Commentary for an Author-Meets-Critics Session on *Causal Reasoning in Physics* by Mathias Frisch sponsored by the Society for the Metaphysics of Science

Alyssa Ney

Pacific APA 2016

1. Introduction

Mathias has written a brilliant, masterful, and engaging book defending the legitimacy and importance of causal reasoning in physics. I learned a lot from reading it and recommend the book highly.

The first half of *Causal Reasoning in Physics* is mostly concerned with responding to arguments that although causal reasoning is legitimate, even indispensable in other branches of science, various features of physical theories make the use of causal reasoning problematic in physics. Mathias considers arguments claiming that causal reasoning requires coarse-graining, multiple realization, local causal relata, and a direction of time, all of which appear to be absent in fundamental physics. Mathias argues that these features which have seemed essential to causal reasoning are either present in physics as well or are not really required for causal reasoning. He also responds in these early chapters to arguments that the interventionist framework many use to interpret and assess causal claims may be inapplicable to the physical domain, particularly because of difficulties in applying the technical notion of an intervention in the context of physics. Mathias argues that there are various ways one may make use of the framework of interventionism in physics and that there is no problem applying the notion of an intervention in even our most fundamental and general physical theories.

Mathias goes on more of the offensive in the second half of the book, positively demonstrating the use of causal reasoning in actual physical frameworks, using examples from linear response theory and electrodynamics. There is an extremely interesting analysis of a debate between Walther Ritz and Albert Einstein over the source of asymmetries in accounts of blackbody radiation and their relation to the thermodynamic asymmetry. Mathias shows how the account of Ritz, in ways similar to what Mathias himself is defending in this book, that the radiation asymmetry is a fundamental causal asymmetry and not reducible to the thermodynamic asymmetry, was not, contrary to the common view, demonstrated false by Einstein, but instead appeared to be eventually adopted by him. And, in a part I will not discuss in these comments, Mathias also criticizes the rival frameworks of David Albert, Barry Loewer, and Huw Price, all of whom would see causal facts as arising from the kind of facts needed to ground thermodynamic asymmetries.

In these comments, I will focus on two issues. The first concerns Mathias's defense of the interventionist notion of causation in the context of physics. This chapter is an impressive response to especially Chris Hitchcock's "What Russell Got Right" and Jim Woodward's "Causation with a Human Face," both appearing in the 2007 *Russell's Republic* volume. Hitchcock and Woodward expressed serious doubts that one could find a foothold for causal reasoning in physics. Mathias critiques several of their arguments, however, I can't help but raise the question of why someone wanting to defend causal reasoning in physics should care so much about demonstrating the applicability of an interventionist theory in the context of fundamental physics. Interventionist models are certainly useful in special sciences like biology and economics where we fail to have anything like strict laws and the use of force concepts is tenuous, but in the context of fundamental physical theories like classical electrodynamics, it is not clear why causal reasoning cannot be directly legitimized in terms of facts about nomic structure and interactions. In what follows, I will show why Mathias has not quite convinced me that the notion of an intervention is applicable in fundamental physics. However, my main concern with this chapter is whether interventionism is even motivated for some branches of physics and why we should think the legitimacy of causal reasoning in a domain at all depends on the legitimacy of applying interventionism to that domain.

The second issue I will be focusing on in these comments concerns the kind of physical theories to which Mathias's arguments apply. Mathias is explicit from his introduction that he intends his conclusions to apply to all established physical theories. He says, "The arguments I want to consider here are meant to invoke features that apply quite broadly to any physical theory that is "on the books" (p. 23). However, quite often, it appears that Mathias's points only apply to nonfundamental physical theories, physical theories which may be mere approximations of the truly fundamental theories, or those branches of physics we can plausibly think of as special sciences because their claims do not hold of everything, everywhere, in all circumstances. I will focus on several of Mathias's key examples that would best make the case for causal reasoning in fundamental physics to draw out these concerns. However, I should emphasize that even if I am right that Mathias's arguments don't apply to truly fundamental physics, this doesn't undermine Mathias's achievement in this book in demonstrating the legitimacy and ubiquity of causal reasoning in physics today.

2. Interventionism and Physical Processes

Mathias responds to three concerns about the inapplicability of the interventionist framework in the context of physics. The core premise of the interventionist framework is that we may discern causal structure by considering the results of interventions in a causal network. For example (and roughly), to determine whether a system's having feature C makes a causal difference to another's having a feature E, we directly intervene on the instantiation of C and see what effects this has on the instantiation of E. What will be most relevant to us here is to note that in Woodward's interventionist framework, an intervention on C breaks the causal arrows between the feature C and its causes. So, for example, if the instantiation of C typically depends on some earlier feature of the system A, the intervention breaks this arrow of causal dependence of C upon A. The intervention directly fixes how the system is with respect to C regardless of whether A is instantiated or not.

The first critique of applying interventionism in the context of physics Mathias considers has to do with this requirement that interventions break causal arrows, that the intervention makes it the case that the system's having of feature C only depends on the intervention; it no longer depends on its having A. The concern is that if it is a consequence of the physical laws that the system's having C depends on its having A, then breaking this dependence would be a violation of the physical laws. And so the intervention would appear to be impossible. Turning to the second critique, several have raised the concern that for the notion of an intervention to be applicable, a system must have an outside from which one can intervene. But our most fundamental and general physical theories, those that contain models of the universe as a whole, do not have outsides. So there is no place for one to intervene from. So, again, interventions would appear to be impossible. And finally, third, what interventionism aims to do is discern causal structure in a system, that is discern paths of causal influence from a system's having one feature (or a constellation of features) to it or another system's having another feature. But there is a concern that since the physical laws relate only global world states, they do not have the resolution to capture the more localized paths of causal influence the interventionist aims to reveal. In what follows, I will argue that Mathias's defense of the claim that our physical theories have the resources to discern paths of causal influence is compelling, however his way of demonstrating this makes the interventionist framework superfluous. First, however, I have questions about Mathias's responses to the first two concerns, that interventions would be impossible in the context of fundamental physics.

Recall that the first concern was that interventions are processes that break causal arrows between the putative cause C and the prior features on which it is supposed to depend. We can see this clearly in the statement of interventionism Mathias quotes in Chapter 4, which comes from Jim Woodward:

> X is a total cause of Y iff under an intervention that changes the value of X (with no other intervention occurring) there is an associated change in the value of Y. (Woodward 2007, p. 73)
>
> I is an intervention variable on X with respect to Y iff I meets the following conditions:
>
> (1) I causes X.

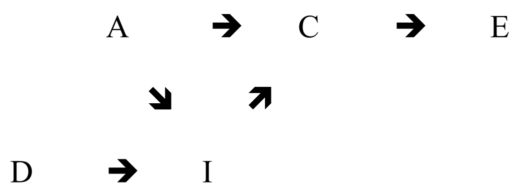
> (2) I acts as a switch for all the other variables that cause X. That is, certain values of I are such that when I attains those values, X ceases to depend upon the value of other variables that cause X and instead only depends on the value taken by I.
>
> (3) Any direct path from I to Y goes through X. That is, I does not directly cause Y and is not a cause of any causes of Y, if any, that are built into the I-X-Y connection itself; that is, except for (a) any causes of Y that are effects of X (i.e. variables that are causally between X and Y) and (b) any causes of Y that are between I and X and have no effect on Y independently of X.
>
> (4) I is independent of any variable Z that causes Y and is on a direct path from I to Y that does not go through X. (Woodward 2007, p. 75)

The requirement of arrow-breaking appears to follow from clause (2). In the context of the special sciences, the fact that interventions serve to break causal arrows is unproblematic since the laws of the special sciences (if there are laws in the special sciences) are widely recognized to not be strict; special sciences have exceptions. However, the laws of fundamental physics are supposed to be strict and so breaking causal arrows would amount to breaking physical laws.

Mathias considers several responses to this concern. One is that it only applies to cases of truly fundamental physical theories that are not supposed to have exceptions. If there is no such thing and all physical theories have exceptions, then there isn't a problem. Now Mathias expresses sympathy with this possibility throughout the book, but this is a contentious claim, one that many would argue is tantamount to rejecting the existence of physical laws, and so what is there to be said about this argument assuming there are such things as exceptionless physical laws? Mathias offers a couple of proposals.

First, Mathias argues there is a way to model interventions in physical systems that involves the satisfaction of (2), but does not involve arrow breaking. To show, this he provides an example involving a feedback mechanism. Suppose that the way someone will intervene on C does not merely depend on the intervener's decision (D) how she wants to intervene, but also on the earlier state of the system (whether or how it instantiates feature A). Then the system can be modeled in the following way:

<u>Figure 1</u>



In this case, the result of the intervention will be to fix the fact about the system's instantiation of C, but this doesn't render it independent of its earlier having A, since the nature of the intervention itself depends on its having A. So this is supposed to illustrate how interventions can work (and satisfy Woodward's clause (2)) without breaking causal arrows.

I was a bit confused about what Mathias had in mind with this proposal.[1] This is supposed to be a case of an intervention that renders C independent of A without thereby breaking causal arrows between A and C. But I am not clear what justification there is for saying the causal arrow between A and C is not broken. One might try to argue that since C depends on I and I depends on A, C depends on A. But these are supposed to be arrows of direct causal dependence. So we would need some other motivation for saying in this case the intervention doesn't break the causal arrow between A and C. If we can talk

[1] I'm setting aside the discussion of stability here because I think getting into that would be too complicated given the time constraints. The concern I am raising is not supposed to be an a priori insistence that causal models have to be stable.

about physical forces and that the introduction of I doesn't remove the existence of these forces, then we can do that, but this seems to rely on prior facts about physical influence and relying on these would seem to make the interventionist framework superfluous. So I am just not sure there is motivation for the kind of case Mathias describes.

An alternative Mathias proposes is that perhaps interventions do not really require the satisfaction of clause (2); that is, we may allow both soft and hard (or surgical) interventions to reveal causal structure, where soft interventions need not be switches; one can intervene while allowing the feature C to still depend on its causal parents (A). Mathias uses a nudging of a cue ball as an example of a soft intervention. The result of this nudge depends not only on its features (the direction and angle of the stick) but also where the cue ball started out on the table when it was nudged. So the nudge is an intervention that not only does not break causal arrows, but also fails to break the dependence of C on its parents. This is a contentious suggestion whether such soft interventions can play the role required of interventions in causal modeling. Mathias argues that the motivation Woodward has provided for clause (2) only justifies a conception according to which an intervention changes the system's instantiation of C, not one that renders C independent of its parents. But there are other reasons to be cautious of replacing the account of intervention above with one requiring only soft interventions. For example, Frederick Eberhardt in a recent paper (2014) notes that it is questionable whether it would be a good idea to use only the notion of soft intervention in an account of causal discovery because it can often be difficult to know when one has carried out a soft intervention successfully. This is due to the fact that there will most of the time be unobserved causal factors that are (justifiably) left out of a causal model.

When we use surgical interventions, we can be sure that a given intervention breaks the causal connection between C and all of its causes, including those that are unobserved and may play the role of confounders. On the other hand, if we are using a soft intervention, one that may or may not break the arrows between C and its causes, we cannot be sure that dependencies on other observed causes are broken. This can prevent us from knowing whether or not the intervention has actually been carried out in the way intended. This isn't to say that soft interventions may not often be useful and reveal causal information, but it isn't obvious that we should simply replace Woodward's account of total cause with one that makes use of a notion of soft intervention.

I now wish to turn to the second difficulty for applying the notion of an intervention in the context of fundamental physics, namely, that interventions seem as if they must come from the outside of the system in question, but our most general fundamental physical theories do not have outsides. In response to this concern, Mathias argues that it would be better to articulate the notion of an intervention in such a way that it does not require something coming from the outside. He discusses the notion of atomic intervention from Judea Pearl. This is formally modeled by the replacement of one structural equation with another that has the variable in question set to a particular value. I wasn't entirely clear what Mathias was proposing here. At first, I thought the suggestion was that we view interventions not as physical processes that have to come from somewhere (in or outside of the system), but instead as mathematical operations that involve a variable being stipulated to take on a particular value. In this case, the "intervention" isn't an action on a system at all, it's just a change in representation. We replace one causal model that has certain patterns of causal dependence with another with a different pattern of dependence. But I don't think this is what Mathias is proposing and anyway it wouldn't solve the problem, which is how the original pattern of dependencies could fail to obtain given they provide an exhaustive and accurate model of physical systems. Assuming we had a complete physical model, there is no way even to consider an alternative yet accurate physical model with additional arrows. What the text seemed to suggest though was that Mathias is thinking Pearl's notion of an atomic intervention gives one a way of viewing interventions as coming from the inside. I don't see how the notion of atomic intervention specifies anything about the "where" of an intervention though. In any event, if we do think of interventions as occurring "inside" systems, then this will return us to the concerns about arrow-breaking described previously.

Turning finally to the third problem for using interventionism to discern causal structure in physics, recall the issue was that physical laws seemed not to allow us to distinguish localized paths of causal influence since they only relate global world states. Yet the interventionist framework aims for finer resolution of causal structure. In response, Mathias argues that physical theories do have the ability to distinguish finer causal routes. Here, following Sheldon Smith, he points to representations in terms of causal Green's functions, which allow us to represent a physical state in terms of a sum of different disturbances that are its causes. This is a useful thing to note for the purpose of finding causal structure in physics, but it again seemed to raised the question: if the dynamical framework of physics already provides what we need to discern causal influences, then why care about saving interventionism in the context of physics? It's the interpretation of the Green's function that is telling us what causes what, not any facts about the counterfactual results of interventions.

So my first concern is that I am not sure Mathias has really shown that the notion of an intervention is viable in the context of fundamental physical theories. And it looks like the physical laws (and whatever principles are needed to distinguish causal from anti-causal Green's functions, which I haven't talked about yet) themselves provide the resources to ground causal facts without any need for the interventionist's theory.

To further develop this point, consider one of the central examples Mathias uses in the book to motivate the indispensability of causal reasoning in physics, an example illustrating the radiation asymmetry. This is John Earman's example of a radio tower broadcasting a signal into the far reaches of space.

In such a case we would observe radiation fields diverging from their source. And yet, although the fundamental equations governing this process are time-symmetric, we don't observe in nature the reverse process, of radio waves coming from far reaches of space and converging on a tower. Why is that? The reason, Mathias argues, is that there is a causal asymmetry in the fundamental processes: radiation diverges away from sources into distant regions of space, but does not converge from distant regions of space onto an object. Mathias notes that one may argue that this causal asymmetry is not fundamental, but may in turn be grounded in a more fundamental principle about probabilistic independence. If we assume radiation coming from distant regions of space is independent ($P(X_1 \wedge X_2 \wedge \ldots \wedge_n) = P(X_1) \wedge P(X_2) \wedge \ldots \wedge P(X_n)$), but radiation coming from a source is not, then assuming the causal Markov condition, the causal asymmetry will follow. Mathias is officially neutral about whether it is the probabilistic independence assumption or the asymmetry itself that is fundamental. He expresses sympathy with a comment of Richard Feynman's that depending on our interests, we may take one or another principle as fundamental and there is nothing more to say.

My point in bringing this case up is to note that I cannot see what interventionism adds that would allow us to discern the asymmetric causal structure in this case. The laws, boundary conditions, and causal assumptions (the asymmetry or independence assumption) allow us to discern the causal structure. If we want more fine-grained information, we may look at the causal Green's functions. Intervention counterfactuals seem beside the point.

2. Fundamental vs. Nonfundamental Physics

I now want to turn to the second topic, about whether Mathias's arguments really succeed in showing us something about the use of causal notions in fundamental physics as opposed to in nonfundamental physical theories. Mathias argues that even in the case where we possess fully general physical laws, causal reasoning is still indispensable.

He uses an example involving the points of light we see in the night sky to respond to an argument from Huw Price and Brad Weslake that "the use of causal notions cannot be justified within the context of the laws and is a scientifically unmotivated add-on." The question is how can we be sure that a point of light we see in the night sky comes from a single source, a star, rather than is source-free radiation streaming from infinity. (I just want to add as a parenthetical comment that Mathias also uses another example, of the representation of proton collisions at the LHC, to make a case for the use of causal reasoning in fundamental physics. I will not discuss this example since it is quite complicated, but will note that what makes it complicated is the fact that our representation of these collisions (Mathias argues) draws on a patchwork of assumptions and principles from various levels of description. So, I don't find this a compelling example of causal reasoning in our most fundamental physics, as opposed to causal reasoning in a patchwork of fundamental and nonfundamental physics.) Mathias argues that to solve this problem using only laws and boundary conditions would require solving a final value problem, one that would require having far more information than we actually possess. Yet, we have no problem concluding that the light comes from a star. How? The answer is that we use causal reasoning. We note that light with similar characteristics appears in a similar position in the sky at nearby times, and at other relevantly related positions at other times and from different locations on Earth. We then use the principle of common cause to infer that the light we see in these different situations is coming from a common source, a star. To rule out the hypothesis that the light we see in the night sky is source-free radiation, we start by noting the reasonable assumption that fluctuations in fields in the remote past of the universe are random. We then infer from this that source-free radiation in space is probabilistically independent, which makes it extremely unlikely for it to exhibit the coordinated behavior necessary to explain our observations. Thus, Mathias concludes, causal reasoning is far from redundant in physics, it is indispensable. The fact that this example is so mundane suggests how frequently we must rely upon causal assumptions (the principle of common cause and facts about probabilistic independence) to generate predictions in fundamental physics.

Here is where one might complain that Mathias has not really shown the indispensability of causal reasoning in fundamental physics. Instead what we are seeing is the indispensability of causal reasoning where our information about fundamental physics is limited. If we were able to access the complete state of the universe at the current time, then we would be able to solve the final value problem and would have no use for causal reasoning in determining whether what we see is a star. If we had access to the full physics, it appears we would not need the causal assumptions in such a case.

But Mathias rejects that causal reasoning is essential only in cases in which we lack relevant bits of knowledge. To see this, we may return to the radio antenna example. In this case, recall, we wanted to know why we see radiation diverging from sources to far reaches of space but not radiation converging from far reaches onto a localized object. In this case, we may assume we have complete information about boundary conditions. Still, Mathias argues, without causal assumptions, we lack the ability to answer the question. To answer it, we need to assume a causal asymmetry that radiation diverges from sources, but does not converge onto them, an asymmetry we have seen may perhaps be grounded in the probabilistic independence of radiation in the past but not the future. This is a key example for Mathias that reappears throughout the book. But I do not believe that even this case shows that causal reasoning is indispensable in fundamental physics. (Although I should say I take no issue with the fact that it shows the indispensability of causal reasoning in nonfundamental physical theories. But I take it that is a less controversial and so less interesting conclusion.)

The reason I am not convinced is that I don't see what reason we have to say that fundamental physics itself fails to permit signals coming from spatially distant locations to converge onto a localized object. This is after all something that is perfectly allowed by the laws. Now, Earman says this sort of event would be "miraculous" which suggests it is an event that if it occurred could not have an explanation. And Mathias wants to agree with this verdict, by claiming that it is ruled out by the correct causal modeling of radiation processes. But given that we are assuming full knowledge of boundary conditions and the laws, we would possess a deductive-nomological explanation of the unusual convergence. Thus I would argue that if it is part of a physical theory that such processes cannot take place, then this won't be a fundamental physical theory, but rather a special science devoted not to the study of all electromagnetic processes that may occur in nature, but rather those that are likely to occur given what we ordinarily observe.

Mathias considers something like this objection in Chapter 6. To defend the claim that the reverse process is physically impossible (as a matter of fundamental physics), he suggests that we need not think that what is physically impossible is only what is incompatible with the laws. One can say that what is physically impossible is what is incompatible with the laws and additional assumptions (e.g. the causal asymmetry he is arguing for). But I don't see that this case provides motivation for this reinterpretation of physical possibility. And so I'd like to hear more.

Reference

Eberhardt, Frederick. 2014. Direct Causes and the Trouble with Soft Interventions. *Erkenntnis*. 79: 755-777.