2] In particular, minimal model explanations are aimed at accounting for why we can use minimal model caricatures to investigate the behaviour of a variety of physically disparate real life systems. Naturally, once one has on one's hands a minimal model explanation one might use it in order to explain the behaviour of particular target systems, but minimal model explanations are, as conceived by B&R, aimed at explaining *patterns* of behaviour exhibited by physically diverse systems.

[2]Batterman himself mentions that there can be two types of why-questions that arise in these circumstances (Batterman, 2002, 26-27). When one asks of a particular event 'why did this happen?', what we mean may depend on whether we consider the phenomenon to be the causal details of the particular event or the general fact that, for instance, struts buckle upon reaching some critical load.

In other words, the explanatory target of minimal model explanations is in the first place *the ubiquity of the macrobehaviour*. If we find a way of rigorously delimiting the universality class associated with a particular kind of macrobehaviour, then the minimal model's membership in our universality class allows us to understand why real world systems behave the way they do. Why? Because we now know that the details that the minimal model elides (such that it strikes us initially as a caricature) are irrelevant for the production of macrobehaviour of interest, and so we know that if a particular system is a member of the delimited universality class, whatever differences obtain between it and the minimal model caricature are not relevant for the production of the behaviour they have in common. In this sense, if we are in a position to explain the ubiquity of the macrobehavior we will understand the relationship between our minimal model and *particular* target systems, but this is not what marks out a minimal model explanation. Minimal model explanations allow us to get a handle on patterns of behaviour that real life systems have in common with certain minimal caricatures, and in doing so of course we come to understand the real life systems involved in more detail. But there is no sense in which, at least as far as minimal model explanations are concerned, the minimal model can be considered to 'explain' the target system.

It does not then make sense to ask, as Lange does, "how does the minimal model's macrobehavior acquire explanatory priority over the target system's macrobehavior?" (Lange, 2015, 295) For B&R, the defining feature of a minimal model explanation is precisely that it provide us with a mathematical technique or process for delimiting the universality class associated with a particular kind of macrobehaviour. Minimal model explanations do not seek to place minimal models in particular kinds of explanatory relations (or give them explanatory privilege) over particular target systems. Rather, they provide us with a principled reason for grouping together physically disparate systems on the basis of the macro-scale behaviour that they exhibit. This in turn allows us to determine whether other systems (for instance, certain computationally convenient and conceptually enlightening minimal models) may also find themselves in such a grouping. The question of what picks the minimal model out as the explanans rather than the explanandum given that it and the target system are both members of the same universality class is ill-formed. In a minimal model explanation, the model itself is not the explanans, nor is any particular target system the explanandum.

That is to say that a successful minimal model explanation focuses not only on a *particular* target system but rather in explaining why it is that a class of systems exhibit the same behaviour at a certain scale. It is natural for the question of why systems of such varying physical configurations exhibit this behaviour to be provoked by the fact that a minimal model that at best caricatures the basic physical structure involved may also exhibit this same behaviour, and this question may in some contexts run alongside our interest in explaining the robustness of the patterns of behaviour that interest us. Insofar as we relate minimal models to the behaviour of target systems individually we often do so because we have principled reasons to see these two systems as exhibiting the same kinds of relevant behaviour, and one is for some contextual reason preferable to employ.

It is important to note here that *absent the minimal model explanation*, which provides us these principled reasons to treat these models as somehow relevantly similar, the minimal model does not 'explain' the target system's behaviour. Rather, the fact that a stripped down caricature such as the minimal model can be seen to exhibit the same behaviour is the catalyst for our seeking an explanation of the robustness of such macrobehaviour. We could, of course, *see* that the systems exhibit the same behaviour in some sense, but we would be left without an answer to the question with which B&R begin: why/how is the minimal model explanatory? Consider the LGA. We see that our caricature of the basic physical structures in terms of a hexagonal lattice and an updating algorithm exhibits the same kind of behavior as fluid moving past a plate. But how do we know that this is anything other than a helpful coincidence of mathematical detail? How do we know how far we may push these similarities in modeling things of engineering significance? It is the renormalisation group procedure (and the minimal model explanation it facilitates) that allows us to understand precisely why such caricatures may be relied upon to accurately capture the behaviour of real-world fluids at certain scales.

We may now see why it is that Lange's third objection is the result of the same misapprehension as his first. Lange asks us: why must an explanation of the target system's behaviour account for additional questions that such an explanation might provoke? The answer: because we are not in a position to *explain* the behaviour of the target system in terms of a minimal model unless we are in possession of answers to the broader questions cited by B&R. In the absence of the renormalization group transformation, we can perhaps (depending on the context and our levels of confidence) use our LGA caricature in order to investigate, predict or compute the behaviour of particular neural systems, but unless we are able to explain the robustness of the pattern we cannot use it to *explain* the behaviour of any particular system.

Putting the above differently, we might say that Lange assumes that *the minimal model itself* adequately addresses the explanatory questions involved in investigating the pattern of behaviour in which we are interested. The problem here is the thought that in some sense the minimal model can feature in a minimal model explanation (which involves delimiting the universality class associated with the macrobehaviour) simply by virtue of some explanatory relation that obtains between the minimal model and the target system. From this vantage point, it seems natural to conclude that the questions to which B&R refer are completely optional, imposed as they appear to be once the bulk of the explanatory work has been carried out. But as we have seen, the minimal model *itself* does not allow us to explain anything, as it were, in isolation. The questions at which B&R gesture must be answered *prior* to our being able to provide any kind of explanation of the patterns observed by appealing to the minimal model caricature.

5 UNIVERSALITY CLASSES AND COMMON FEATURES

Before returning to the bigger picture, there is one more objection to address. Lange suggests that the account offered by B&R in fact collapses into a common features account, albeit at a different level, since what what is of explanatory importance turns out to be the mere property that all of the systems share of being mapped to the same fixed point by a certain transformation. Is this not, Lange asks, a common features account in disguise?

Recall that B&R categorise common features accounts by the fact that they do not distinguish between the following two questions:

1. What accuracy conditions are required for the model to explain?
2. In virtue of what does the model explain?

It is perfectly plausible, for instance, that an account of model explanations may gesture to something that, in the right philosophical light, looks in some sense to be a ′common feature′ without necessarily being a ′common features account′ in the sense that it might treat these two questions differently. The question, for B&R, is not whether there exist any relationships between the models of interest that may plausibly be called 'common features', but rather whether whatever 'common features' we may observe are capable of providing answers to the questions that characterize minimal model explanations. That is to say: B&R do not claim that the model and the target systems will not have *any* features in common (and in fact we should expect that they probably will), but rather they deny that such common features perform the explanatory work in minimal model explanations.

Suppose then that we accept that B&R's universality classes provide us with a 'common feature' even in some flat-footed sense. It does not then follow that the answers to the above questions are going to be the same. In fact, for B&R they will not be. Our answer to the second question, in this case, would invoke the common feature that our systems within our universality class share, that of being mapped to the same fixed point by a mathematical abstraction procedure that provides us reason to group these particular systems together. Yet this common feature does not furnish us with the accuracy conditions that are required for the model to explain.

In other words, just because some aspects of a particular account of what makes particular kinds of models explanatory may be rephrased in terms of common features, it does not mean that the account itself is thereby a 'common features account' in B&R's sense. If we grant Lange's suggestion that the property of being mapped to the same fixed point by a particular transformation is sensibly a common feature of a variety of models, at best we may answer the second question above with reference to this common feature. We may even be able to do less than answer the second question, since such a property is explanatory precisely because we are able to tell a story about the mathematical abstraction procedure's ability to reduce our situation down to the relevant details. Consider the LGA case. Without a story at hand *either* about the renormalization group procedure's ability to pare down the variables in the right way, or more details and specified accuracy conditions regarding the kind of fluid behaviour the minimal model is expected to capture, the mere citing of the common feature of being mapped to a particular fixed point in our space of systems is not sufficient to answer either of the questions that B&R suggest that an account of minimal model explanations must address.

6 CONCLUSION

In the preceding sections, I have argued that in objecting to B&R's suggestions regarding minimal model explanations Lange misunderstands the thrust of their account of minimal model explanations. In particular, I have suggested that this misunderstanding consists in Lange's seeing minimal model explanations as relating special types of models to *particular* target systems rather than, as B&R urge, seeing minimal model explanations as looking to explain robust patterns of behaviour that are exhibited by a variety of physically diverse systems.

In a sense, then, Lange's response to B&R is instructive. Scientists routinely observe that behaviours on a certain scale are robust across systems of varying physical configuration, and as B&R argue it is precisely the pattern, rather than any one target system, that they are often concerned to explain. Lange's objections illustrate the way in which fixating on relations between models and individual target systems is not always a fruitful way to understand the role that models (in particular minimal models) play in our explanatory endeavours.

References

Robert W. Batterman. Asymptotics and the role of minimal models. *British Journal for the Philosophy of Science*, 53(1):21–38, 2002.

Robert W. Batterman and Collin C. Rice. Minimal Model Explanations. *Philosophy of Science*, 81(3):349–376, 2014.

Nigel Goldenfeld. *Lectures on Phase Transitions and the Renormalization Group*. Frontiers in physics. Addison-Wesley, Advanced Book Program, 1992.

Nigel Goldenfeld and Leo P. Kadanoff. Simple lessons from complexity. *Science*, 284(5411):87–89, 4 1999.

Leo P. Kadanoff, Guy R. McNamara, and Gianluigi Zanetti. From automata to fluid flow: Comparisons of simulation and theory. *Physical Review*, 40: 4527–4541, Oct 1989.

Marc Lange. On "Minimal Model Explanations": A Reply to Batterman and Rice. *Philosophy of Science*, 82(2):292–305, 2015.