A HUMEAN ACCOUNT OF LAWS & CAUSATION

PhD Thesis in
History & Philosophy of Science

Toby Friend
Science & Technology Studies
University College London

August, 2017

I, Toby Friend, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

....

January 16, 2018

Abstract

The thesis proposes a new account of laws of nature and token causation within the Humean tradition. After a brief introduction in §1, I specify and defend in §2 a Humean approach to the question of laws and causation. In §3 I defend the view that laws are conditional generalisations which concern 'systems' and detail further issues concerning the scope, content and universality of laws. On the basis of the discussion concerning laws' logical form, I argue in §4 against a view of laws as mirroring the structure of causal relations. Moreover, I show how this conception is implicit in the best system account of laws, thereby giving us reason to reject that account too. §5 presents an alternative 'causal-junctions conception' of laws in terms of four causal features often associated with laws: component-level and law-level dispositionality, and variable-level and law-level causal asymmetry. These causal features combine to demarcate a central class of laws called 'robust causal junction laws' from which other laws can be accounted for. §6 provides a Humean analysis of the causal features used to characterise robust causal junction laws. This is done first by providing an analysis of dispositions in terms of systems and laws, and second, by providing an analysis of causal asymmetry in terms of relations of probabilistic independence. In §7, I then provide a nomological analysis of token causation by showing how the causal junctions described by robust causal junction laws can be chained together in a particular context.

Acknowledgements

This thesis is the culmination of work dating back to long before I began my PhD. For the inspiration to pursue the subject of laws and causation I must first thank the regular attendees of the *London Causality Reading Group*, in particular: Adam White, John Pemberton and Phyllis Illari. Our sometimes-regular meetings from 2009 introduced me to an enormous amount of literature and a friendly but keenly philosophical environment for discussion. I should also thank Robert Northcott for his excellent supervision for my MPhilStud thesis, a precursor to my PhD. Thanks must also go to Emma Tobin for encouraging me to apply to UCL after examining said thesis and for helping me secure funding.

My PhD research has been generously funded by the AHRC without which this work certainly would not have been possible. I am also grateful to the AHRC for a further Study Visit and Conference Fund award which allowed me to have an extended stay in Helsinki to attend the summer school on Metaphysics of Science organised by Tuomas Tahko and present at the CLMPS conference 2015. Substantial financial assistance has also been repeatedly provided to me from the Science and Technology Studies department enabling me to attend and present at multiple conferences.

Since I've been at UCL, I've had the opportunity to make contact with many philosophers and research groups which have been enormously influential. Special thanks must go to Anna Marmodoro and all those on the *Powers Structuralism in Ancient Ontologies* and *Metaphysics of Entanglement* projects at Oxford for allowing me to be involved with the research, reading groups and presentations. Also I'd like to thank Andreas Hüttemann and those on the *Causation, Laws, Dispositions and Explanation* project in Cologne for assisting me to come and visit on numerous occasions. Of those connected with this latter project I should thank in particular my good friend Florian Fischer who also invited me to his brilliant Summer School in Bonn and has let me sleep on his couch many times. Further thanks to Andreas in Cologne and Daniel Kodaj at Oxford for their very helpful comments on an early draft of my paper *Laws are conditionals*, much of which is repeated below in §3.1.

My department in UCL has been a fantastic place to work for the past four years, made so predominantly by the awesome PhD cohort and our graduate tutor Chiara Ambrosio. Special thanks go to the philosophers who have helped me shape my ideas on a number of occasion and particularly to Rory Jubber for providing me with comments on an early version of §2.3.

Luckily, I was able to draw on the expertise of three fantastic supervisors: Phyllis Illari, Emma Tobin and Luke Glynn. The care and interest they've shown in my work has been incredible and they've all gone way beyond the call of duty for the time they've provided me with—I'll never forget it. Extra special thanks goes to my primary supervisor, Phyllis, who has been helping me develop my ideas on all this causation stuff for about eight years now! I've been honoured to be Phyllis's first primary supervisee at UCL and sure hope I can do her proud.

Above all, I can say without hesitation that the person most invested and supportive during all this research has been my beautiful wife Leila. Leila has listened to me garble through inchoate versions of this thesis for uncountable hours and has always been unnervingly adept at providing constructive criticism, inspiration and encouragement. I couldn't have done any of this without her, let alone stayed happy and healthy during it.

A final thanks is also extended to my daughter, to whom I'm grateful for kindly avoiding coming prematurely while I've been completing this thesis. Having said that, I also can't wait to meet her.

Contents

List of Tables

List of key definitions and schemas

1 Introduction

In comparison with the concept of 'causation', the concept of a 'law of nature' is a relatively new one. Although precursors to it can be found in works dating back at least to the ancient Greeks (Giere 1995), its coinage is generally attributed to Descartes, picked up soon after by Newton in his works on Optics (Zilsel 1942, Needham 1951). In contrast, causation has been a perennial topic of enquiry for natural philosophers dating back to the earliest proto-scientific endeavours (Hulswit 2002, ch.1).

Despite being relatively new to the conceptual repertoire of natural philosophers, the discovery of a law of nature is widely considered one of the most highly prized scientific achievements. Consider the following from Feynman:

> From a long view of the history of mankind—seen from, say, ten thousand years from now—there can be little doubt that the most significant event of the 19th century will be judged as Maxwell's discovery of the laws of electrodynamics. (2011, §6)

What, then, is the relationship between causation and laws of nature? Here, things become controversial. Some have thought the relationship to be very close indeed. Helmholtz, for example, claimed that,

> Our demand to understand natural phenomena, that is, to discover

> their laws, is a different way of expressing the demand that we are to search for the forces that are the causes of the phenomena. The lawfulness of nature is conceived of as causal relationship. (1896, 40, translation in Frisch 2014, 1)

Helmholtz seems here to be drawing a very close relationship between laws and causation. To change the emphasis a little: understanding the world is a process of discovering laws of nature, knowledge of which is given to us through the search for causal relationships. Others have been less sympathetic to such a close association. For instance, some have believed that causation is not the right concept to use in current-day science, and physics in particular (Kirchhoff 1876, Mach 1900, 1905, Russell 1912, Hempel 1965, Norton 2009, Price and Weslake 2009; see Frisch 2014 for a discussion); while others have believed it is rather *laws* which are of limited value Beatty (1980), Giere (1995), Mumford (1998), Machamer et al. (2000), Woodward and Hitchcock (2003).

I, for one, am on Helmholtz's side. More specifically, I believe that laws of nature are conceived of in terms of certain specific type-level causal relations. This result may seem to disappoint those of a Humean-empiricist bent who so desperately seek to rid science from worrying metaphysical notions like causation (e.g. Hume 1738 [1978], 1777 [1993], Russell 1912, Carnap 1928, Hempel 1965), but I will argue that it need not. Laws may be *conceived* of in causal terms, but we can still remain believers in a non-causal *analysis* of both laws and causation. In fact, I will use the observations about the causal conception of laws to inform a number of issues which have been raised among those influenced by Hume concerning how best to conceive of both phenomena. The result is an account of laws in terms of specific and independently analysable type-level causal relations which can then be put to use in a competitive and comprehensive analysis of token causal relations.

The analysis I have in mind here is *metaphysical*. It need not accurately capture aspects of the epistemology or conceptual role of causation, but it must provide conditions which are at least extensionally adequate.[1] Although I will refrain from using the term, many Humeans would consider such an analysis a 'reduction' of causation to something more basic or fundamental.

This process of reasoning is at odds with many standard approaches to questions of causation and laws. The contemporary philosophical literature acknowledges a deep divide between Humeans and non-Humeans. Traditionally, Humean theorists start with a limited set of non-causal facts from which the laws are derived. With the laws in place, a metaphysical analysis of causation is then sought typically drawing on laws in some manner. Regularity analyses (e.g. Hume 1777 [1993], Mackie 1980, Grasshoff and May 2001, Strevens 2007, Baumgartner 2008, 2013a), counterfactual dependency analyses (e.g. Lewis 1973a, Hitchcock 2001b, Halpern and Pearl 2001), probabilistic analyses (e.g. Reichenbach 1956, Suppes 1970, Lewis 1986a, Menzies 1989, Kvart 2004, Fenton-Glynn 2011, 2016) and some conserved quantity theories (e.g. Salmon 1977, 1994, 1998a, Dowe 2000) all seem to be 'off the shelf' theories of causation readily available for Humeans. The only problem—a *big* problem—is picking one which is extensionally adequate.

In contrast, the non-Humean theorists often start with the assumption that the former approach is hopeless. Causation is simply too basic to be analysed in the way the Humean wants. Instead, non-Humeans begin with a set of explicitly causal facts and derive an openly circular analysis of causation (or else provide no analysis at all (see Anscombe 1971). Interventionist analyses (e.g. Woodward 2003), certain mechanist analyses (e.g. Machamer et al. 2000, Bechtel and Abrahamsen 2005, Craver 2007, Illari and Williamson 2013), dispositionalists and powers-based analyses (e.g. Mumford 1998, Molnar 2003, Heil 2005, Bird 2010) are all developed from this starting assumption (although strength of belief in it varies). Only with such analyses

[1]Presumably *mere* extensional adequacy is insufficient. The conditions must be informative, general and exhibit some degree of modal robustness.

in place—and depending on their fondness for the concept—do non-Humean theorists consider developing an account laws too (e.g. compare Bird's 2007a tolerance of laws with Mumford's 2004 rejection of them).

In such terms, this division is not sufficiently refined to capture all the available views. Indeed, the view presented below falls through the cracks. Going by tradition, then epistemically my sympathies lie with the non-Humean. I think we have good reason to believe causal relations are basic to our perceptual experience of the world (see Michotte 1963, Anscombe 1971, Menzies 1998, Carey 2009) and I also suspect that causation is *conceptually* basic too (as have, for various other reasons Tooley 1987, Kant 1998, Edgington 2011). Moreover, I think that an adequate account of laws should pay attention to this fact. Ultimately, laws are generalisations we deem important because they capture particularly useful and salient causal information. However, metaphysically, my sympathies lie with the Humean, since I think there are also good reasons for expecting causation to admit of metaphysical analysis in terms of non-causal facts. A crucial point I want to capitalise on is that siding with the idea that causation is conceptually and/or epistemically basic does not preclude a Humean metaphysical analysis of it (simplified to 'analysis' from hereon). More significantly, if we embrace the epistemic and conceptual priority of causation we open ourselves to a more plausible and penetrating Humean account of both laws and causation.

But things are still more nuanced than that. As indicated, it is typical among Humeans to treat both type-level and token-level causation as metaphysically downstream of the laws (ignoring so-called 'causal laws' Cartwright 1979). The Humean typically relies on an independent analysis of laws (even if only implicitly) in order to develop an account of both token-level causation (e.g. Lewis 1973a, 2000, Salmon 1977, 1998c,a, Dowe 2000), and type-level causation (e.g. Reichenbach 1956) or both (e.g. Suppes 1970, Baumgartner 2008).

My Humean account of laws and causation will not work quite like this. In

§5 I will propose four type-level causal features which I think can be put to use in accounting for a certain class of laws, what I call the *'robust causal junction laws'*. Other laws are then to be accounted for derivatively. But where causation is invoked in order to ascend to an account of laws, it will then be thrown away (like the ladder spoken of in Wittgenstein 1961, §6.5.4), in order that a non-causal account of it in terms of laws can then be provided. This is carried out in §6 where I will provide a Humean analysis of these features. Crucially, these analyses will not themselves draw on the laws in their own metaphysical analyses. Once laws have been accounted for in this way, I then use them to provide a nomological analysis of *token* causation in §7. Consequently, the new Humean picture orders the direction of analysis rather differently than before.

This constructive work only emerges from §5 onward. Before that I intend to make some methodological and critical remarks and lay a significant amount of groundwork concerning laws' logical form. In §2, I defend Humeanism as a methodological position interpreted as an abstinance of reference to three kinds of 'necessary connection'. I then move on to §3, which discusses laws' logical form. Laws, I will argue, have the form of a conditional with an antecedent predicate clause satisfied by instances of a system-type; hence all laws are 'system-laws'. In §4, I explore and criticise what I take to be a dominant conception of laws as mirroring the structural form of causal relationships. Showing why this is problematic will also allow me to comment on and ultimately reject the 'best system account' of laws which dominates much of the discussion on Humean approaches to laws.

The positive spin on the relationship between causation and laws comes in §5 when I consider more carefully how laws are causally conceived offering two pairs of causal features I believe to be inferentially tied to laws: two of 'dispositionality' and two of 'causal asymmetry'. With these causal features in mind I am then able to offer my own distinct proposal for a 'causal-junctions conception' of laws defined in terms of *robust causal* *junction laws*, a particular class of law which exhibit all of the described features. In §6 I then set about providing an analysis of the causal features drawn on in the new account in Humean terms. §6.1 argues for a purely extensional conditional analysis of laws' dispositionality and §6.2 argues for a probabilistic analysis of laws' causal asymmetries. In combination, these analyses support a *Causal-Junctions Account* of laws presented in §6.3. With a fully Humean account of laws in hand, I then show in §7 how we might (metaphysically) analyse all token causation using the idea that some causation occurs within systems described by robust causal junction laws (i.e. is 'intra-system causation') all other causation is a chaining of the former kind (i.e. is 'inter-system causation') achieved via some observations about event-identity. Since the former is a limiting case of the latter, the result is an account of causation which interprets all causation as 'inter-system causation'. Hence, the thesis will have provided both a Humean account of laws as system-laws and of token causation as inter-system causation. §8, provides a brief summary.

Figure 1.1 (p.19) represents the direction of reasoning that this thesis aims to establish from §5 onwards. The arrows indicate the direction of analysis and the numbers indicate the sections in which the respective analysis or account is developed. The result of the proposed accounts of laws and token causation is a more nuanced Humean position which avoids the brash rejection of what I take to be salient and crucial connections between the two. By adopting this new position I hope the Humean can respect many of the intuitions which have motivated rejection of it in the past.

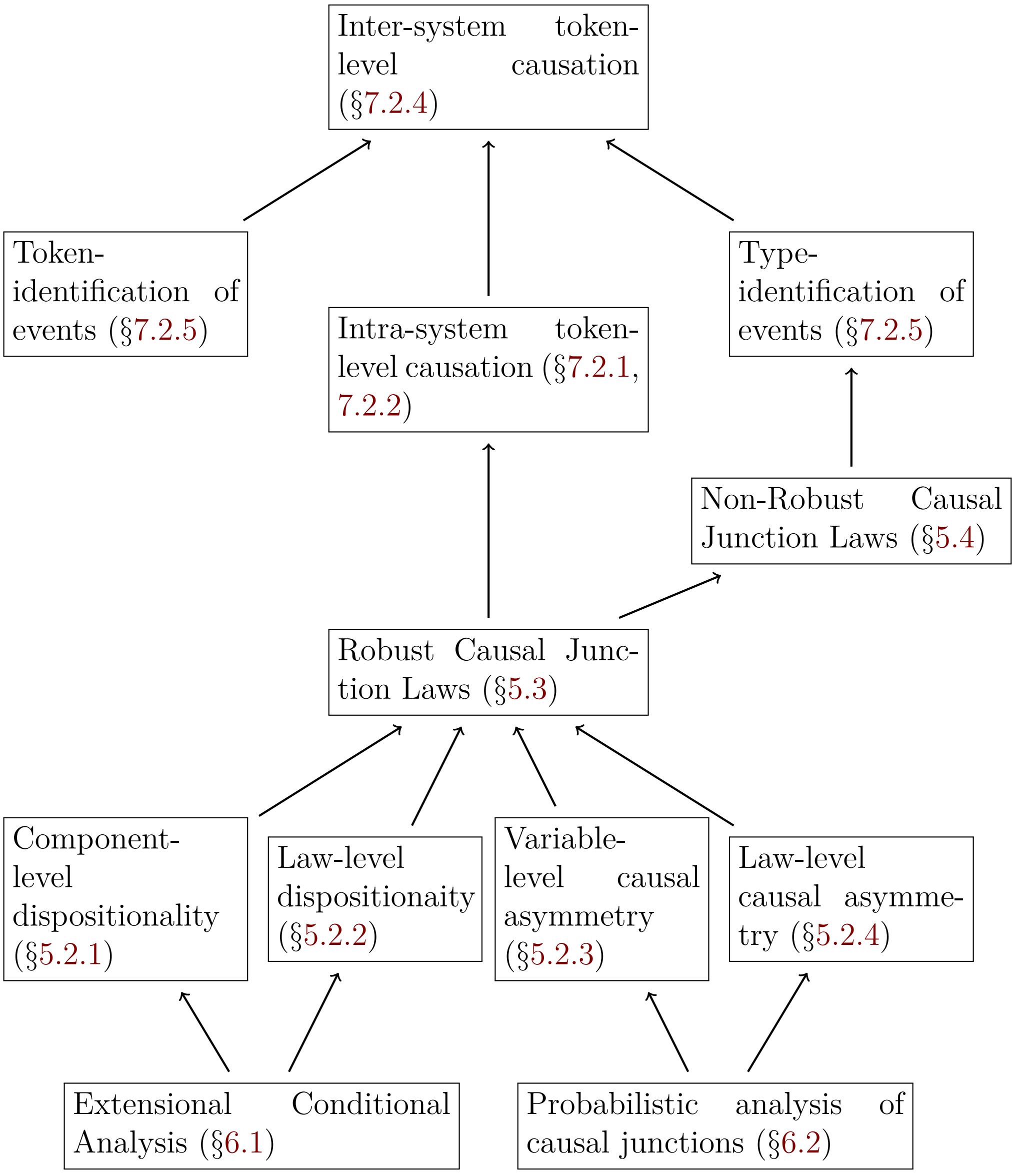

Figure 1.1: Thesis structure from §5 onwards

2 Humean methodology

In contemporary terms, I am no committed Humean. It seems to me that there are a number of arguments which should at least remain open as to whether the sort of necessary connections eschewed by a thoroughgoing Humean really exist. Nonetheless, when it comes to laws and causation, I do not see any benefit to introducing unHumean posits. In this chapter I will argue for a form of methodological Humeanism. This is a methodological approach to philosophising which restricts the available materials which can form the source or explanation of certain phenomena—in our case, laws and token causation—to the contents of the so-called 'Humean mosaic' (defined below).

According to Hume (1738 [1978], 1777 [1993]), he could observe neither the exercise of powers, nor necessary relations. In general, Hume thought that,

> all events seem entirely loose and separate. One event follows another; but we never can observe any tye between them. They seem conjoined, but never connected. And as we can have no idea of any thing, which never appeared to our outward sense or inward sentiment, the necessary conclusion seems to be, that we have no idea of connexion or power at all. (1777 [1993], ch.7, part II)

For Hume, discussion of causation and necessary connections were conflated into a singular path of reasoning: we can't perceive a worldly instance of the relation and so we can't conceive clearly of one, so we're better off not assuming one. Beebee has suggested that Hume was 'so sure of the connection between experimental reasoning [i.e. concerning, e.g., our inferential habits] and causation that it does not even occur to him that experience might reveal events to be connected in a way that does not enable one to infer the second from the first' (2006b, 80; see also Mackie 1980). But whatever the reason for Hume's lack of demarcation, it is clearly one we can nowadays make. For instance, one might coherently accept primitive causal connections but reject primitive modal connections or vice versa.

Nor should we stop here in demarcating types of connection Hume was dissatisfied with. For nowadays, sympathisers with Hume take it as incumbent to defend a position opposed to primitivism about causal connections (e.g. Anscombe 1971, Edgington 2011), subjunctive conditional connections (e.g. Lange 2009), higher-order connections between universals (e.g. Armstrong 1983, Bird 2007a, Swoyer 1982, Tooley 1987) and lawlike connections (e.g. Carroll 1994, Maudlin 2007).

Of course, it is common to think that these various connections are closely related. So the saying goes, laws support counterfactuals (a kind of subjunctive conditional), causation is often justified by drawing attention to corresponding counterfactual dependencies, and higher-order connections are often considered in discussion of laws. Nonetheless, it seems logically coherent to consider the denial or acceptance of each of these independently of any other. One could, for instance, be an error theorist about any of the connections whilst defending a nontrivial account of the others.

Even when taking the Humean position to be one which denies all these connections as primitive, the influence of Hume's sceptical attitude has been pervasive. For example, his distaste for metaphysical posits which go beyond the empirically accessible was a characteristic feature of the entire positivist-empiricist movement (e.g. Mach 1900, 1905, Carnap 1928, Hempel 1965). The positivist programme, perhaps best exemplified in Carnap (1928), was an attempt to build up an account for all rationally justifiable theory in terms of a basic set of constituent claims which made reference only to those aspects of reality which could be directly perceived (or, later, were the basic constituents of a rational physics). For the positivists, some aspects of theory would be necessary, by virtue of their analyticity. However, the necessity that comes with analyticity arguably does not conflict with the general Humean denial of primitive causal, subjunctive, higher-order or lawlike connections. Indeed these were all connections the positivists denied. Either they could be defined in terms of more epistemically reasonable connections (e.g. constant conjunction) or they were to be rejected from theory.

The positivists were motivated by epistemological concerns, but in recent times Humeanism as a philosophical stance has taken on a more metaphysical flavour, most notably in the works of (Lewis 1986g, 1994). Lewis admitted that many of the papers which contribute to his *Philosophical papers*,

> fall into place within a prolonged campaign on behalf of the thesis I call "Humean supervenience." [...] Humean supervenience is named in honor of the greater denier of necessary connections. It is the doctrine that all there is to the world is a vast mosaic of local matters of particular fact, just one little thing and then another (1986e, ix).

It is fair to treat the 'necessary connections' Lewis denied as including all the various disambiguated types of connection Hume was concerned with. This lead Lewis to attempt to provide Humean treatments of causation Lewis (1973a), of subjunctives Lewis (1986b), and laws (Lewis 1973b). Lewis also expressed a distaste for higher-order connections (see Lewis 1983) and was at pains to show how the 'work' such connections could do might be done drawing only on the resources of the mosaic.

Seeing the world as a Humean mosaic led Lewis to believe not only that there were no primitive necessary connections, but that there were no strictly necessary relations between spatiotemporally distinct entities at all. For Lewis, the world's facts could be *recombined* in any possible combination (provided there is enough spacetime). More precisely, 'the principle is that anything can coexist with anything else, at least provided they occupy distinct spatiotemporal positions' (Lewis 1986f, 88).

This *Principle of Recombination* is one of the central principles of Lewis's Humeanism and reflects Hume's own commentary on necessity. For Hume reasoned that ''tis an established maxim in metaphysics that whatever the mind clearly conceives includes the idea of possible existence, or in other words, nothing we imagine is absolutely impossible' (1738 [1978], Book I, part II, section II), where, 'in order to know whether any objects, which are joined in impression, be separable in idea, we need only reconsider if they be different from each other; in which case it is plain they may be conceived apart" (Book I, part II, section III). From such extant conceivability of separation Hume surmised that 'there is no object which implies the existence of any other' (Book I, part II, section IV).

However, at least since Kripke (1980), we know better than to assume that conceivability can be used conclusively to justify reasoning about possibility. Lewis accordingly avoided grounding his view on such evidence. We might expect Lewis to replace the original evidence for Humeanism with something else, but as MacBride notes, 'it is a curious fact that the proponents of the contemporary Humean programme—Lewis included—having abandoned the empiricist theory of thought that underwrites Hume's rejection of necessary connections provide precious little by way of motivation for the view' (2005, 127). In fact one might wonder what reasons there could be for denying the existence of the sorts of necessary connections which would prevent an unrestricted principle of recombination (see, e.g. Wilson 2010).

It's worth pointing out that the principle of recombination is not *essential* to a view in the spirit of Hume which denies primitive causal, subjunctive or higher-order connections. One might, for example, reject the primitive connections and yet also reject recombination because one holds an error theory of modal talk. Or one might think that certain necessary mathematical facts which are neither worldly nor analytic must be considered when modalising about laws (see McKenzie 2013). We must, therefore, separate the presumption of recombination from the desire to restrict the available solutions to philosophical problems of law, causation and counterfactuals to those which base their accounts solely on the mosaic.

In order to avoid falsely aligning directly with any existing philosophical tradition or giving a misleading interpretation of Hume, I will define my own approach in terms of the content of the mosaic directly. The 'Humean methodology' is the method of limiting oneself to reference to the materials in the mosaic in order to account for certain phenomenon of interest—in the case of the current text, *laws* and *causation*. Anyone following this methodology I will call a 'Humean'. As it turns out, Lewis and other supporters of Humean supervenience are Humean in this sense, and I will draw heavily on their discussions of laws. But crucially, there are other ways to be Humean which may be preferable in light of certain issues.

If we are Humean, what are we precluding from the mosaic? Based on the disambiguation I provided above, we have at least the following three distinct kinds of necessary connections to consider.

1. Primitive causal connections: a connection between particular causes and their particular effects.
2. Primitive subjunctive connections: a connection of modal dependence, e.g. between states of affairs or events.
3. Primitive higher-order connections: a connection between entities which can have instances (e.g. universals).

We could add to this list a fourth primitive connection of 'lawhood'. Primitive laws have been advocated on occasion as an exasperation with the progress of the philosophical debate over laws (see Carroll 1994, Maudlin 2007). I have no principled arguments that such primitive connections are unreasonable. My only reason for denying them is that I believe there is an account available which doesn't make reference to them (see §5 and §6).

Each of the three connections listed above have been posited for distinct reasons and as far as I'm concerned each should be denied from the mosaic when laws are under scrutiny. In the following sections I will explain why I think this is justifiable. Three points are worth mentioning beforehand.

First, what is denied in each case is not the existence of the connections altogether. What is to be denied is rather the presence of these connections in the mosaic, which is to form the raw materials (quite literally) with which laws and causation are to be accounted for. Essentially, the Humean at least wants to avoid taking such connections as a *primitive* in their analysis.

Second, it is typical for Humeans to consider the mosaic to exhaust the fundamental, either because it is conceptually prior (as the positivist-empiricists thought) or because it metaphysically grounds everything else (as Lewis seemed to think). However, I do not take this to be a requisite for being a methodological Humean. Although I will not argue for it here, there seems to me the logical space to deny the fundamentality claim and still believe that building an account of laws and causation which are limited to the mosaic is methodologically advisable. What I aim to show in §2.1, §2.2 and §2.3 is that there are good reasons to avoid referring to these connections when trying to provide an account of laws and causation—this is what I mean by not taking such connections as primitive. One might think these arguments are sufficient to infer the lack of fundamentality of such primitives. Nonetheless, the space seems there to defend the more nuanced claim that there exist such fundamental connections despite Humean analysis being available.[1]

Third, the methodology has been defined in terms of reference to a limited class of phenomena, namely, those which appear in the mosaic. Of course, I've made reference many times already to causal relations and I will be discussing them further along with the other connections which don't appear in the mosaic. Does this mean I'm not adhering to the methodology? No. The point is not to avoid reference to such connections altogether, but only in *the final analysis*. This would render such connections primitive. My final account of laws and their type-level causal features is provided in §6 and of token causation in §7.

2.1 Against primitive causal connections

There are good reasons to adopt a methodology which avoids reference to primitive causal connections over and above the relata in causal relations. Although saying what they are is a delicate issue. Traditionally, this Hume-inspired attitude is motivated by the belief that we don't observe causation, rendering it *at best* a theoretical posit. But this is controversial. In the introduction I admitted that my sympathies concerning the epistemic priority of causation in perception lay with the non-Humean: I think we observe causation all around us and that it plays an important role in how we conceptualise the world. If one has sympathies with Anscombe (1971), for instance, we might believe that one can observe the cat is lapping up the milk (understood as a causal fact) as well as anything else in the perceivable world. Moreover, citing experiments by Michotte (1963), Leslie and Keeble (1987), Schlottman and Shanks (1992), Peter Menzies (1998) concludes that,

> to the extent that there is a prevailing orthodoxy among cognitive

[1] One might, for example, suspect causation must be fundamental in order to have any hope of an informative solution to Putnam's model-theoretic arguments against a non-trivial interpretation of our language (see Putnam 1977, 1980). Even so, one might also think—along Humean lines—that an analysis of causation is possible in which causal talk can be removed from the analysans (hence not primitive). Despite the controversy which would surround such a view, it seems to me logically coherent.

> psychologists, it is that there is a modular, informationally encapsulated system that is responsible for the perception of causal relations in special circumstances. (1998, 354)

But notice that to deny that causal connections are observed is not the same as denying that we can come to know of the existence of a causal relationship via direct perception. One can agree, for example, with Anscombe (1971, 137) that someone can see that the cat drank up the milk or that the dog made a funny noise or that things were cut or broken by whatever we saw cut or break them. But this does not commit one to claiming that they see the causal *tie*—a thing in and of itself—between the cat's drinking and the milk's being drunk up, or the dog's movement and the noise, etc. As Armstrong (1968) pointed out, seeing that some fact Fa is the case, need not imply seeing the particular a. Similarly, we might think that seeing that there is a causal relation does not imply seeing the causal relation as a distinct entity.

Moreover, one can consistently distinguish *perception* (whether veridical or not) interpreted as impacting on perceivers' inferential states, from *observation* interpreted as the process whereby information about the world is transferred (via distribution of photons across the retina or airwaves on the eardrum, chemical reactions in the olfactory receptors, etc.) to the brain. This latter interpretation of observation may be philosophically limited when it comes to making sense of how we form perceptual beliefs, but it is nonetheless helpful to make sense of the limitations on the physical process which mediates our mind and the world. We know, for instance, that certain inert chemicals will have no observable odour because they do not react with our olfactory glands, and that objects in a vacuum make no observable sound because they cannot transmit sound waves. When it comes to the causal relation, it seems highly plausible that Hume's claim, interpreted as one about observability in this sense, remains correct. For there appears to be nothing more than the cause and the effect which are suitable candidates for impacting on our sense-organs in a relevant way. No extra tie between them will be observable. If this is correct, then while there may be a sense in which we perceive causal relations, we can still agree that there is no observable *connection* between causes and effects.

In defense of this last point, it's worth considering some of the ties that have been mooted to exist between causes and effects. On the one hand we have those ties suggested to be metaphysically motivated, such as Tooley's (1987) primitive causal relation C or Armstrong's (1983, 1997, 2004a) nomic necessitation relation N. On the other hand, we have those ties suggested to be physically motivated, such as Aronson (1971) and Fair's (1979) transfer of energy/momentum. Even if one countenances these connections, the question remains whether they are able to impact on our sense-organs independently of the relata which they connect. I, for one, cannot bring myself to entertain the idea that a transfer of energy or momentum or a universal of nomic necessitation is observable in this sense.

It seems to me, therefore, that Hume-inspired scepticism concerning the observation of causal connections can be maintained in conjunction with some fairly liberal attitudes towards our epistemic access to causation. But a final point is that even if causal connections *were* observable (or acceptable for some other reason) there is still practical benefit in a methodology which prohibits making use of reference to them. This is because the existence of a primitive causal tie between causes and their effects is neither sufficient nor necessary for making *use* of causation.

Notice that if causal relations weren't able to be conceptualised under generalisation then we would be severely limited in what use we could make of them. Even if a causal tie in each token case were observable, this would limit us to being able to tell when some causal event had happened. We would have no ability to foresee causation or engineer anything to behave predictably. It is only once we allow our causal knowledge to fall under explicit generalisation that we are able to do these things, by providing us with the knowledge that instances of a general consequent property will accompany instances of a general antecedent property. The causal tie is, therefore, insufficient to render causation a practically useful aspect of reality.

For a related reason, a causal tie is also unnecessary to make use of causation. When we intervene on the world to change it, we want to change what events are taking place (or what facts are true). It is irrelevant to such an intervention whether or not our world is one in which the interventions which are performed are ones mediated by some real primitive causal tie. All we need to know is that the intervention will result in the desired outcome. But this is something which could feasibly take place in a Humean world; if our world were really just the mosaic void of primitive causal connections it could nonetheless be the case that the same actions are predictably accompanied by the same results as any of the unHumean worlds dreamt of by those in favour of primitive causation.

The Humean methodologist sees the insufficiency and lack of necessity of causal ties in the practical utility of causation as further motivation to look for ways to analyse causation in non-causal terms, i.e. in the terms of the mosaic. Ultimately, I will suggest this can be done when we have a Humean account of laws.

2.2 Against primitive subjunctive connections

Another connection the Humean denies from the mosaic are primitive subjunctive connections, specifically, modal relations of dependency which ground conditionals of the form 'if it were the case that [...] then it would have been the case that', such as counterfactuals. Many have pointed out that laws of nature are counterfactually stable in a way that accidental generalisations are not (e.g. Hempel 1965, Lewis 1973a, Armstrong 1983, Carroll 1994). More specifically, a criterion of 'nomic preservation' (NP) is often proposed (see Carroll 1994, Goodman 1947, 1954 [1983] for some early versions; Lewis 1973a, from p.72 considers and rejects counterfactual stability in this form). The specific version of this which I am happy to endorse owes a lot to the work of Lange (2009) and Roberts (2008) and can be presented as follows (where '$\succ$' indicates the subjunctive conditional).

> **(NP)** A proposition l is a law only if it is a member of the largest nonmaximal set Λ of sub-nomic propositions such that all nested counterfactuals $q \succ l, r \succ (q \succ l), ...,$ for antecedents $q, r, ...$ logically consistent with Λ, remain true in any scientific context.

I say 'I endorse' NP, but that is not to say that I think we should take counterfactual conditionals (a kind of subjunctive) as primitive. For it does not seem to be the subjunctive nature of laws which is doing the explanatory work of telling us *why* the propositions among Λ count as a law. What follows in this section is a discussion of why this Humean view is reasonable.

First some clarifications. In NP, 'nonmaximal set of sub-nomic propositions' means that the set doesn't contain every sub-nomic proposition there is, and a proposition is 'sub-nomic' so long as it makes no claims of lawhood ('no inertial body moves faster than the speed of light' is sub-nomic, 'that no inertial body moves faster than the speed of light is a law' is not sub-nomic'; see Lange 2009). That laws satisfy NP is motivated firstly from the idea that it is only under a counterfactual supposition which conflicts with some law or other that it would be correct to infer that some (possibly distinct) law is no longer true in that context. The world can change in myriad ways and it still be true that copper is conductive. Only if we counterfactually suppose that copper is *not* conductive, or that electrons are neutrally charged, or that some other of a related group of laws is false would it be true to infer that copper is not conductive. The laws are the set of all such groupings, hence the *largest* nonmaximal set.

NP requires not only that laws remain true under consistent counterfactual suppositions, but also that they remain *counterfactually stable* under any number of nestings of consistent suppositions. Not only is it the case that had I worn different socks today then copper would still have been conductive, but also that had I eaten a different breakfast then had I worn different socks then copper would have been conductive. The idea behind this extra criterion is that laws should remain *laws* under consistent suppositions rather than merely true. To be explanatory, one might think, a proposition should be counterfactually robust. Hence, even under counterfactual suppositions, it must be the case that laws are true under counterfactual suppositions, and so on.

Although I accept the forgoing point about laws' stability, I doubt (and suspect Lange also doubts) that there is any proposition which is a member of the largest nonmaximal set of propositions which remains true under counterfactual suppositions consistent with the set *but fails to remain true under further counterfactual nestings of suppositions consistent with the set.* If NP is suitable test for lawhood at all, I will assume it suffices to simply check whether a proposition remains true under the relevant suppositions rather than entertaining infinite nestings of suppositions as well.

Lange (1993b, 2009) has argued that there is no context in which laws do not remain true under consistent counterfactual suppositions. But a series of examples from Roberts (2008) should, I think, persuade us otherwise. Roberts points out that in cases where there are some particularly salient 'extra-scientific truths' in a conversational context, then we should not always expect the laws to remain true in all logically consistent counterfactual suppositions. For example, assume that we are in a world in which there is a god who created the universe and that we are concerned with the god's decision to make ours a world which supports life. In such a context, we are likely to endorse a counterfactual conditional like, 'had the world been $500^o K$ hotter, there would still be intelligent life', since we would assume that God would have made the laws differently to how they are in the context in order for life to remain supported.

If Roberts' reasoning is correct, we should not expect preservation across modal space to capture a feature of laws which holds in every context. Roberts proposes instead that we limit the contexts in which laws are nomically preserved to the 'scientific contexts', i.e. those in which 'the point of the conversation is to exposit, apply, or expand empirical-scientific knowledge' (p.198). I will assume the Roberts strategy here is the right one (see Roberts 2008, ch.7 for more detail on what a scientific context is).

Another feature of NP is the claim that laws form the largest *nonmaximal* set. The idea here is that the set that contains every proposition whatsoever, viz. the maximal set, will trivially be one whose members remain true under counterfactual suppositions logically consistent with it. Any counterfactual supposition will be inconsistent with some member of the maximal set so any counterfactual supposition which isn't inconsistent (of which there are none) will be one in which the members of the maximal set remain true.

Lange has suggested, and I am happy to agree, that what distinguishes the set containing laws is that it is the only *nonmaximal* set which exhibits this characteristic. Take the proposition 'all the pears in this basket are sweet', abbreviated 'p'. Let us assume p is true (i.e. there is a basket being referred to and the pears in it are all sweet). So p is actually true, but nonetheless it is not true under any counterfactual supposition logically consistent with it. E.g. had I placed an unripe pear in the basket, p would not be true. Nor would it clearly remain true under any supposition $\neg q \vee \neg p$, for some arbitrary true proposition q, e.g. 'the number of this page is 33'. Although consistent with p's truth, there are presumably some scientific contexts under which the supposition of $\neg q \vee \neg p$ would lead us to infer p's falsity, e.g. in a context where I am poised over the basket with an unripe pear. But since q was arbitrary, for p to be a member of that set whose members remain true under logically consistent suppositions, the set will have to also contain all true propositions whatsoever and hence will be maximal. Laws don't work like this. For any arbitrary true proposition q and law l we know that within scientific contexts, the supposition that $\neg q \vee \neg l$ is going to allow us to infer $\neg q$ unless $\neg q$ itself is inconsistent with some law. The laws eke out an intermediate level of necessity between pure contingency and metaphysical necessity by virtue of being the only nonmaximal set of propositions which remain true under (nontrivial) counterfactual suppositions consistent with the set.

I for one am happy to agree with Lange that NP defines a necessary condition for being a law, viz. being a member of Λ.[2] But what I don't agree with, and this is where the Humean claim appears, is that the subjunctive truths which place this condition on laws actually *explain* why laws are special. For the supporter of the Humean methodology, such subjunctives are something the laws explain, not vice versa. To make that Humean attitude plausible it is of prime importance to show how a Humean methodology can determine the laws. This is the task of §6. However, this might seem an unproductive exercise if one is entrenched within the view that NP does all the explanatory work one could desire for laws. In the rest of this section, I want to undermine this view by pointing to both to its insufficiency and irrelevance in making sense of laws.

Notice that the way NP is expressed, it says only what laws must be like, not what must be laws. This is because there are further categories of proposition which fall into Λ and yet do not contain only laws. One such category is that of the 'broadly logical truths', including logical laws, e.g. the law of excluded middle; purely mathematical truths, e.g. '1+1=2'; definitions, e.g. 'all vixens are female foxes'; and identities between the referents of logical constants, e.g. 'Eric Blair is George Orwell'. All broadly logical truths are metaphysically necessary, i.e. they are propositions which remain true under any supposition

[2] Notice that NP is inconsistent with the typical Humean view of laws popularised by Lewis (1986b) in which laws are typically broken by small 'miracles' under counterfactual supposition, even if the supposition is logically consistent with every law. This divergence from the norm has repercussions for views about backtracking, etc., which I cannot explore here. But see (Roberts 2008, especially ch.6) for a useful discussion.

whatsoever. Because of this, they are counted among the laws according to NP. However, I assume it is implausible that such examples are laws of nature.

In order to get a sufficient condition for laws, it might be thought that we can stipulate that the laws are the *contingent* propositions of *Lambda*. However, as tidy as this response is it faces some challenges.

One challenge emerges with a common form of argument that *all* laws are broadly logical truths. Despite their differences, Armstrong (2004b), Bird (2007a), Sellars (1948), Shoemaker (1997) and Swoyer (1982) all motivate their belief in the strongest necessity for laws on the grounds that the very identity of properties depends on it. In brief, since categorical properties are implausible (because they imply 'quiddities' see, e.g. Black 2000) properties must be non-categorical, i.e. get their identities through relationships with other properties (in the case of Bird, Shoemaker and Swoyer) or particulars (in the case of Armstrong and Sellars). But since the relationships concern identity, they must be metaphysically necessary.

In defence of laws' contingency, the complaints with categorical properties which motivate these metaphysical positions can definitely be challenged (see Schaffer 2005b, Roberts 2008). And the argument from property-identity to laws being broadly logical truths is also dubious given the apparent contingent status of physical constants (Hendry and Rowbottom 2009, Keinänen 2011). Moreover, some have pointed out that certain features of theorising in scientific practice lends credence to the view that laws are contingent, as for example when different force-laws or Newtonian-worlds are considered (Lange 2004, Hendry and Rowbottom 2009; see also §2.4-§2.5 Roberts 2008).

Still, a different reason one might worry about the condition of contingency is if one had specific examples of propositions which were plausibly both laws and broadly logical. After all, there only needs to be one law in Λ which isn't contingent to render the added condition on laws' contingency false. For example, one such argument put forward by Bird (2001) aims to show that salt's ability to dissolve in water is both lawlike and metaphysically necessary by virtue of the natures of the two materials (although a series of responses may have blunted Bird's original argument; see Beebee 2002, Psillos 2002, Bird 2002). A more historically motivated argument for particular noncontingent laws comes from (Hanson 1958, ch.5) who pointed to a number of places where the total force law (i.e. Newton's second law of motion) has been treated and even explicitly referred to as a definition in scientific practice.[3]

One might argue that the definitional status of a proposition ultimately undermines its status as a law. Take the example of Atwood's machine, referenced by Hanson, in which two masses are suspended by either end of the same massless thread hanging over a frictionless pulley. A well-known textbook of Hanson's time claims 'Newton's second law is a definition and hence incapable of proof [...] the Atwood machine is essentially a device for measuring the acceleration of gravity g by determiniation of [acceleration] a rather than a set-up for the verification of Newton's second law' (Kolin 1950, 46-7). By treating the total force law as a definition, force is eliminated from any essential role in making sense of empirical enquiry. This might lead us to wonder whether it is really appropriately called a law in these circumstances. Perhaps this is why the authors Hanson cites who support a definitional reading draw various caveats over its being titled a law. Poincaré said that 'the [second] law of motion [...] ceases to be regarded as an experimental law, it is now only a definition' (1905, 100) and Humphreys and Beringer (1950, 38-9) claim that 'Newton's laws are not physical laws [...] but are definitions of the basic concepts of physics'.[4]

[3]Admittedly, Hanson also points towards a number of other uses of the total force law which are not definitional. In various contexts the law may not be a definition but a prerequisite for scientific thought, a core component of a particular world-view or a summary of a large body of experience (pp.99-100). In none of these latter cases is the law as necessary as the broadly logical truths. Nonetheless, if there are definitional uses of laws, it seems plausible that some law could exclusively have such a use, or at least, exclusively have those uses in scientific contexts.

[4]The reality of forces is still today a controversial issue (see Creary 1981, Cartwright 1983, Ward 2009). I will, however, continue to assume the realist position. If this turns out to be untenable then I suppose my use of force laws as examples in what follows may ultimately be poor choices, but the general points should hold nonetheless. The key point

I don't think we should be so quick to reject some laws being definitions. Whatever we think of the total force law, a different example which seems to be definitionally true in any context is the Hardy-Weinberg equilibrium law.

> **Hardy-Weinberg equilibrium:** For any x, if x is a population of reproducing organisms, the allele and genotype frequencies will remain constant in the absence of other evolutionary pressures.

This law was originally proved mathematically and so is arguably true by virtue of the meanings of the terms and (an extension of) logic alone. It's not clear to me, however, that the relationships described by the law could be removed from reasoning so easily as force seems to be (see also the discussion of the symmetry principles in McKenzie 2013).

Regardless of the outcome of this debtate, it's important to notice that the discussion of whether or not definitions can be laws does not seem to concern any modal feature, but rather something about the status of our knowledge of them. One might argue, as Poincarè did, that definitions are not empirical or experimental enough to count as a law since they are too easily removed from reasoning. But the point has little to do with the modal profile of the propositions themselves. This suggests that what explains the status of definitions as laws is not a fact about necessity but rather about the knowledge we have of how to manipulate them in reasoning. Insofar as we explain why definitions are or are not laws, there does not seem to be any direct need to call upon subjunctive facts.

Another category whose propositions fall within Λ but which are plausibly not all laws is that of laws' logical entailments. Since NP is deductively closed, any proposition p is true under any counterfactual supposition in which a proposition which entails p is true. Hence, NP entails that any proposition entailed by a law is in Λ. If it is a law (and so in Λ) that copper is conductive,

is whether or not we should treat definitional readings of laws as compatible with those in which the proposition's status as a law remains.

then the proposition that any copper statue of a cat made by a sculptor with hazel eyes is conductive on a Tuesday is in Λ as well. Similarly, if it is a law that copper is conductive, then the proposition that copper is conductive or it is not conductive will be in Λ. Although controversial, I take it that our pre-theoretical inclination will be to consider neither of these consequences to be plausible candidates for laws. For one thing, unlike the laws from which they are derived, these consequences do not seem to provide suitable explanations of, say, the conductivity of a copper statue of a cat by a hazel-eyed sculptor on a Tuesday (see Roberts 2008, §2.4 and Lange 2009, 16 for a discussion of this issue).

We should not let the long shadow cast by certain approaches to laws (e.g. Ramsey 1950 (2^{nd} edition, Lewis 1973a) incline us to accept that laws are closed under deductive entailment in this unintuitive way. As far as I know, this feature of laws has never been argued for directly but only as a consequence of accounts which entail it being proven otherwise plausible. In fact, given that there is a strong inclination to deny deductive closure from direct considerations, I am tempted to take this as a reductio on such accounts.

In order to weed out unwanted propositions from the class of laws, we could hope to stipulate that laws should not be overly specific or too disjunctive (perhaps according to Gricean conversational maxims of informativeness). Another option is to require that laws only cover relationships between natural kinds (e.g. Harré 1993). Further still, one might suggest that it is the relationships themselves which are natural (e.g. Armstrong 1983, 1997). Finally, there is the option I will myself be suggesting in §5, that laws are demarcated by the particular causal information they are associated with. But whatever our choice at this stage, whether it be a constraint on logical form, naturalness or something causal, notice again that it does not seem to be the subjunctive nature of laws which is doing the explanatory work of telling us which of the propositions among Λ counts as a law.

Lange's attempt to determine the laws with subjunctives is limited by the fact that many non-laws appear to have the same subjunctive features that laws do. In terms of NP, once we have Λ, it appears we still require some further method of individuation to separate the laws from the non-laws. But even when we consider responses to this apparent insufficiency, it does not seem to be subjunctive facts which play a role in deliberation over lawhood. This should, I think, lead us to suspect that the counterfactual stability which gives laws their subjunctive flavour is derivative of some other property of laws which may be specifiable in Humean terms.

2.3 Against primitive higher-order connections

I now move on to consider primitive higher-order connections, which a Humean methodology precludes from among the determinants of an account of laws along with primitive causal and subjunctive connections. A higher-order connection relates entities which can have instances. In particular and I will focus mainly on higher-order connections between universals (n-adic for any n), entities referred to (if they exist) by nominalisations of relational property terms, e.g. 'proximity', 'resemblance', etc. Although I also aim to critique here so-called dispositional essentialist views which bear significant similarities.

The need for a higher-order connection in accounting for laws has long been mooted on the grounds that without it, the laws must be grounded in the very regularities they aim to explain (Armstrong 1983, Bird 2007a, Maudlin 2007, Mumford 2004, Swoyer 1982). Many Humeans, particularly those of a metaphysical stripe, can be interpreted as understanding laws to be metaphysically explained by, exist 'in virue of', or be *grounded in*, the first-order regularities which exist in the mosaic (e.g. Beebee 2006a, Lewis 1986g, Loewer 1996, 2012). But *prima facie*, it is laws which explain why the world is as it is including the mosaic itself. So it is that Maudlin argues,

> if the laws are nothing but generic features of the Humean Mosaic, then there is a sense in which one cannot appeal to those very laws to explain the particular features of the Mosaic itself: the laws are what they are in virtue of the Mosaic rather than vice versa. (2007, 172)

The problem posed to the Humean is, therefore, one of circularity: for the Humean, facts in the mosaic explain why the laws are as they are, but the laws are also supposed to explain why the material world is as it is and for the Humean that is precisely the mosaic. In §2.4 I will present what I think is the right Humean response to this problem drawing, in part, on recent work by Loewer (2012). However, here I want to consider whether or not being non-Humean actually avoids the problem of laws' explanatory circularity. Since this problem is what I take to be the main motivation for positing primitive higher-order connections it will be my main focus. But it's also worth saying why I don't think the previously considered primitive connections will help either.

Accepting a primitive causal tie might explain why certain features of the mosaic are the way they are (in conjunction with reference to other features). But they could not alone provide a ground for the *laws*. The laws at least require a regularity to hold, and the mere fact that causal connections exist between events is not sufficient to entail that any of those exist. After all, the connections could exist between thoroughly random pairs of property-instantiations, with no generalisable manner at all. But if the regularity is what we are citing as further grounds for laws beyond primitive causal connections, then the circularity problem has not been removed by positing such connections.

Lange's understanding is, as we have seen, that the laws are grounded in primitive subjunctive facts. But can we explain why the mosaic is as it is without risk of circularity by reference to those subjunctives? I for one cannot see how this would work. If the mosaic is different to how we think it is, then surely the subjunctives will be too. For example, if, unknown to us, the mosaic actually contains magnetic monopoles, then it may not be the case that if a magnetic field had existed at the centre of this page then it would have been dipolar (i.e. we might be mistaken in reasoning counterfactually in accordance with Gauss's law of magnetism). The fact that we are led to posit different subjunctives under different assumptions about how the mosaic is suggests that the state of the mosaic partly explains what subjunctives are true. So, even if subjunctives have an element of irreducibility about them, at least part of the explanation for them is that the world is as it is. If that's the case, even Lange's own view won't get him out of the circularity.

The alternative and popular direction taken by many who are concerned about these explanatory problems concerning laws is to posit higher-order connections between universals. One philosopher well-known for doing this is Armstrong (1983, 1997). As he put it, 'we need to put some "distance" between the law and its manifestation if the law is to explain the manifestation' (Armstrong 1983, 41). Consequently, Amstrong posited a contingent relation of 'nomic necessitation' $N(F, G)$ which would 'add to the first-order facts' (Bird 2007a, 91) by relating properties F and G whose instances' constant conjunction is understood to be lawlike. According to Armstrong, the relation of nomic necessitation N is a universal itself of a higher-order than those it relates. Moreover, the nomic necessitation relation is a *primitive* and so cannot be analysed in terms of its relata or its relatas' instances. According to the view, laws are made true by instances of this nomic relation, thereby distinguishing them from purely accidental regularities.[5] It is this move which is supposed to establish the much desired 'gap' between laws and regularities and so open up a space for showing how

[5] In fact, Armstrong thought we should refer to the atomic fact $N(F, G)$ as the law itself. I ignore this complication, hopefully at no cost to the argument.

laws can explain the particular features of the laws' 'manifestation' in the mosaic.

Armstrong's view has come under considerable criticism and Armstrong eventually abandoned it in this form (see Armstrong 2004b, 2005). Lewis was one of the first to express concern:

> N deserves the name of 'necessitation' only if, somehow, it really can enter into the requisite necessary connections. It can't enter into them just by bearing a name, any more than one can have mighty biceps just by being called 'Armstrong'. (Lewis 1983, 366)

In other words, by simply naming the relation one of 'necessitation', we don't thereby come to see why it should play a necessitating role. A more carefully worked out expression of this problem was presented by (see Van Fraassen 1989, pp.96-8) in the form of a dilemma between available solutions to an 'identity problem' and an 'inference problem'. If $N(F, G)$ necessitates first-order regularities $\forall x(Fx \rightarrow Gx)$ by definition then at best it puts a name to whatever it is we are trying to explain and at worse simply misidentifies the relation as one satisfied by any accidental regularity (e.g. the relation of extensional inclusion). Assuming the relation is not merely extensional inclusion, then we have a problem of *identifying* it. But if, on the other hand, the $N(F, G)$ does not necessitate by definition, then there is the problem of saying why it should be the case that we are able to *infer* the regularity from $N(F, G)$.

In response, Armstrong argued that a regularity in the mosaic, say between things being F and being G, is explained '*because* something's being F *brings it about* that that same something becomes G'. In this context 'brings it about' is supposed to be understood causally: 'the required relation is a causal relation, the very same relation that is actually experienced in the experience of singular causal relations, now hypothesized to relate types not tokens' (Armstrong 1993, 422; see also Armstrong and Heathcote 1991, Armstrong 2005). For Armstrong, this solves the identity problem because the necessitation relation is not understood as a mere *ad hoc* posit but a real natural feature of the world which holds between universals rather than particulars, viz. *causation.*

Despite being posed to him as a dilemma between identity and inference problems, Armstrong presented this response as a solution to *both* problems. As he remarks,

> if a certain type of state of affairs has certain causal effects, how can it not be that the tokens of this type cause tokens of that type of effect? The inference is analytic or conceptual.' (*Ibid.*)

Assuming that the token instances of the first-order regularities which we wish to explain simply are such causal relations (something to be queried anyway, in§4), then Armstrong is here taking the inference problem to be solved too.

We might wonder, however, whether this really has made the problem go away. More recently, Bird has shown that since, for Armstrong, universals are categorical (i.e. they have no non-trivial modal properties) he is committed to it being metaphysically possible that $N(F, G)$ holds (for some universals F and G), and yet also that $\neg\forall x(Fx \rightarrow Gx)$. Hence it is not possible to deduce the regularity—i.e. $\forall x(Fx \rightarrow Gx)$—from $N(F, G)$ alone. What is needed, it seems, is some sort of *even-higher-order* connection N' between the regularity and the nomic necessitation relation in order to support the inference. But if we stick to categorical relations only, then Bird argues we will only have the same problem at a higher level and a vicious regress ensues.

Bird then offers his own solution to the inference problem. Talking of Armstrong and other philosophers who have taken a similar approach to laws (e.g. Dretske 1977, Tooley 1987), Bird remarks that they,

> take nomic necessitation to be soft, a contingent relation, because of which they cannot answer the Inference Problem. The Inference Problem is solved only if necessitation has an essence (essentially, if $N(F,G)$ then $R(F,G)$), (2007a, 97)

In this context, $R(F,G)$ is the "extensional inclusion relation" 'between F and G that holds whenever $\forall x(Fx \rightarrow Gx)$' (Ibid.).

One could be forgiven for admitting some confusion over this response. First, the contingency Bird refers to cannot be that of $N(F,G)$ itself. For it is not in itself problematic if $N(F,G)$ does not hold in all worlds, only that it can apparently hold at worlds where the regularity doesn't hold. Therefore, it would seem more appropriate to attribute the failure to answer the inference problem to the *categorical* nature of N rather than the contingency of its instances. Second, is $R(F,G)$ a reasonable posit on behalf of Armstrong's account? After all, if R is a genuine categorical universal (as they all are for Armstrong), then if there is an inference problem from $N(F,G)$ to the regularity $\forall x(Fx \rightarrow Gx)$ there will be one between $R(F,G)$ and the same regularity for the same reason. Therefore, introducing $R(F,G)$ into the discussion does not appear to be very helpful and we should interpret Bird's solution as the suggestion that essentially if $N(F,G)$ then $\forall x(Fx \rightarrow Gx)$.

Given these observations, Bird's solution to the inference problem (put simply) aims to provide a justification for rejecting the categorical nature of the necessitation relation. I will come to consider the details of this justification shortly. But before that I want to consider whether such a rejection (justified or otherwise) even addresses the problem we should be at pains to solve. For the initial problem with Humean approaches to laws concerned explanation: specifically, how the laws can both be grounded in regularities in the mosaic and also explain instances of those regularities. Armstrong wanted to try to ground laws in something other than regularities so that an informative explanatory relationship could be reinstated. On the one hand, if the inference problem which faces Armstrong's account is understood in terms of a failure of *explanatory* inference, then we can start to see how its conclusion would seem particularly troubling. For the account was supposed to improve on the Humean view by showing how explanations of regularities in terms of laws were to be possible. But on the other hand, if the inference problem is interpreted as one concerning a failure of *logical* inference, i.e. entailment, we may wonder whether it's even a problem worth solving.

Indeed, I'm tempted to think that solving the inference problem interpreted as one of logical entailment is completely orthogonal to what's really at stake, namely, an inference problem interpreted as one of explanation. No doubt much of the confusion has stemmed from the repeated use of the term 'necessitation'. But if I am right, whether the grounds of a law necessarily implies the regularity is irrelevant to whether or not it explains the regularity. After all, entailment is certainly not *sufficient* for explanation. From the fact that we know the cat is on the mat we can infer (because it is entailed) that the cat is on the mat, but the fact that something is known doesn't explain what is known. This should lead us to wonder whether Bird's closure of the logical gap by positing an essence to laws, or the properties therein, is sufficient to close the explanatory gap. Certainly it is not a rejection of the categorical nature of properties *alone* which confers explanation. And notice also that it is far from obvious that entailment is *necessary* for explanation. If there is something in the world like primitive causation (as Armstrong appears to think), we might be open to explanations which don't license a logical entailment to what they explain. What might explain that the cat is on the mat is the presence of sunshine on the mat (the mat is nice and warm), but such a fact need not support in any obvious way an entailment that the cat will be on the mat.

We would, I think, do well to restate the issue with Armstrong's account in terms of explanation directly. In these terms, the dilemma posed by Van

Fraassen amounts to the difficulty in avoiding triviality through simply identifying $N(F,G)$ with whatever explains instances of lawlike regularities in the mosaic while at the same time enabling whatever account of $N(F,G)$ we do provide to inform us of why those instances are happening.

Posed in these terms, is Armstrong's account successful? I think the answer is still 'no'. But the reason why will help us see why the response of rejecting categoricalism is no improvement. Recall that Armstrong's response to the explanatory dilemma is to identify N with the causal relation holding between universals. With such a move Armstrong appears to both avoid trivially identifying the nomic relation $N(F,G)$ with whatever explains why any F is a G and avoid misidentifying the relation with the extension of all token causal relations. But now it is incumbent to say why such an identification avoids an inference problem. For let us grant that causes explain their effects, and even that the existence of a causal relation between cause and effect explains why the cause explains the effect. None of these concessions serve to explain why a causal relation which exists between two entities, in particular between one universal and another, should explain a causal relation between any other two entities, specifically an instance of the first universal and an instance of the other.

Here, it might be pointed out that actively looking for such an explanation is not required. As Armstrong exclaimed, 'if a certain type of state of affairs has certain causal effects, how can it not be that the tokens of this type cause tokens of that type of effect? The inference is analytic or conceptual' (1993, 422). Personally, I do not share Armstrong's intuition of analytic or conceptual inference. Moreover, I'm concerned that if it is analytic or conceptually true that causal relations at the type-level (i.e. at the level of universals) entail causal relations at the token level (i.e. at the level of the first-order regularities) then the response is simply making the very mistake of triviality that the identity problem picks up on. But regardless, now we know that it is explanation we want and not mere entailment, such a response will seem anyway unhelpful. For even assuming Armstrong's conceptual intuitions, why should we grant that the existence of a causal relationship between types of state of affairs *explains* why the tokens of these types exhibit causal relationships. For all that has been said, couldn't it be the other way around or couldn't they share a common explanation or couldn't they be logically related in a way completely distinct from explanation altogether?

To make the point more salient, consider briefly a different view of laws: that laws are definitions (proposed by Le Roy 1901; see also Sellars 1948). If that were the case, then two predicates which feature in the expression of a law will be related conceptually: part of what 'being F' means is what 'being G' means. In §2.2 I gave the example of the total force law, which has been argued to be true by definition: the value of force is *defined to be* equal the value for change in momentum. For the sake of argument, let's grant that this is indeed the case with the total force law. Let's also assume that like all analytic truths, definitional truths are true of necessity (even though there appear to be counterexamples to this; see Williamson 2007). It follows by necessity that the regularity between force and change in momentum must hold, hence the inference problem interpreted as one of entailment is thereby solved. But what about the inference problem interpreted as one of explanation?

I claim that it is *not* thereby solved (the following discussion is inspired by Williamson 2007, ch.3). Say we have a number of definitions at our disposal including that the value of force is by definition equal to the value of change in momentum. Now there must be some process by which these definitions come about, be it by convention, by expert executive decision or by God's decree. But suppose this process were to have been different, would we thereby expect any difference in the way objects behave? For instance, suppose we decided tomorrow to redefine force to be merely a change in mass. Apart from changing our naming conventions, would we have to radically revise the way we live in the world, reengineering all our vehicles, recalculating the trajectories of all known projectiles, etc.? Presumably not. When we change a definition, we change what part of the world our words refer to, not the world itself. Just as whatever it is that explains why each bachelor is unmarried is not that being a bachelor entails being unmarried, what *makes it* the case that the value of force is equal to the value of change in momentum is not the fact that it is true by definition (though it may in fact be), but something about the way the world is.

Consideration of such a conception of laws shows that merely assuring a logical entailment from law to regularity (or a regularity's instances) is not enough to solve the inference problem interpreted as one of explanation. So why should one expect to solve such an inference problem by rendering an unexplanatory higher-order connection non-categorical, as Bird advises?

Take for instance the view Armstrong eventually came to accept (Armstrong 2004b, 2005) in which universals are identified partly by the particulars which instantiate them (the details are remarkably similar to an under-cited paper by Sellars 1948 who argues for an individuation of properties on Leibnizian grounds). Armstrong therefore ended up denying his earlier assertion that universals should be categorical admitting that 'if one thinks of nomic connection as a higher-order relation holding between first-order universals, and one extends the partial identity idea to this predication, then laws of nature become (strictly) necessary—an unexpected result, but not unwelcome to me' (2004b, fn.8). Nonetheless, securing the entailment of regularities from laws makes no improvement on the earlier account when it comes to solving the inference problem interpreted as one of explanation. In fact, it can seem quite the opposite. After all, according to the new view, universals are *partly identified by* particulars, so that first-order facts are once again playing a very key explanatory role in higher-order universals' (including nomic relations') individuation.

Bird's own solution may seem to go beyond merely assuring a necessary entailment of regularity from law. For his position concerns the *essence* of universals. The idea seems to be that laws should be grounded in the natures of the properties themselves (see also Swoyer 1982). Although Bird is not explicit about it, it is popular now to consider essential truths to be rather different from necessary truths (e.g. Fine 1994, 2012, Gorman 2005). On this understanding, the essence of some nomic relation is just the sort of thing which would indeed explain its consequences. For instance, if *being* H_2O is the essence of *being water*, then *being* H_2O should explain why water does the things it does (e.g. boiling at certain temperatures, beading, forming hexagonal crystal lattices, etc.).

But even under this plausible interpretation of essence, it's hard to see how Bird's solution is supposed to help us see how $N(F, G)$ would thereby come to explain the first-order regularity.[6] For if it part of the essence of the nomic relation $N(F, G)$ that the regularity $\forall x(Fx \rightarrow Gx)$ follows, then we have the very same explanatory problem posed to Armstrong's contingent necessitation relation. If the essence of a nomic relation or property is simply to be *identified* as an entailment of regularity then at best the description of the essence simply puts a name to the very explanatory mechanism which higher-order theorists are trying to describe and at worse the description of essence picks out something higher-order theorists are explicitly trying to avoid, viz. the regularity itself.

Clearly the essentialist wants to avoid this, perhaps by deeming the essence of higher-order nomic relations to be something like a genuine natural relation between universals. But if that's the case, we once again have the problem of saying why it is relevant to the regularity. Bird's fully considered account effectively places the higher-order connections in the identity relations between universals (see esp. Bird 2007b).[7] Via analogy with graph-theory (similar to

[6] Bird's more developed suggestion is that we in fact do away with the N relation and consider the essence of the first-order properties themselves; I don't see how this affects my point here and ignore the complication for the current discussion.

[7] Perhaps this stretches the suitability of the term 'higher-order', but the identity relations exist between universals all the same.

Dipert 1997), a structure of such identity relations may indeed be sufficient to individuate the properties (although see Barker 2013). But now the problem becomes one of saying how we are supposed to make sense of the explanatory relevance of these universals to the regularities which fall under them. I can see no reason to suspect that simply by finding a relational way to individuate the universals we thereby entail, let alone *explain* why the regularity occurs.

To provide an esoteric example, perhaps the natural relation between universals which individuates them according to the graph-theoretic approach is the ordering God has given them in her Platonic heaven. There does not appear to be any reason to think that *given that fact alone* the natural relations between universals will entail that any kind of ordering between the instantiations of those universals. And even if some case can be made for an entailment between relations among universals to relations among particulars, there is still the requirement that the entailment be an explanatory one. Given the assumptions, I see no reason to prefer one direction of explanation over the other. After all, couldn't it be the case that God orders the universals in that way *because* of the way the particulars are distributed in the mosaic?

Perhaps much about the complaints I'm leveling at the utility of higher-order connections boils down to the long-argued for complaints from nominalists about metaphysical realist positions (e.g. Quine 1948, Daly 2005, Devitt 2009). The general feeling is that these further metaphysical entities are not posited in the right kind of way to warrant any belief in them as explanatory of the phenomena we set out to explain. Ultimately, if metaphysical devices like higher-order connections between universals are to do the explanatory work for which they have been posited, they require an explanatory basis entirely distinct from mundane first-order phenomena and yet be able to explain that phenomena. If there's a moral to draw, it's that when metaphysicians like Armstrong claim there needs to be some 'distance' put between what's explained and what's doing the explaining, metaphysical realists are liable to posit so much distance in searching for the explanans that they can't then chart an explanatory relationship back again. And if they can't do this, it's unclear why one should feel compelled to abandon the Humean approach to which they are opposed.

Of course, it might be argued that assuring a 'distance' between laws and what they are supposed to explain is an insufficient but nonetheless necessary step in accounting for the explanatory relationship. Perhaps that is correct. In which case, if there is explanation to be had, we'd better adopt *some* more exotic metaphysical position than Humeanism. But to push back on this sort of way of arguing, one might simply question the very motivation. For instance, nominalists and property-realists will dispute, whether or not resemblance is something which requires explaining (see the discussion in Pereyra 2002). For the realist this is something strange which demands an expanded ontology. The nominalist can object to this on the grounds that real properties are mysterious and that the explanatory relationship is not obvious simply given the posit, but they may also object that resemblance is simply not something which requires explaining. After all, why should we be surprised that some things resemble each other and others don't? Would it be less strange if nothing resembled anything else? Similarly, the Humean might respond not only that it is far from obvious how higher-order connections explain their instances, but also that instances are not something which really need explaining in such a deep metaphysical way anyway.

This last idea becomes key, I think, in giving the Humean's best response to the circularity issue we started the section with, to which I know turn.

2.4 Getting mosaical about laws and causation

I advocate a Humean methodology. That is, I believe an account of laws and causation should be given which makes no reference to primitive causal, subjunctive and higher-order connections. Being Humean about causation has been motivated in two ways. First, by a defence of the Humean idea that a causal tie or connection cannot be observed in instances of causation, and second, by pointing out the irrelevance of primitive causation in providing us with the kinds of utility we typically take causal knowledge to provide us with. It is noteworthy that neither of these arguments entail the non-existence of a primitive causal relationship. It remains logically possible that a causal tie really does exist between all and only causally related events. (This is why my Humeanism is only methodological rather than existential or reductive.) Nonetheless, I have argued that a posit of primitive causal tie would bear no use to us in making use of causal knowledge in prediction and intervention nor for making sense of how laws come to be. Such ties should, therefore, be considered as good as non-existent for the sake of discussion of laws.

Being Humean about laws was motivated as a result of dissatisfaction with the alternative proposals, in particular, those employing primitive subjunctives and primitive higher-order connections. In §2.3 I discussed an argument which aims to show that if one accepts a Humean methodology, one cannot thereby explain the on-goings of the mosaic in terms of laws. Whether or not that is true, it seems to me that no alternative philosophical account of laws has persuasively offered a way to do this either. Given this result, we lose the impetus to invoke metaphysical posits like higher-order connections of necessitation (as Armstrong has suggested) or individuative higher-order manifestation relations between properties (as Bird has suggested). So, while I think the arguments presented in §2.3 are more careful than some of the passing complaints launched at non-Humean approaches (e.g. Lewis 1983, Van Fraassen 1989, Braddon-Mitchell 2001, Ward 2002), they amount to much the same conclusion.

But what of the intuition that laws can explain the mosaic? Here I want to discuss a number of answers the Humean might provide and some potential problems with them. While I consider the verdicts on all of these to be inconclusive, I will suggest that a promising response is to accept some degree of failure for laws to explain the mosaic whilst showing how laws can nonetheless be explanatorily useful in other ways.

One good place to start is with a distinction drawn by Loewer (2012) between 'metaphysical explanation' and 'scientific explanation'. According to Loewer, metaphysical explanation is an atemporal determination relation of some kind, often considered to be a relation of grounding of the explained fact in a more fundamental fact (e.g., Rodriguez-Pereyra 2005, Schaffer 2012, Fine 2012). Scientific explanation is different. It need not explain by drawing attention to more fundamental facts but, 'typically, shows why an event occurred in terms of prior events and laws' (Loewer 2012, 131). By drawing this distinction Loewer hoped to cast doubt on whether there really is a circularity worry for Humeans. The mosaic metaphysically explains the laws and the laws scientifically explain the mosaic; two different relations, hence no circularity.

A recent argument from Lange aims to bring the relevance of this kind of distinction into question. Lange begins by positing a *prima facie* plausible 'transitivity principle', that the metaphysical grounds of any entity scientifically explains whatever that entity scientifically explains. He then shows that if laws are grounded the mosaic, it will lead to the commitment that the mosaic scientifically explains itself or laws do not explain at all—ultimately, a no better position for the Humean than before Loewer's distinction.

> [I]f the law that all sodium salts burn with yellow flames helps to scientifically explain why a given flame is yellow, but this law is grounded in the Humean mosaic, then [...] the Humean mosaic would have to help scientifically explain why the given flame is yellow. The Humean mosaic, in turn, consists of the fact that the given flame is yellow [...] together with a host of other such facts; a giant fact capturing the complete Humean mosaic holds in virtue

> of various local Humean facts, one of which is the fact that the given flame is yellow. Hence, [...] it follows that the flame's being yellow helps to scientifically explain itself. That cannot be. (2013, 258)

I will assume that reasonable criticisms of Lange's argument will not gain purchase by querying the grounding relation itself. For although Lange's transitivity principle and subsequent examples are given in terms grounding, it doesn't seem to rely on any particularly metaphysically contentious relation (a 'big-"G"-grounding' in Wilson's 2014 terms). All that is needed from the relation referred to as 'grounding' by Lange is that there is an explanatory relationship between regularities in the mosaic and the laws (in that order). That is certainly something Loewer seems to accept. In Salmon's 1984b terms, the explanatory relationship must obviously be *ontological* rather than epistemic. But granting that, the principle would be no less plausible if we read 'grounds' as 'causally explains' or 'constitutively explains'. In fact, it seems a limit-case of the transitivity principle holds also if we interpret grounding simply as identity. After all, surely anything identical with an explainer of X is also an (ontological) explainer of X.

A more productive line of criticism grants the grounding relation but queries the transitivity principle. Hicks and van Elswyk (2015), Miller (2015) argue that in cases where a multiply-realisable entity, e.g. a lion, is grounded in some realiser, e.g. the lion's particular token atomic composition, the realised entity may scientifically explain facts which the realiser does not, e.g. the number of prey animals in a given region. But by amending the transitivity principle to be contrastive (something Lange considers independently motivated) Lange (forthcoming) shows that the relations of explanation in apparent counterexamples to the transitivity principle have a structural difference to that which exists between laws and the mosaic for Humeans. Whereas it is the lion's presence *rather than absence* which helps scientifically explains the number of prey in the area, it is not the case that the electron's presence *rather than absence* helps metaphysically explain the lion's presence *rather than absence*. But according to Humeanism it seems to be the case both that a generalisations' lawhood rather than non-lawhood helps explain the presence of confirming instances in the mosaic rather than disconfirming instances *and* that the presence confirming instances in the mosaic rather than disconfirming instances explains those generalisations being laws rather than non-laws.

Another objection to the transitivity principle from Hicks and van Elswyk is that it 'summarily dismisses the anti-reductionism of Fodor (1974)' (2015). Their reasoning is that since grounded entities may not be reducible to what they are grounded in, what they are grounded in cannot help scientifically explain what the grounded entities help scientifically explain. Lange concedes the irreducibility of some grounded entities but denies that this supplies a reason to reject the transitivity principle. It is possible, Lange argues, that a grounded entity might scientific explain some other phenomenon because it unifies the latter in a way which the grounding entities cannot. But this only serves to show that the grounded entity can scientifically explain the phenomenon in a non-superfluous way, and does not entail that a 'sufficiently full and complex account' in terms of the grounded entities might serve to provide *some* scientific explanation of the phenomenon as well. It seems, therefore, that the objection is insufficient to motivate a denial of the transitivity principle.

A different sort of a query with the transitivity principle aims not to cite examples or types of example which disobey it but rather asks for a reason to assert it. Clearly instances of transitivity exist in explanation between grounding explanations and scientific explanations. But why should we think they *always* do? In private conversation, Luke Fenton-Glynn has expressed this thought suggesting we might rather opt for 'transitivity principle *lite*' which stipulates that if a helps metaphysically explain b and b helps scientifically explain c, *and a, b and c are all distinct*, then a helps scientifically explain *c*. If we accepted this principle rather than Lange's stronger principle, we could argue that laws' grounding in the mosaic renders the two indistinct, hence the transitivity cannot be established in this case. This sort of response to Lange seems coherent to me, although I suspect some argument for the modified principle is required in order to persuade those who already accept what I take to be the more immediately intuitive version of the principle that Lange provides. And presumably merely reiterating the merits of Humeanism is unlikely to be sufficient.

Ultimately, I suspect that Humeans will have to accept the force of Lange's argument. Laws cannot be both grounded in the mosaic *and* explain it. Assuming that Humeans cannot avoid the grounding claim (which need not require invoking any metaphysically loaded relation beyond that which exists between universally quantified facts and facts about their instances) then it seems the Humean may have to avoid trying to show why laws explain the mosaic. But I don't think this should pose a major threat. After all, in the foregoing section I suggested that non-Humean accounts have also failed to provide any convincing account of such explanation. Moreover, the Humean can incorporate a dismissal of the idea of explanation along with a well-rehearsed complaint that non-Humeans cling too heavily to the outdated metaphor of laws laid down by God to govern the procession of nature (Loewer 1996, Beebee 2000). Very few metaphysicians accept this in any literal sense today, a fact which entitles Humeans to legitimately ask whether there is value in clinging to any vestige of it as mere metaphor. As with the idea of laws as God's decrees, the Humean might also deem as misguided the idea that laws are the sorts of things which explain why nature's mosaic behaves as it does.

But a Humean might want to say something more conciliatory about where the idea that laws explain comes from. If this is the aim, a better position for the Humean to take might be to shift the focus away from trying to show why laws explain the mosaic and instead point towards a different explanatory relationship in the vicinity. I will propose my own suggestion shortly, but before that I should mention one attempt which I think ultimately fails from Skow (2016). His suggestion is that when we cite a law in an explanation of some fact or occurrence, the law itself does not provide the explanation, but rather citing the law 'implicates that some of the reasons why' (p.22) a particular phenomenon occurred are the same as those of other instances of the laws' antecedent type.

> If it does not just happen to be true that all sodium salts burn yellow, if it is a law of nature that they do, then by citing this law in response to the question why this powder burns yellow one implicates that some of the reasons why this powder burns yellow are the same as the reasons why all sodium salts in the actual, and also in other physically possible worlds, burn yellow. (2016, 22)

Obviously, unless *being sodium* has something to do with burning with a yellow flame, the reference to the law will seem irrelevant in such an explanation. So then what makes *being sodium* relevant? Here Skow is not clear. Drawing on Lange (forthcoming) I think we can undermine the possible options. Skow's ultimate concern is (roughly) to defend the claim that all explanation is either causal or grounding explanation. On the first option, the law that all sodium salts burn with a yellow flame implicates the fact that all sodium salts have something in common—a certain electron structure—which *causes* a yellow flame upon burning. But here it looks like we have simply resorted to another (more fundamental) law to explain the yellow flame, viz. that all salts with that electron structure burn with a yellow flame. This is exactly what Skow (and the Humean approach under consideration) was trying to avoid!

A different tack offered by Skow is to suggest that 'facts about the laws help ground facts about causation' (p.28). Rather than laws explaining facts about the mosaic, they are higher-level explainers, metaphysically *explaining why* causes explain their effects. So, for example, the law that sodium salts (or salts with a certain electron structure) burn with a yellow flame *grounds* the fact that each sodium salt causes a yellow flame when burnt (see Marshall 2015, for a similar view). But this is dubious. Just as we might be suspicious of someone inferring a rule on the basis of some premises by making use of that very rule, similarly we should be suspicious about explaining some fact on the basis of certain explanans by making use of the explained fact. If the Humean accepts that laws help ground facts about why one event in the mosaic causes another, *and* that those causal events ground the laws, then they seem to be once more in an undesireable circle.

Perhaps there is a way out of this worry. Taking inspiration from Papineau's (1992) discussion of inductive inference we might aim to distinguish 'rule-circularity' from 'premise-circularity' in the case of explanation noting that there may be reasons to mitigate the former despite problems with the latter. Regardless, in what remains of this section I want to discuss a different approach which moves away from trying to explain the mosaic entirely. My hope is that it will still resonate to a sufficient degree with the intuitions which drive the Humean's circularity worries.

The idea is that independently of whether or not laws can explain the first-order on-goings of the mosaic, they can nonetheless explain our *epistemic status* regarding those on-goings. Begin with the idea, accepted by many Humeans and non-Humeans alike, that the world is one in which inductive inferences can be performed. Based on observed data, we are able to come to know, via inference, generalisations which hold in unobserved cases as well. For Humeans, there is no particularly deep reason for this, the world just turned out facilitate such inferences. This has been a bone of contention with Humeanism in the past, but I take it that recent responses have undermined the idea that there is anything particularly problematic about Humeanism when it comes to justifying inductive inference (see especially Beebee 2011).

If inductive inference weren't a justified form of reasoning, we would not be able to come to know what was going to happen in the future, distant past or very far away. Of course, we might predict some outcome and get it right by mere coincidence, but I take it for some inductive inference to count as *knowledge*, the circumstances must be such that the inference is *justified*. Since Gettier (1963), we now know that such justification cannot be a mere matter of internal reasoning. In the case of inductive inference, the regularity must actually hold in general across observed and unobserved data *and be lawlike*. An induction that the ten coins in my pocket are silver as a result of me testing two may turn out right, but it will not be justified and so not count as knowledge. An induction that light has the same speed in all reference frames as a result of me testing a small portion of light sources (e.g. with the Michelson-Morley setup) may indeed be justified and so count as knowledge. This is because the regularity concerning the coints is not lawlike, whereas the regularity concerning light is.

So, for the Humean, the fact that the mosaic grounds inductively inferable laws is part of the reason we are able to make *knowledgeable* predictions. If laws weren't grounded in the mosaic, we would have a significantly reduced capacity (if not entirely eradicated!) to perform knowledgeable inductions and consequently also to knowingly predict events beyond the data readily available to us.

Consider an example. Humeans will typically accept that there are many ways the world could have been such that if I had let go just now of my pen held in my outstretched arm it would accelerate (roughly) in the direction of the centre of the earth's mass. Some of those ways are such that nothing has ever accelerated in the direction of the centre of the earth's mass before or ever will again, other ways things do so only rarely. Some ways the world comprises nothing at all through spacetime but for a single event in which the earth, me and the pen come into existence for a brief enough time for me to let go of the pen and it accelerate towards the earth. I take it that in such ways the world could have been, it would be incorrect to attribute *knowledge* to me that if I were to drop my pen it would accelerate (roughly) in the direction of the centre of the earth's mass. The Humean might plausibly attribute this to the mosaic in each case. The mosaic has insufficient facts to support the kind of inductive inference required to confer on me knowledge of the pen's acceleration, even if it will so accelerate.

Now compare these ways the world might have been with the way the world actually is. The mosaic of this world is such that *every* uncharged object accelerates towards the centre of mass of the collection of objects whose gravitational field it is in. More specifically, in the way the world actually is, this generalisation is a *law*. Because of this, I am able to come to know it via inductive inference (though not necessarily *my* inductive inference!). Consequently, I *know* that if I let go of the pen held in my outstretched arm, it will accelerate (roughly) in the direction of the centre of the earth's mass.

Notice that there is no circularity here. The mosaic (let's say) includes the fact that I dropped my pen which helps metaphysically explain the gravitational regularity by constituting an instance of it. Indeed, for the Humean the mosaic will explain more than that, for it will explain *why the regularity is a law* (the details on why *this* is are the subject of §4.2, §5 and §6). The fact that the gravitational regularity is a law explains why I am justified in inductively inferring instances based on other knowledge. By transitivity, this ultimately means that the mosaic explains my knowledge. But this isn't circular. The mosaic doesn't explain *why things occur* in it (which would be circular) but it explains *why we know things occur in it*.

It's important to be careful with this alternative story about laws' explanatory power. For a typical Humean, people with inductive knowledge are themselves part of the mosaic and so (in a sense) is our being in such a knowledgeable state. But the fact a mental state is one of knowledge (rather than justified belief) is grounded in facts which extend beyond the internal features of the knower.

For instance, to be knowledge that some inductively supported prediction will actually occur it must in fact occur, and this is not only because knowing p entails p. For the Humean, inductively supported predictive knowledge is justified on the basis that what is predicted is inferred from a law which will be grounded partly in the occurrence of that very prediction. Therefore, the Humean will say that knowledge that an outcome is going to occur on the basis of induction is justified partly by virtue of the fact that it does occur.

Is *this* circular? I can see no reason to think so. Notice that the kind of justification being considered here is *external* to the knower. Classic externalist accounts of knowledge have often included criteria for justification grounded in the truth of what is known. For example, the criterion of a causal relationship between the knower's belief and what is known (Goldman 1967) and the criterion of truth-tracking conditions between the knower's belief and what is known Nozick (1981). The requirement that there be a law according to which some knowledge-conferring inductive inference is performed is no different in this respect. Of course, the suggestion that the justificatory component of knowledge also grounds the truth of what is known would be problematic if it required the knower to justify *to themselves* (i.e. internally) what is known on the basis of that very knowledge. But that is not what is being suggested. The knower who makes an inductive inference to an unobserved case justifies this to themselves by reference to *observed* data. But whether or not they know in this case has partly to do with whether or not they are inferring on the basis of a law grounded in the very unobserved case they are predicting.

If the above reasoning is correct, laws do explain *something*—not facts about the mosaic itself but rather why we are able to know via prediction facts about the mosaic. Knowledge of what to predict is certainly a paradigm epistemic status that laws help provide us with. But it will also be important for the sake of the account of laws developed below to keep in mind that laws also provide us with knowledge about how to *act*. After all, we are not like Dummett's trees, unable to manipulate the world around us. And much of our ability to do so intentionally and successfully is supported by our knowledge of the kinds of generalisations we can inductively infer (Woodward and Hitchcock 2003). I know that letting go of the pen will (in the context) lead to its acceleration out of my hand. I know this partly because of my knowledge (albeit, perhaps, in 'folk' terms) of the law of gravity. Contrast this with a case in which I drop the pen on the blind faith that it will fall or as a spontaneous reaction (as a child might) or mere caprice. Knowledge of the laws not only allows us to make informed predictions but informed *goal-directed actions* as well.

As with justified prediction, for the Humean a goal-directed action motivated by the laws and their consequences partly ground the very generalisations one acts on the assumption of. This means that the justification for the knowledge of how to act to achieve one's ends is partly grounded in the action and its consequence. But as with prediction, such justification is external to the knower, and there is no subsequent risk of justificatory circularity.

In sum, the sort of explanatory issues discussed in §2.3 can be sidestepped if we no longer aim to support the premise that laws explain their instances. This seemed to lead to trouble whatever account of laws we took. The suggestion I have just given is that we treat laws as instead explanatory of *our knowledge of* the mosaic. This includes the kind of knowledge we can have when we make predictions on the basis of laws and the kind of knowledge we have when we consider a goal-directed action. Obviously this won't satisfy those who consider it of paradigm importance for the Humean to explain the ongoings of the mosaic directly, but it seems to go some considerable way to placating the concern that laws for the Humean are explanatorily impotent.[8]

[8]One might wish to reflect at this stage on whether any of this impacts what we can say about the *governing role* of laws. Humeans have generally been dismissive of this characteristic of laws (see Loewer 1996, Beebee 2000), citing the intuition that laws govern as residual of a misguided theological metaphysics. A more nuanced position has been put forward by Roberts (2008), who suggests that we can be Humeans and accept the governing role of laws so long as we take it to merely be the characteristic defined by NP (see §2.2. If that's all there is to governing then I am happy to go along with Roberts and accept

2.5 Humean methodology and the rest of this thesis

With the Humean methodology established, I can move on to develop an account of laws and, subsequently, of token causation. Having said that, it is worth drawing attention to those conclusions drawn in what follows which do not require the acceptance of a Humean methodology or else can be accommodated by non-Humeans without too much adaptation. One obvious conclusions is that drawn in §3.1 which explores the logical form of laws. While I have the Humean position in mind when characterising laws' logic (for instance, as a conditional concerning first-order particulars), the general argument that laws are conditionals does not require this. Nor is this required by the discussion of the individuation of system-types in §3.2.1. Also, §5 itself will not attempt to make any Humean analysis of the causal features associated with laws (a task left to §6). As I will point out, the conclusions of §5 might therefore be used to formulate a non-Humean account of laws with type-level causation as a primitive. Similarly, I will avoid framing the analysis of token causation presented in §7 in any explicit Humean light, although I openly remark that the success of the Humean analysis of laws in §6 would render the analysis in §7 Humean also.

In general, since Humean methodology is defined negatively, in terms of the prohibition of certain primitives from the analysans of phenomena of interest, views which reject Humeanism by introducing the prohibited entities will not obviously come into conflict with the accounts developed here. Indeed, one might suspect that if the accounts I provide are satisfactory with the minimal resources of the mosaic, they will fare just as well under any richer ontology.

---

that laws govern, since I accept NP. However, I suspect this form of governance will not be enough to satisfy everyone's conception.

3 The logical form of laws

This chapter explores the logical form of laws which will help us to put some constraints on and frame questions for what's to follow. §3.1 argues that laws are conditionals with a particular form (the vast majority is published in Friend 2016). One aspect of this conditional form is that the antecedent predicate denotes a 'system-type'. I subsequently move on, in §3.2, to discuss the nature of system-types and their relation to laws. Another aspect of laws' logical form is their universality. In §3.3 I discuss a number of ways to save the truth of laws interpreted as universally quantified by what I take as a traditional modification of the 'system-predicate' and my own preferred method of modifying the consequent 'behaviour-predicate'. I end in §3.4 with a discussion of what further work needs to be done to get a philosophical account of laws given their logical form.

3.1 Laws are conditionals

An often assumed, but rarely argued for, view of laws of nature is that they are quantified conditionals. The ubiquitous schema '*All Fs are Gs*' dominates much philosophical discussion on laws but rarely is it shown how actual laws mentioned and used in science are supposed to fit it. Instead, what can seem embarrassingly toy examples like 'all ravens are black' are employed for discussion. Ignoring the complexity of real cases has some value, but it also comes at a price, and today there is increasing literature arguing that laws have been misrepresented: either we have been wrong about their logical form or else we are misguided in supposing them to be important features of science (see, e.g. Cartwright 1983, Dupré 1993, Giere 1988, 1995, Mumford 2004, Maudlin 2007).

My project in this section will be to offer some argument for the view that laws do indeed have a quantified conditional form. Throughout the discussion, I will be drawing on examples from science. Picking examples is a tricky business. Philosophers tend to have their favourite laws (often either toy examples or else the most strict laws of physics) and disregard many of the statements called 'law' across the whole of the sciences. My approach is consciously opposed to this: cast the net wide and remove unwanted catch afterwards. For case-studies, I attempt to pick a wide variety of statements referred to as 'law' from a variety of sciences. Philosophers (and some scientists) may balk at some of the examples, if they are obviously ridden with counter-instances, not general enough, not from an austere enough science, etc. While I am happy to accept that a full account of laws may see fit to disregard some of my examples my feeling is that to make such a move before an exploration of logical form has returned its conclusions would be arbitrary or a result of unjustified bias.

Having said that, there are clearly some statements we call laws which scientists themselves would disregard as important objects of philosophical or scientific enquiry. For instance,

> **Bode's law:** All planets 1 to m orbit the sun in an ellipse with a semi-major axis $\alpha = 0.3 \times 2^n + 0.4$ astronomical units, where $n = -\infty$ when $m = 1$ and $n = m - 2$ for all $m > 1$.

Bode's law is largely discredited from capturing any generality worth investigating. First, it is false. Many planets do not satisfy its prediction; second, there are no known plausible explanations as to why such a law would hold even approximately. While I take it that there may be numerous such statements in science, I will assume its fairly clear which they are. In what follows I will choose examples partly based on what scientists call laws, but also filtering for those which are actually employed in scientific practice. Bode's law, so far as I know, is not employed in that capacity and never was to any significant degree.

A separate issue regarding picking examples concerns representational phenomena in science which are not accorded the honorific 'law' but nonetheless seem, philosophically speaking, to be of similar interest, e.g. mathematical models, principles, rules, theorems, equations basic to scientific theory. One telling example is the Schrödinger wave-equation, which I argue below is associated closely with a law. In §3.2.2 I will also discuss specifically the notion of models in science and their connection with laws. I leave it as implicit in what follows that similar remarks can be carried over to many other representational phenomena.

Laws have long been described in philosophical literature as conditionals stating that if anything has some quality, then it has another. Typically, this is more formally represented by the following famous schema.

$$\forall x(Fx \to Gx) \qquad \text{(Schema 1)}$$

Some past employments of Schema 1 can be found in Hempel and Oppenheim 1948, Sellars 1948, Kneale 1950, Ayer 1999 [1956], Fodor 1991, Harré 1993; more recently, the schema can be found in, e.g., Cohen and Callender 2009, Mitchell 2000, Nickel 2010, Beebee 2011, Carroll 2012.

Schema 1 characterises the logical form of laws as universally quantified conditionals where F and G are replaced by predicate clauses. Traditionally, the domain is understood broadly and the conditional '$\to$' is taken to be *material.* By arguing that we should schematise laws in the form of Schema 1 that is not to deny that there are distinct logical forms which are equivalent to them (e.g. that obtained by conjoining a law with a logical truth). What follows is a discussion of the appropriateness of Schema 1 as a natural way of rendering laws which reveals the particular type of inference they license (i.e. from antecedent to consequent).

Today Schema 1 is often the go-to shorthand for philosophers' general discussion concerning laws' logical form, purporting to capture the most famous of examples of pseudo-law as 'all ravens are black' and 'all emeralds are green'. But though ubiquitous, the schema has become the focus of ridicule from scientists and more practice-focused philosophers of science. After all, the pseudo-laws just provided are not laws in any sense in which a working scientist would consider and Schema 1 can seem at best an oversimplified philosopher's caricature.

Along these lines, Maudlin (2007) has complained that,

> the fundamental law of Newtonian mechanics, the mathematical consequence of Newton's first two laws, is $\mathbf{F} = m\mathbf{a}$ or $F = md^2\mathbf{x}/dt^2$ or, most precisely, $\mathbf{F} = md(m\mathbf{v})/dt$ [sic[1]]. The fundamental law of non-relativistic quantum mechanics, Schrödinger's equation, is $i\hbar\partial\Psi/\partial t = \hat{H}\Psi$. No doubt these can be tortured into a form similar to $\forall x(Fx \rightarrow Gx)$, but it is hard to see what the purpose of the exercise would be. (2007, 11)

Certainly, it is not immediately obvious why we should expect equations—the mainstay of representations of law in many areas of science—to be representing information most appropriately captured by Schema 1. But if I am right, there must be some way to do so. Furthermore, if the defence is to count for a significant proportion of laws it had better not be too torturous a procedure.

[1]There are too many variables for mass here. Presumably Maudlin intended $F = d(m\mathbf{v})/dt$.

It is clear, however, that there has been almost no justification of the repeated use of this schema to represent laws in the philosophy of science literature, and since ubiquity doesn't entail correctness, the time has come to provide some.

There are a number of features of Schema 1 which require defending. One feature of logical renderings of laws in general which has invited a lot of discussion and criticism is their universality or 'strictness' (see, for a brief sample, Anscombe 1971, Cartwright 1979, 1989, Harré 1993, Giere 1995, Hüttemann 2007). As a consequence, the use of the universal quantifier '∀' in Schema 1 is a legitimate point of concern (see Fodor 1991, Pietroski and Rey 1995, Schurz 2002, Spohn 2002, Nickel 2010, Reutlinger 2011, Hüttemann 2014, for some responses). In §3.3 I will have something to say about this. The immediate task, however, is to argue for what comes after the quantifier: the conditional with antecedent and consequent predications of the same variable.

My argument for the conditional feature of the schema of laws aims to show that a conditional is implicit in our very understanding of laws even when they do not appear to explicitly have that form. But before providing such an argument, I must pause again to further qualify the project. For even granting that laws are conditionals, we may yet wonder what conditionals they are. Typically, when laws have been represented with Schema 1 the conditional is interpreted as *material*. Hence, the conditional is truth-functional: it is true so long as nothing which satisfies the antecedent also fails to satisfy the consequent. This interpretation has raised concern even with those sold on the conditionality of laws. When the '$\rightarrow$' is understood as a material conditional, schematising laws with Schema 1 renders them structurally equivalent to trivial universal conditionals which are true because nothing satisfies the antecedent (e.g. 'all unicorns are purple') or true because everything satisfies the consequent (e.g. 'all electrons are actual'). Furthermore, all material conditionals are logically equivalent to their contraposition, which can intuitively seem less worthy of the title 'law' (see Hempel's 1945 ravens paradox). These characteristics can seem to indicate the material conditional as being too liberal to capture the form of laws of nature.

Other interpretations of the conditional in lawlike statements have been provided which are more restrictive. Some have treated laws as inferences about kinds (proposed in Lowe 2009), or as inferences about abstract models (proposed in Giere 1988), or as an assertion of the existence of a relation of necessitation between two first-order universals (as we have seen, Armstrong 1983, understood laws to be *identified* with such existences). There have also been more logically exotic suggestions that laws should be understood as non-truth-functional conditionals (along the lines suggested by Stalnaker 1968) or under a suppositional theory (suggested, e.g., in Mackie 1973, ch.4). Still other renderings employ an even *more* liberal conditional than the material conditional, such as those which treat laws as generics in which the inference from antecedent to consequent is qualified to hold only in 'normal' circumstances (e.g. Schurz 2002, Drewery 2006, Nickel 2010).

I have yet to be persuaded that the material conditional is not appropriate for laws, although I will refrain from arguing this here. For those of us who believe laws to have the form of *some* conditional or other, a more pressing concern is to justify *that*. Of course, it might be the opinion of some who believe that laws are conditionals that there is no way to show this until we have the logic exactly right. I disagree, and although I will write with the material conditional in mind, I intend my comments to be support the conclusion that laws are conditionals regardless of which particular version of conditional this may turn out to be.

3.1.1 The look of laws

Certainly *some* of the statements titled 'law' in science very clearly have the form of a quantified conditional. Take the following examples (many of my examples throughout are drawn from Lange 2000, 30-33),

> **Farr's law:** All epidemics follow a positive, symmetrical curve of infected population over time.

> **Werner's law:** Geological strata are ordered with the older under the newer.

> **Wallace's law:** Every species has come into existence coincident both in time and space with a preexisting closely allied species.

Although none of the above expressions contain explicit 'if' and 'then' clauses, they can clearly be understood as telling us that for all things, if they are of a certain type then they will be a certain way, i.e. there is an antecedent classificatory predication of the kind of system the law is about (e.g. an epidemic, a stratum, a species), and a consequent predication of the kind of behaviour (understood broadly) systems of the antecedent type are expected to exhibit (e.g. a certain dynamics of size over time, a directionality with respect to the centre of the earth, a relationship with other species at certain times in the species' history). They can thus be rendered into the form of Schema 1, e.g.,

$$\forall x(\text{Geological-stratum}(x) \rightarrow \text{not-inverted}(x))$$

Despite laws like the above, which clearly are available to a conditional interpretation, many do not look, at first glance, to have this form at all. The above quote from Maudlin draws our attention to two laws often represented by equations. Consider also the following (allowing the variables to be defined in their usual way):

Equation for Ideal gas law: $PV = nRT.$

Equation for Newton's law of gravitation: $\mathbf{F} = G\frac{Mm}{r^2}.$

Equation for Ohm's law: $V = IR.$

Equation for Yoda's power law: $W = C\rho^{-3/2}.$

Gauss's law: $\oiint_S \mathbf{D}.dA = \iiint_V \rho.dv$

Moreover, there are laws which are neither characteristically represented by equations nor in the form of a conditional, such as,

> **Gresham's law:** Overvalued currency replaces undervalued currency in economic circulation.

> **Buys Ballot's law:** Low pressure in the northern [southern] hemisphere is on the left [right] hand when facing downwind.

> **Faraday's first law of electrolysis:** The mass of an element liberated or deposited on entering into reaction at an electrode during electrolysis is proportional to the quantity of electricity passing through it.

We may feel uncomfortable admitting all the above into a philosophical discussion of laws of nature. After all, Gresham's law seems dangerously close to counting as a 'causal law' in Cartwright's (1979) sense; this would be more perspicuous if, as is often the case, 'drives out' is substituted for 'replaces'.[2] Also, Buys Ballot's law might seem unduly singular to count as

[2] I assume throughout this work that the laws under enquiry are more appropriately captured by Cartright's contrasting category 'laws of association', which 'tell how two qualities or quantities are co-associated' (1979, 419).

an instance of a law, if considered implicitly restricted in application to Earth. Moreover, we might expect that many of the laws above admit of a mechanistic explanation or reduction to simpler laws. For instance, we explain the Buys Ballot law in terms of the Coriolis effect. It is conceivable that certain metaphysical motivations might lead one to demand the full generality and non-derivability of laws, in which case one could complain that some or all the above statements fail to count as proper *Laws of Nature*. Nonetheless, all these generalisations do seem earnestly called 'law' by scientists working in their respective fields (they are not like Bode's law in that respect). Consequently, the project will be to see how far we get without taking action on the basis of such complaints. I will assume that the above examples are all as good examples of scientific law as each other, although I will be arguing that they are represented in a misleading way.

If the reader is happy to follow my lead in treating the above non-conditionally represented statements as *prima facie* examples of laws, why should we think that laws ultimately have the form of a conditional? My argument is really very simple. These laws describe how things behave. But none describe how *everything* behaves, so the characterisation of behaviour must be restricted to some antecedent 'system-type'. Hence the conditional. While I take this brief summary to sound plausible up to a point, it clearly needs some more developed exegesis. The argument is perhaps most forcefully applied to those laws typically represented by (and misinterpreted as) equations. So I recommend we start by looking there.

3.1.2 Equations and their laws

What do the symbols either side of the '=' sign in the ideal gas law represent? If the symbols were interpreted as naming determinate properties, the expression might appear like an ill-formed identity (since identities are dyadic) rather than an equation. Anyway, it is hard to conceive how the gas constant R could be a property, after all what kind of a thing would instantiate it? P, V, n and

T are not obviously determinate properties either, because they are *variables* taking any non-negative real number as their values. However, the meaning of the gas law cannot be simply the algebraic statement that the product of any two numbers equals the product of a constant multiplied by two other numbers—that is a trivial truth of mathematics! And we know this cannot be the case since the variables in the ideal gas law, unlike purely mathematical equations, are associated with scientific units, e.g. Newtons-per-meter-squared (Nm^{-2}) for pressure, and degrees-Kelvin (K) for temperature.

A step in the right direction is to understand $PV = nRT$ as describing a group of functional relationships over *variable*-properties understood as sets of possible values for pressure, volume, quantity and temperature. For example, one function included in the group would be that which takes any value in a set of pressures as input and outputs some value in a set of temperatures. But this cannot be the whole story because functions, unlike laws, are not the sorts of things which can be truth-evaluable: a function is an operation not a statement. We *could* understand the equation as indicating a statement of the form 'there exists such a function between sets of pressures and sets of temperatures with the following results', but this again is an uninteresting fact, since there exist infinite such functions. If laws were the equations themselves, it is hard to see either what they represent or how they could be useful.

What is implicit in the formulation of the laws expressed as equations is the added clause that the group of functional relationships described by them are *true of things*. The set of temperatures, pressures, volumes and quantities of substance are those belonging to one and the same entity *to which the equation may be applied*. And the relationships described don't just exist in some abstract sense, but reflect the relations between the actual specific pressure, volume, temperature and quantity of such an entity at some moment in time. This suggests that the information bound up in equation $PV = nRT$ is something to be predicated of an entity, something an entity can satisfy or not. In general, we may understand equations to indicate a certain high-order behavioural property which entities have when and only when a specific group of relations determined by the equation holds between various (usually quantifiable) lower-order variable-properties of that entity.

In the case of the ideal gas equation the relevant behavioural property is instantiated by something so long as the functions described by the equation map from actually instantiated variable-properties to actually instantiated variable-properties. Consider the pressure-function $P(x, t)$, the volume-function $V(x, t)$, the temperature-function $T(x, t)$ and the quantity-function $n(x, t)$, each of which take as input an entity and time and outputs a real number. This output defines the determinate value of the respective variable-properties pressure, volume, temperature and quantity of substance. The relevant higher-order behavioural predicate for the ideal gas law would be one which is satisfied by entity g at t_1 *only if* the compound product-function $P{\times}V$ which takes as inputs the outputs of the $P(x, t)$ and $V(x, t)$ functions and outputs the same value when $x = g$ and $t = t_1$ as does the compound product-function $n{\times}R{\times}T$ which takes as inputs the outputs of the $T(x, t)$, $n(x, t)$ functions and the constant R.

So, the equations employed to stand for laws are informative of the behaviour *of things*. But laws cannot have the form 'Bx' where x is a variable ranging over entities and B is a behavioural predicate indicating that whatever satisfies it bears certain internal relations, for this is simply ungrammatical. Laws are supposed to be *general*, so a plausible binding of the variable is with the universal quantifier: $\forall x(Bx)$. However, it is certainly not the case that scientists use the ideal gas law, or Newton's, Ohm's or Yoda's law to say that *everything* bears these relationships between their properties. Concerns of absolute generality aside, to be the sort of thing which can have its components related in the way described by the behavioural predicate, that thing must at least *have* instances of those property-types, and many entities don't have any of them. After all, what's the pressure of a government, or the volume of a thought? But even when granting only the systems which have the relevant property-instances, the equations (or other formulae standing for a set of functional relationships) described by the behavioural predicates are not indiscriminately attributed by scientists to all such systems. An apple has an internal pressure, volume, quantity of substance and a temperature, but we don't expect to use the equation $PV = nRT$ to describe its behaviour. Similarly, a lamppost has weight and density, but one doesn't ask what the respective growth-constant is in order to predict its behaviour with the equation in Yoda's power law.

Assuming we want to preserve truth (*pace* Bach 1994) as best we can while still retaining generality, laws like the ideal gas law, Ohm's law and Yoda's power law will need to be captured by a more qualified schema than $\forall x(Bx)$, given a relatively broad domain. Note that this is a different problem from that which arises when we consider counter-instances to the law. One major complaint with treating laws as universally quantified is that many of them would thereby appear falsified by the existence of counter-instances. I have eschewed finding a solution to this problem in the present discussion (this is the subject of §3.3). However, the present issue does not concern these kinds of counter-instances. Countries, thoughts and apples are not counter-instances to the ideal gas law, since they were never even in the running for candidate systems to be described by it. Indeed, we might understand the issue presently under discussion as that of explaining *why* these examples are not in the running for being counter-instances to the ideal gas law (in contrast to a particular real gas in Boyle's air-pump).

There appear to be two plausible options to qualify the general schema in order to make sense of this fact. One option is to understand the domain quantified over by the universal quantifier as restricted to only those things scientists are wont to ascribe the behavioural predicate to. In the case of the ideal gas law, for instance, the restriction would be to ideal gases (or those gases which may be appropriately modelled as an ideal gas). Such a response would be similar to that often made in order to make sense of utterances like that of Yogi Berra when he said that "nobody goes there anymore, it's too crowded." Without some suitable distinction in domain restriction of Yogi's first clause from that of the second his utterance appears contradictory.

I do not see why domain restriction to less-than-universal contexts should be prohibited wholesale from the expression of laws. As Maudlin (2007, 12) points out, it does not appear to be contradictory to assert that some law describes physical states *around here*, or *for the last 10 billion years*, and we might deem it appropriate to attribute such restrictions to limitations on the domain rather than any explicit antecedent condition. However, my present criticism is with the idea that domain restriction is *sufficient* to limit application of the behaviour-predicate. I will offer a few suggestive points to think not.

First, laws are supposed to be re-statable across a variety of contexts and, assuming they don't change (see Lange 2008, for a discussion), across time too. Laws are repeated in textbooks and science classes, they are derived, explained, invoked for calculation and written down by students. But further restrictions on the domain of the sort we have been considering are typically considered to be determined by contextual parameters (Stanley and Szabo 2000). For example, if the domain of quantification is restricted more in an utterance of "everything is gone" when uttered on return to my burgled home than when uttered at the end of the universe, then the extra restriction (e.g. to the things in my house or in my jewelry box) will be provided by the context rather than by linguistic features of the utterance. Laws are not the sort of expressions which should be susceptible to so much influence from context. An electrician, for instance, wants Ohm's law to be about the same things from one situation to another. Otherwise, it is hard to see why she should hope to apply all her understanding from case to case.

Second, it simply does not seem to be the case that laws are expressed in such a way that they ascribe a certain behaviour to everything and expect domain restriction to do all the work of cutting out unwanted denotation.

Invariably, when the terms P, V, n, and T are defined, in relation to an application of the equation $PV = nRT$, they are defined, respectively, as the pressure/volume/quantity-of-substance/temperature *of a specific gas in a cavity*. In other words, the terms represent variable-properties of a certain type of thing. Similarly, when the terms in $V = IR$ are defined, V is defined as the potential difference measured across a conductor, I as the current through the same conductor, and R as the resistance of that conductor. Here, the variable-properties are of *one and the same electrical conductor*.

The explicit reference to what type of thing the variables are supposed to be of suggests another way to qualify the schema. This option is to keep the domain broad and the context open, but to condition the ascription of behaviour on the satisfaction of some further predicate S, e.g. 'being an ideal gas' or 'being an electrical conductor'. Such a predicate would pick out a kind of system from which the behaviour can be lawfully inferred. To take this option is to understand laws as conditionals of the form $\forall x(Sx \rightarrow Bx)$. For instance, under the proposal, what is missing in the ideal gas law as an unrestricted generalised predication of behaviour is the added clause that the relationships entailed by the equation in the behavioural predicate *exist between the pressure, volume and temperature of any ideal gas*, i.e.,

> **Ideal gas law*:** For all things, *if* a system is an ideal gas in a cavity *then* its temperature, pressure, quantity of substance and its volume will satisfy the equation: $PV = nRT$

Note that being an ideal gas in a cavity is not simply identical to satisfying the equation $PV = nRT$, since presumably other assemblies could satisfy it as well, given peculiar enough circumstances. In §3.2.1 I will suggest what exactly the identity criteria are for types of system like ideal gases. But it should be clear already that a proper understanding of the utility of the expression '$PV = nRT$' requires us to understand it in the context of *Ideal gas law**, i.e. a conditional. The law predicates a type of behaviour of something if we can predicate something else of that thing: that it is an instance of a certain type of system.

Similar reasoning goes for the other equations often passed off as themselves laws. Although, as Maudlin says, it would be torturous to put the equations into the form of Schema 1, this does not mean that laws themselves don't naturally take this form. For instance, Ohm's law of resistance is not an equation, but a conditional which states that all components with a constant value of resistance in an electronic circuit will experience a voltage-loss and current distributed in accordance with the equation $V = IR$; similarly, Yoda's power law is not an equation but a conditional which states that all plant seedlings in a single sewn area will grow with their properties of dry-weight and density of sewn seeds being related by the equation $W = C\rho^{-3/2}$. Newton's law of gravitation is a conditional which states that all two-mass systems will experience gravitational components of force proportional to the product of each other's mass according to the equation $\mathbf{F} = G\frac{Mm}{r^2}$. Gauss's law is a conditional stating that all closed surfaces are such that the flux density across them ($\oiint_S \mathbf{D}.dA$) is equal to the charge contained within ($\iiint_V \rho.dv$).

The discussion of laws represented by (and often interpreted as) equations has resulted in a number of important conclusions. First, to understand an equation as something scientifically informative, it is necessary to interpret it as describing a set of functional criteria relating something's physical properties. This observation suggests the idea of a 'behaviour-predicate' corresponding to the equation which is defined to be satisfied by an entity if and only if its physical variable-properties are related according to that equation. Second, laws' generality demands that they make a general claim about the behaviour of things, but since it is plainly false that everything satisfies the behaviour-predicates corresponding to all the equations considered so far, the predication has to be qualified. I have suggested that the most plausible way to do this is by conditioning on what we might call the 'system-predicate', which is satisfied if and only if an entity is of an appropriate type. In order to highlight the importance of these two types of predicate and their characteristics I substitute the traditional schematic letters F and G in Schema 1 for S and B, respectively, i.e.,

$$\forall x(Sx \rightarrow Bx) \qquad \text{(Schema 2)}$$

Since the discussion made no essential assumptions specific to the choice of equations used, we appear to have an argument that laws represented by equations are in fact more appropriately rendered in the form of a conditional in the form Schema 2. For most of the rest of this thesis I will focus exclusively on laws whose behaviour-predicates are characterised by an equation, such as those mentioned here. Shortly I will extend this reasoning to provide more reason to think all laws *must be* conditionals. But before that it's worth pointing out how similar reasoning to that presented above carries over to the other recalcitrant laws mentioned earlier.

3.1.3 Miscellaneous laws

We now move on to those laws which superficially, at least, do not appear to be conditionals or equations. The first thing to notice is that the examples given for this 'miscellaneous' group of laws assert the existence of some relationship between variables. Gresham's law relates the variable of quality of currency to time; Buys Ballot's law relates height and direction of pressure with location in hemisphere and direction of wind; Faraday's first electrolysis law relates the mass of a liberated element with a quantity of electricity. Whilst not *equations* as such, these laws are nonetheless similar in that they describe functional relationships.

Similarly to the argument for equations, we can note that these functional relationships will not be understood properly unless they are interpreted as relating the actual instances of various properties by individual entities. Therefore, we can infer a high-order behaviour-predicate for each putative law satisfied by an entity if and only if the relation defined in the functional relation holds between the lower-order variable-properties which feature in that relation. For example, the above expression of Gresham's law indicates a high-order predicate 'is such that its quantity of undervalued money in circulation is negatively correlated with its quantity of overvalued money in circulation and the quantity of overvalued money is increasing'.

The next step is to notice that a full explication of the law would not predicate such a behavioural property of everything, but rather on some specific system-type. For Gresham's law, the relevant system-type is arguably a closed economy. Consequently, a more explicit rendering of Gresham's law would be,

> **Gresham's law*:** All closed economies are such that their quantity of undervalued money in circulation is negatively correlated with their quantity of overvalued money in circulation and the quantity of overvalued money is increasing.

Similarly, for the Buys Ballot's law, the relevant system-type might possibly be the Earth itself; for Faraday's first electrolysis law, it would be an assembly comprising an electrode and electrolyte. Therefore, the same argument as above works to show the implicit conditional in all the considered laws above.

3.1.4 Why all laws are conditionals

I have just provided an argument for understanding some of our laws as conditionals. There appeared to be a straightforward way to interpret conditionally the ideal gas law, Ohm's law, Yoda's power law, Gresham's law, Buys Ballot's law and Faraday's first law of electrolysis (among other laws). The argument relied on the understanding that the behaviour described by a formula (such as a numerical equation) by which those laws are typically represented is not a behaviour attributed to everything in a relatively unrestricted domain, but rather to a specific class of systems (the ideal gases, the electrical conductors, etc.) which the behaviour is antecedently conditioned on. This style of argument works for individual cases. However, besides inductive motivation, it does not provide conclusive reason to think that *every* law need incorporate such conditioning.

Indeed, we might think that some of the most austere behavioural formulae do not need conditionalising in this way, since they really do apply to everything. Consider the physicalist view that physics consists in a project to determine those formulae which most accurately describe the behaviour of every single system in existence by some basic formulae and generalisations. Perhaps the Schrödinger's time-dependent wave-equation is expected to be among those formulae.

$$i\hbar\partial\Psi/\partial t = \hat{H}\Psi$$

Schrödinger's equation seems in many respects a paradigm representation of a law (see for example, Maudlin 2007, 11).[3] Assuming that there is indeed a law associated with Schrödinger's equation, we might follow the reasoning made above insofar as we recognise the equation to represent some information concerning the behaviour of things, i.e. we identify a behaviour-predicate such as 'is such that its Hamiltonian is related to the time-derivative of its wave-function according to the equation $i\hbar\partial\Psi/\partial t = \hat{H}\Psi$.' Let's abbreviate this S_t. Now, the physicalist might believe that there there is a non-conditional law,

$$\forall x(S_t x).$$

[3]That is, despite the fact that neither the equation nor the information it represents are typically referred to as 'a law' in scientific practice. While I expect that it is relatively harmless to use the equation as an example, some have demanded that we take a closer look at the practice of scientific denotation of the title 'law' (e.g. Mumford 2004, esp. p.129).

While such a generalisation may in fact be true (although only if physicalism is true), I argue that it is highly implausible as a rendering of any statement likely to be called a law given the current state of physics. I offer four points.

First, it is well known that there are technical difficulties extending current analysis involving the wave-equation (quantum field theory, quantum chromodynamics, etc.) to large-scale entities. Solutions to the Schrödinger equation are just about manageable for Hydrogen and Helium, but even for the immediately more massive atoms, computational methods are required. When we come to consider small molecules or anything more complex, approximation methods like that from Born and Oppenheimer are required in order to make useful predictions or explanations. So, regardless of whether this is considered as 'mere computational intractability' or not, the generalisation $\forall x(S_t x)$ certainly doesn't reflect the *practice* of contemporary physics. But more importantly, if there is any reason at all to doubt that all macromolecules behave in the way dictated by the S_t predicate, any scientist averse to unnecessarily bold conjectures is going to hedge their bets with regard to the application of the Schrödinger equation. After all, why commit to saying the S_t predicate applies to everything when one can be far more confident of a weaker claim at no cost of utility: that S_t is satisfied by those entities which are composed only of relatively few standard-model particles? Until our numerical capabilities are developed enough to derive and test solutions to the Schrödinger equation (in its unapproximated form) for macromolecules, there is no *scientific* reason (supported and motivated predominantly by empirical evidence) to believe a generalisation like $\forall x(S_t x)$—although we may believe such a law for other (philosophical) reasons.[4]

Second, relativistic quantum mechanics may include the application of

[4]The view expressed here is not meant to be as sceptical as that of Cartwright's (e.g. 1999, 2). Where we are able to apply classical physics more straightforwardly than quantum physics, I think we may still have scientific reason to believe the quantum theory to be in some sense the more correct.

Schrödinger's equation to contexts of special relativity, but not *general* relativity; the latter remains an as yet unresolved problem for physicists (the tensor fields described in Einstein's field equations imply definite values for both position and location of objects whereas the Heisenberg uncertainty principle implies such precision is impossible). As with the problem of computational intractability we may believe a reconciliation between the two theories is possible. But we must also admit that there is no current scientific support for laws which entail a reconciliation. Again, why should a scientist commit themselves to the S_t predicate applying to everything when they can be more confident of a weaker claim at no cost of utility? One might believe the predicate applies universally, but such a belief will not be justified by scientific discovery alone.

Third, even if some reconciliation can be achieved and physics becomes united under one body of theory, we would still be far from justified on the basis of physics alone in our commitment to the view that the theory applied to *everything*. The project of physics is to describe the behaviour of the *physical* world: plausibly, that which has location, size, momentum, charge. It is an open question what fits into this ontology; presumably the entities of the standard model and the Higgs particle are there, but what about fields and gravitons? Whatever we decide, the truth of physics is not hostage to physicalism. The truth of the standard model, for instance, is not thought to stand or fall on the debate over platonism in mathematics or mind-body dualism; its truth is independent of these debates. Physicists' laws and theories must, therefore, include a restriction of some kind to avoid purporting to describe the behaviour of mathematical objects, incorporeal minds, or indeed any type of entity mooted to be irreducible to physical stuff. Perhaps a domain restriction might be plausible at this stage to exclude some of the obviously non-physical stuff. But again, why render laws as something so liable to contextual influence when an explicit restriction on system-type would make it instantly more general? Once more, the point here is not that physicalism is false, nor that we have no good reasons to believe it. The point is that we make no explanatory gains by writing unconditional universality into the conjectures of physics. Physics provides some behavioural claims about a particular class of entities, it is up to us as philosophers, biologists, policy-makers, etc. to decide whether or not anything falls outside this class.

Fourth, it's not even clear to me that a universally accepted physicalism would lead us to consider the fundamental laws as unconditionally universal. Physicalism is typically understood as a claim about the sufficiency of physics (e.g. Davidson 1995, Papineau 2001). But it is not typically understood as incompatible with an unrestricted mereology, and not all objects which exist without such a restriction will be characterised well by the physical laws. For instance, call the object comprising the Earth up to some specific time and Proxima Centuri for all time thereafter *Earturi.* If the laws of physics were thought to apply to it, it would be a counterexample to many, e.g. conservation of mass/energy, relativistic limits, Schrödinger wave evolution, etc. So if $\forall x(S_t x)$ is to be a law, physicalism needs not only to be an established part of scientific theorising, but it needs to be conjoined with further highly philosophically controversial theses, such as the rejection of unrestricted mereology. Once more, even if these controversial views are correct or have strong philosophical motivation, it does not seem to be part of *scientific* reasoning that these views are true. Since science seems to be the best indication of what to call a law, we should be careful not to let philosophical biases infect those ascriptions.

I take these points to show that Schrödinger's equation, along with any other prominent formulae in physics, is not applied unconditionally, but rather only to a specific class of systems. It may prove difficult to provide necessary and sufficient conditions for such a system, but the *practice* of quantum physics reveals at least that the class is not understood to include general relativistic systems, mental states or platonic forms. It *is* known to contain systems which can be predicted and explained (potentially including those reasonably approximated; see §3.3) by the formula $i\hbar\partial\Psi/\partial t = \hat{H}\Psi$. Let '$Q$' denote the property of being in such a class of systems. We then get the following expression for the lawlike knowledge physicists have relevant to applications and interpretations of the Schrödinger equation—what we might call 'Schrödinger's law'.

> **Schrödinger's law:** $\forall x(Q \rightarrow S_t)$

This reasoning significantly strengthens the argument that all laws are conditionals, specifically of the form Schema 2. From here on I will assume this is the correct schema for all laws. Of course, a discussion of laws' logical form is still only a portion of the story of them. To this end, I turn now to discuss in more detail the nature of the subject of laws, viz. *systems*.

3.2 Laws are system-laws

This section aims to elaborate further on the content and scope of laws centering on the idea that all laws are 'system-laws'. Discussion on the logic of laws of nature has often seen it appropriate to distinguish 'system-laws' from other laws. Schurz, for example, distinguishes 'system laws' from,

> those fundamental laws of physics which are not restricted to any special kinds of systems (be it by an explicit antecedent condition or an implicit application constraint). (2002, 367)

This distinction has the potential for at least two interpretations. According to one, system-laws are so-called if they have an antecedent condition which contains a singular term, i.e. a term which picks out an individual rather than a class of individuals. It can seem to be this reading Schurz has in mind when he says shortly after the previous quote that '*system laws* [...] refer to *particular* systems of a certain kind in a *certain time interval* Δt' (my emphasis). It is in this sense that generalisations found in biology which appear to reference specific times and space (e.g. Earth over the last 3.8 billion years) have been distinguished from those more general laws of nature (Beatty 1995, Rosenberg 2001). Schurz also mentions Kepler's laws and Galileo's law of free fall, both of which condition on the system of application being, respectively, our solar system and the Earth. (We might think that Buys Ballot's law similarly counts as a system-law for the same reason.)

If this is the correct interpretation, then discussion in 3.1 on laws' conditionality is not directly relevant. Although the arguments provided there aimed to show that all laws have a conditional form they have not demanded that the antecedent incorporate a singular term, although this might be the case for some laws. However, according to a second interpretation, a system-law is so-called simply because it is a conditional *simpliciter*. This interpretation seems to match with Schurz's claim that the fundamental laws of physics are not system-laws because they are not 'restricted to any special kinds of systems (be it by an explicit antecedent condition or an implicit application constraint)' and his regarding the total force law and special force laws as not counting as system-laws 'at the cost of being *per se* not *applicable* to *real* systems, because they do not specify *which* forces are active' (Schurz's emphasis). If we read 'special' and 'real' in these quotes as implying that the relevant systems fall under some specific *class*, then we have reason to think of *any* law in the form of a conditional as a system law.

Under the second, broader interpretation, the argument presented above for the conditionality of laws entails that *all laws are system-laws*. The total force law (Newton's second law) and the special force laws (e.g. Newton's gravitational law) are (*pace* Schurz) system-laws in this sense since they express conditionals, both having the antecedent condition that the system is *massive*, or perhaps more appropriately *'Newtonian'*. The argument for these points draws on the fact that the behaviours these laws describe are not expected to be fully general. Instead, they condition behaviour on entities characterised by a particular body of theory (e.g. classical mechanics) not assumed by the practitioners who use them to capture everything in existence.

I will use the second interpretation of 'system-law' in what follows assuming that the arguments provided in §3.1 provide reason to think all laws are system-laws. But despite our ability to draw such a generalisation about laws, we still don't know much about what precisely a system or system-type is. In §3.2.1 I will argue for a criteria of identity for system-types in terms of conditions of assembly and in §3.2.2 I will argue that systems are real-world entities (as opposed to abstractions or fictions). This latter discussion will raise the problem of non-universality which will then be addressed in §3.3.

3.2.1 What's a system-type?

In this section I will develop a more precise understanding of what is meant by the notion of a system by providing a general criterion of identity for system-types. In what follows I first say what criteria of identity are. I will then suggest that we should individuate system-types by 'conditions of assembly', which are further divisible into 'conditions of componency' and 'conditions of organisation'.

Frege taught us that to be able to individuate something for the purpose of naming it, we should possess a 'criterion of identity' which would tell us for anything whatsoever whether it is the same as that thing. In his words,

> to use a symbol a to signify an object, we must have a criterion for deciding in all cases whether b is the same as a, even if it is not always in our power to apply this criterion. (1950, 73)

The exact lesson to take from this claim is controversial. We may wonder whether it reasonably applies to entities which fail to be individuated by sortal concepts (e.g. water, wine, and other mass-term denotees). Moreover, we may even wonder whether Frege's goal was really to help us individuate objects. For instance, Lowe has argued that,

> it is a misconception to suppose that a criterion of identity is in general a principle whose primary purpose is to be invoked in settling particular questions of identity or diversity concerning individuals: it is not [...] an *epistemic* or *heuristic* principle for the discovery of particular truths. Rather, [...] criteria of identity are *semantic* principles whose grasp is essential to an understanding of certain sorts of general term. (Lowe 1989, 13)

I cannot comment here on whether Lowe has accurately captured Frege's intention, as seems to have been his aim. However, I agree with Lowe that the criterion of identity can be put to service in elucidating semantic facts regarding general sortal terms. To use Frege's well-known example, we learn something informative about the concept of *direction* when we learn that if two things are directions, then they are identical if and only if their lines are parallel. Whether these semantic facts are something to be considered definitive (i.e. analytic) I will remain silent on. For what follows, I assume a criterion of identity for the denotation of a term at least elucidates the extension of the term.

We can utilise this feature of the criterion of identity to elucidate something informative about the concept of *system-type.* Drawing on Lowe's interpretation, we are looking for a way to fill out the ellipsis in the following statement:

> For all things, if they are both system-types, then they will be identical if and only if ...

So, what we are looking for are general features of system-types which elucidate the concept of system-type for us.

By asking for a criterion of identity for system-*types*, rather than systems, I am focusing on how we individuate kinds of systems rather than systems themselves. I think the question of what criteria we use to individuate any particular system may be about as difficult as the question of how we individuate individual persons. If I replace a single resistor in a circuit with one of the same value of resistance, do I still have the same token electrical circuit? If I break a covalent bond between two molecules in a gas, do I still have the same token gas? If Neptune spun out of its orbit around the sun, would Earth still be in the same solar system? These questions are interesting, but very difficult and, I suspect, contextually contingent. However, it appears that a more straightforward and perhaps more scientifically crucial question concerns whether something is an electrical circuit in the first place, rather than, e.g. a purely mechanical system or a bundle of electrical components in a bag. While it may be of limited interest to a theoretical chemist whether the adjustment of a carbon lattice from a graphite structure to a diamond structure preserves the same token system, it is crucial to know whether it is the same system-type. After all, it is the instantiation of a particular system-type which renders a system's behaviour describable by a particular law: if the system-type changes, so may the inferences one is able to make about how it will behave.

If we focus on types of system, it might seem that formulating a criterion of identity is easy: two system-types are identical if and only if they have the same instances. However, for practical purposes, we desire a criterion different from just a list of all the instances a system-type has. Scientists build systems, uncover them in nature, discover their behaviour and inner natures, test their boundary conditions, make predictions about where we will find them and explain facts based on their having those natures (more on this in §5.2.1 and §5.2.2). This suggests, I think, that scientists require knowledge of something more general about what systems must be like to instance a system-type.

If we are to avoid criteria composed merely of a list of instances, then the identity criteria for a system-type should make generalisations about what characteristics all and only their instances have. Note, that this demand for generality holds even if the scientific knowledge follows a pattern of explanation via exemplars. Perhaps scientists will explain an entity's behaviour by calling on its similarlity to an exemplary system which they can demonstrate indexically. Nonetheless, there will need to be *some* general characteristics shared between entity and exemplar in order to be able to say why the exemplar is even relevant in such a case.

So, what might these general characteristics of system-types be? One idea we can quickly reject is that the behaviour associated by law with a system-type features in those criteria. After all, if the behaviour did feature in the identity-criteria, laws would be completely trivial, conditioning some behavioural characteristic on entities defined by their having that very behaviour. Having said all that, it is certainly not the case that the behaviour of a system-type is irrelevant to our epistemic access to it. An important aspect of differentiating system-types from arbitrary classes of assemblies of objects with no obvious scientific interest surely concerns the behaviour engaged in (when law-abiding) by their instances. It is the behaviour of an electrical circuit when compared with the behaviour of a bunch of wires and power-sources thrown together in any gerrymandered formation that makes the former and not the latter of scientific interest (indeed, the specific value of laws' behavioural formulae is what forms the basis of the account of laws developed in §5). Moreover, it may in fact be the behaviour which allows us to keep attention on the system-types of interest. For example, it is the vapour-trail in a cloud chamber which keeps our attention on the system (a moving charged particle) which explains it. Arguably, it is in general what things do that allows us to be aware of them at all—a system-type with only causally inert instances could never be recognised let alone incorporated within a scientific theory.

But if behaviour is not to play a role in system-types' individuation, how *should* we individuate them? My suggestion is the following.

> For all things, if they are both system-types, then they will be identical if and only if *they have the same conditions of assembly.*

A condition of assembly is a requirement on how a system-type's instances are to be constructed, whether synthetically or organically. They divide into conditions on what types of *component* the assemblies which instance the system-type contain and conditions on how those components are *organised* in the assembly. Table 3.1 (starting p.92) lists some laws, their system-predicates, behaviour-predicates and the conditions of assembly for system-types.

Table 3.1: Laws and their parts

Law	System-type (S)	conditions of assembly		Behaviour (B)
		Components	Organisation	
Total force law	is an inertial body	x: inertial body y: net field of force	x is in y	has variable-properties satisfying, $\mathbf{F}(y) = m(x)\mathbf{a}(x)$
Gauss's law for Electricity	is a closed surface	x: surface	x is closed	has variable-properties satisfying, $\oiint_{S(x)} \mathbf{D}.dA = \iiint_{V(x)} \rho.dv$

Lorentz-force law	comprises a charge in a field	p: Electrical field; q: Magnetic field; r: point-charge	r is in p and q	has variable-properties satisfying, $\mathbf{F}(r) = q(\mathbf{E}(p) + \mathbf{v}(r) \times \mathbf{B}(q))$
Ohm's law	is an electrical conductor	x: conductive component y: source of charge z: voltage potential	y is in x z is across x	has variable-properties satisfying, $V(x) = I(x)R(x)$
Law of mutually gravitating bodies	comprises two mutually gravitating bodies	p is a massive object; q is a massive object;	p and q are mechanically isolated from all other objects	has variable-properties satisfying, $\mathbf{F}_{p,q} = Gm_p m_q / r_{p,q}^2$
Ideal gas law	comprises an ideal gas in a cavity	p is an Ideal gas; q is a cavity	p is materially isolated within q	has variable-properties satisfying, $P(p)V(q) = n(x)RT(p)$
1^{st} Law of thermodynamics	is a thermodynamic system	x is a collection of matter	x is closed	has variable-properties satisfying, $\Delta U(x) = Q(x) - W(x)$

Law of pendulum motion	is a classical pendulum	p is a massive bob; q is a light inextensible string; r is a frictionless pivot	system is in a uniform gravitational field; p is attached to one end of q; r is attached to the other end of q	has variable-properties satisfying, $T(x) = 2\pi\sqrt{\frac{l(q)}{g}}$
Snell's law	comprises electromagnetic rays across a medium boundary.	p is an EM ray; q is an EM ray; $r1$ is first medium; $r2$ is second medium;	$r1$ and $r2$ form boundary; p and q joined at boundary	has variable-properties satisfying, $\frac{\sin\theta(p,r1)}{\sin\theta(q,r2)} = \frac{n(r2)}{n(r1)}$
Compton shift law	comprises a photon-electron collision	p and r are photons; q is an electron	p incident on q; r emitted from q	has variable-properties satisfying, $\lambda(r) - \lambda(p) = \frac{h}{m(q)c}(1 - cos\theta)$

Archard's wear law	comprises a load on a surface.	p is a load; q is a surface; r is the softest of p or q;	q resists p normal to the load-force and perpendicular to the surface	has variable-properties satisfying, $Q(r)=\frac{KW(p)L(p,q)}{H(r)}$
Malthusian growth law	is a population of living organisms	p is a population of living organisms	the members of p cohabit a single environment	has variable-properties satisfying, $P(p,t) = P_0(p,t)e^{rt}$
Lotka-Volterra law	is a predator-prey population	p and q are distinct populations of living organisms	p and q occupy habitable environment	has variable-properties satisfying, $\frac{dp}{dt}=p(\alpha-\beta q)$, $\frac{dq}{dt}=-q(C-Dp)$

I will make use of the following schema for representing the criterion of identity for system-types (using $\sum_1^n X_i$ to indicate a conjunction X_1 & X_2 & ... & X_n),

$$\forall x\Big(Sx \leftrightarrow \Big(\sum_1^n C_i x \;\&\; \sum_1^n O_i x\Big)\Big), \qquad \text{(Schema 3)}$$

S is to be replaced by a system-predicate, $C_1x, C_2x, ..., C_nx$ by conditions of componency and $O_1x, O_2x, ..., O_nx$ by conditions of organisation.

For the sake of what follows, I presume the conditions of assembly for a system-type are simply definitive of that system-type. Hence, I will on occasion make use of explicitly expressions of laws in which the system-type is replaced by its conditions of assembly (this will prove particularly useful when drawing attention to potential causal relationships in §7). However, I am open to relaxing this connection if careful consideration of modal contexts suggests a weaker relationship. What's crucial is that the conditions of assembly accurately individuate system-types informatively *as a matter of fact*. One use this will have is for cases where conditions of assembly for some system-types are subsets of conditions of assembly for larger system-types. For instance, the system-type for the Lorentz force law arguably has conditions of assembly specifying that a system comprise a charge in a magnetic and electric field. One of these entities is a magnetic field, which satisfies the conditions of assembly for Gauss's law of magnetism. Hence, any assembly which satisfies the conditions of assembly for the system-type in the Lorentz force law will have a nested sub-assembly which satisfies the conditions of assembly for the system-type in Gauss's law of magnetism.

Requirements on components which feature in the conditions of assembly for a system-type may be obviously dispositional, e.g. being frictionless, being a predator, being inextensible, or else at least superficially categorical, e.g. being a photon (I explore the distinction between these properties further in §5.2.1 and §5.4.4). It may be that some assembly comprises objects which are fungible or perform more than one role in the system. For example, the resistors in series of a circuit can swap places without changing the system from being an instance of an electrical circuit. To put it anthopocentrically: the system-type does not care which component is doing what. What's crucial is that each role is performed by some suitable component.

That there are requirements on arrangement which feature in the conditions of assembly for a system-type should be obvious. As Cartwright says, 'a bunch of resistors and capcitors collected together in a paper bag will not conduct an electric current' (1999, 57). Requirements on arrangement are most naturally thought of in terms of the orientation and proximity of the various components—the components of an electrical circuit must be in contact across their terminals, the predator and prey population must be in the same habitat and capable of crossing paths, etc. It is hard to imagine requirements on arrangement which are not of this sort, but I don't think we should prohibit them out of hand. For one thing, the requirements on arrangement of instances of system-types which are sensitive to relativistic phenomena may require a more subtle characterisation rather than one of relative spatiotemporal orientation and location.

Often the conditions of assembly for a system-type are quite specific. All the examples in Table 3.1 specify a particular quantity of components systems need in order to fulfill certain tasks. However, some system-types mentioned in laws are not so specific. Consider, for instance, Dalton's law.

> **Dalton's law** If something is a mixture of non-interacting ideal gases, then the total pressure it exerts is equal to the sum of the partial pressures exerted by each individual gas.

This implies a functional relationship between the pressure exerted by distinct components (gases) of a system (mixture). But the system-predicate which features in this law's antecedent denotes a system-type whose conditions of assembly do not specify how many gases its instances should have. The law holds for mixtures of any number of gases. Similarly, the conditions say very little about the arrangement of the components, the gases must merely be 'mixed'.

In summary, the criterion of individuation I suggest for system-types does not draw on the behaviour associated with them by law. This criterion identifies system-types by their conditions of assembly which divide into conditions of componency and conditions of arrangement. By picking out individuative aspects of system-types which extend outside of the information provided in the respective laws for each system-type, we have established that system-types are conceptually rich in a distinct way from the inference to behaviour licensed by laws. Arguably it is this which licenses their explanatory power over the lawfully inferred behaviour which they exhibit (see §5.2.4 and §6.2.5).

The notion of a system-type so characterised might seem to resonate with concepts already present in the literature, in particular that of *mechanism* (Glennan 1996, Machamer et al. 2000, Bechtel and Abrahamsen 2005, Illari and Williamson 2013) and *nomological machine* Cartwright (1989, 1999). First, all three can be understood as comprising objects, sometimes called 'components' or 'entities' which may be defined dispositionally. Second, all three are partially characterised by their organisation or arrangement, which is understood to be partly explanatory of the behaviour witnessed in or by them. Third, all three can be 'nested' within further instances and can 'house' further instances within them. Fourth, all three are expected to behave in some functionally relational way. For nomological machines, the contained components will be mutually manifesting their capacities to affect each other in various ways; for mechanisms the contained components will be engaged in polyadic activities (or interactions/interactivities); and for system-types, the contained components will be engaging in behaviour often described quantitatively by equations.

There are also some crucial differences between system-types, on the one hand, and nomological machines and mechanisms, on the other. In one sense, system-types have stricter conditions than both nomological machines or mechanisms. The behaviour associated with types of system is restricted to that which can appear in the consequent of a law, while nomological machines and mechanisms are not in principle so restricted. Hence, toilet cisterns and synaptic transmission can count as instances of types of mechanisms and nomological machines but perhaps not instances of system-types. In another sense, system-types are less strict. Nomological machines have been characterised as needing to exhibit a certain amount of stability, and mechanisms have been characterised as being productive of a higher level phenomenon. Systems are not so characterised. Their criteria of identity comes from their conditions of assembly alone and the behaviour they are lawfully expected to exhibit is neither necessarily stable nor a phenomenon at some 'higher level'.

More significantly for our purposes, unlike a type of nomological machine or type of mechanism (at least under some interpretations), the behaviour of instances of a system-type is not part of its identity criteria. As was noted above, system-types' conditions of assembly are a matter of the particular components they have and their organisation. The activities or manifest behaviour of a system-type's instances is something expected given these conditions' satisfaction and it is a feature of laws that they capture this expectation. But the behaviour itself is not definitive of the system-type. This fact is, I believe, a significant merit of explanation in terms of systems and will be capitalised on in developing the causal conception of laws (see esp. §5.2.4).

Conditions of assembly are clearly important for saying what makes one system-type distinct from another. But it's worth pointing out what conditions of assembly do *not* do. Criteria of identity for system-types are not alone sufficient to sort system-types from mere kinds of assembly. With liberal enough scope, there will be conditions of assembly for any gerrymandered combination of objects in some specified arrangement. But it seems that system-types and their instances, if they are to be distinctive of law-like behaviour, should be rather more special. Something about the specification of organisation and componency conditions which constitute the conditions of assembly for the antecedent property in Kirchhoff's laws means that they supply the criteria of identity for a genuine system-type. But the specification that the same components be in a paper bag do not—there are no laws which have *being a bunch of resistors in a bag* as the antecedent condition (though there might be true generalisations which do). So far I have not provided any reason to say what the difference is, i.e. what makes system-types special. In particular, the criteria of identity for system-types does not seem sufficient to draw the distinction.

In what remains of this section I want to discuss two broad options for how to go about drawing a distinction between system-types and mere kinds of assembly. The first aims to find first some prior criteria for what makes an assembly of objects a system-type and then call a law any true (or true enough) generalisation which holds for them. The second aims to find first some prior criteria for what makes a conditional a law and then call a system-type whatever is denoted by the antecedent predicates of those conditionals. I will show why there are good reasons to opt for the second option.

Pursuing the first option, one could take a brute realist position concerning system-types: system-types are part of the world's natural kinds, thereby supporting inductive inference, counterfactuals, etc. On this approach we simply affirm that the kind of assembly, e.g., *pears in this basket* doesn't count as a genuine system-type and so the proposition that pears in this basket are sweet is not a law.

Tempting as this might be, it does not appear to be particularly informative. After all, it provides us with no criteria to assess borderline or contentious cases. Also, it seems to prohibit out of hand the possibility that there might be laws for non-natural assemblies. Laws of economics like Gresham's law, for instance, don't obviously seem to have system-types which can be classified as natural kinds. At least, the relevant system-type is not natural in any informative sense, other than that it is the kind which appears in the antecedent of a law. But if we are to be brute realists about what distinguishes system-types from mere assemblies, then it seems nature is about the only alternative mechanism to do it. We could, of course, deny that there really are laws of economics. But if we do that, I think we lose a grip on what a law is. Laws, including those of economics, are conditionals of a certain class which allow us to do certain useful things such as predict, explain and control the (natural or social) environment. One might divide the class of laws a in two, claiming that only a portion are real laws since only their antecedent predicates denote real natural kinds. But the question remains what it was about the 'unreal' law that made them useful in the same way? This is the question I think we should be concerned with and for that reason I reserve the term 'law' for that broader class.

Another approach still in line with the option of starting with an account of what makes system-types special is to find explicit criteria for their distinctive features. Some ideas, rejected by Hempel in a closely related discussion, are that properties like 'being a pear in this basket' or 'being a member of the Greenbury school board 1964' are prohibited from being system-types because they have too few (or unitary) instances, or that they mention specific objects or locations (pp.339–342). I agree with Hempel that these are poor methods of demarcation. Some genuine system-types aren't precisely instantiated by any instances, e.g. being an ideal gas in a cavity; and other genuine system-types very explicitly mention locations, e.g. being a freefalling body on Earth.

I don't presume to have engaged in an exhaustive investigation of potential ways one might determine what's distinctive of system-types prior to an account of laws. But we can certainly see a further general problem with any attempt to define the laws on the basis of such an approach. The problem is that even legitimate system-types can appear in conditionals with the right logical form, and yet those conditionals aren't laws. It might, for instance be that every electrical circuit is located this side of the Oort cloud. But despite having a genuine system-type in its antecedent, such a generalisation is no law. Or take a classic example: presumably *being bulk gold* is legitimate system-type if anything is, from which genuinely lawlike behaviour can be inferred. Nonetheless the generalisation that all bulk gold is smaller than a mile in diameter does not seem to be one of them. The existence of such cases entail that even if we did have a way to specify what made a system-types distinct from mere assemblies without reference to laws, we still would not have thereby solved the problem of saying what distinguishes laws from accidental generalisations.

The problems with this system-type-first approach should, I think, be clear. The second option starts instead with a prior account of what makes laws themselves special. With that in place, we might then hope to work backwards and define the distinctive feature of system-types as whatever assemblies are denoted in the antecedent of a law. This suggests that a final account of system-types will remain deflationary: what makes system-types distinct from mere assemblies is that they are the types of assembly which feature in the antecedent of laws. Moreover, this approach suggests that there are no system-types which do not feature in laws. Although the word 'system' may be used for other kinds of assembly, the notion of a 'system-type' would therefore have a stricter sense of being *by stipulation* a type of assembly of objects which appears in the antecedent of a law.

Of course, if one believes that laws can be contingent, it is consistent with this second approach that there are system-types which have instances in worlds which do not obey the law of this world in which they feature as an antecedent condition. Nonetheless, in actuality the notion of a system-type would never be applied to characterise some class of actual systems without there being some type of behaviour those system-types are lawfully associated with. Moreover, this actual connection with behaviour does not mean that behaviour is part of system-types' criteria of individuation. Consider an analogy. Arguably, the criterion of individuation for properties is their intension (the function from worlds to individuals). We might wish to distinguish from among the properties the *intrinsic* properties as peculiarly special. Intrinsic properties are, let's assume, distinguished from other properties by being independent of accompaniment by other properties. But we need not call upon this fact for individuation of intrinsic properties, since intension remains sufficient and necessary for the individuation of any property whatsoever. Similarly, a particular characteristic of behaviour might be crucial for saying what distinguishes system-types from mere types of assembly but the above criterion of individuation in terms of conditions of assembly need not refer to such a characteristic. In a slogan, conditions of assembly locate the system-type, association with behaviour by law is what makes system-types special.

The account of laws developed in §5 and §6 aims to characterise a particularly important kind of behaviour understood as a 'robust causal junction'. It is this which I think ultimately explains what makes laws, and consequently system-types, special. But we are still not finished discussing the nature of system-types, specifically whether they can really be understood to concern real-world entities.

3.2.2 Laws concern real-world systems (not models)

In §3.2.1 I argued that system-types are individuated by their conditions of assembly. The examples I gave strongly suggest that the conditions are the sort satisfied by real-world entities rather than abstract entities. In this section I want to comment on a view growing in popularity that laws should not be understood in this way, but rather to be about *abstract models*, or some kind of entity distinct from the real-world systems we intuitively take them to concern.

In the last forty years it has become common among philosophers of science to downplay the importance of laws in scientific practice. The reason for this often has something to do with laws' apparent failure to precisely describe those systems to which they are typically understood to concern, i.e. there is a failure of laws' universality (see Cartwright 1983, 1999, Dupré 1993, Giere 1988, 1995, Mumford 1998, Lowe 2006). No gas precisely obeys the ideal gas equation, no pendulum precisely obeys the classical pendulum formula, no mutually gravitating bodies precisely obey Newton's gravitational equation, etc. This 'problem of non-universality' can seem to render many laws *false*, which under most perspectives is seriously undesireable.

In order to save truth, some philosophers have argued that laws must be understood to concern entities *other than* those real-world systems to which we normally take them to apply. For instance, Giere suggests that we understand laws as,

> statements that characterize the structure of a theoretical model (p.42) [... where ...] a theoretical model is an abstract structure that may be characterised in a variety of ways utilizing a variety of linguistic resources (p.41) [...] The Boyle-Charles gas law, for example is to be understood as asserting that the pressure, volume, and temperature of an ideal gas are constrained to satisfy the relationship PV=nRT. Note that this statement is not to be understood as a universal generalization concerning real gases, but as part of the characterisation of what it is to be an ideal gas. (1988, 42)

As Giere notes, this avoids the problem of non-universality by rendering laws about something alternative to real-world systems about which they are true by definition. Similar points have been made by Cartwright who argues that 'the basic scientific laws do not literally describe the behaviour of real material systems' (1983, 203) and by Lowe (2006), who identifies the subject of laws to be abstract particulars, viz. *kinds*, in the hope of preserving their truth.

These moves to shift the subject of laws away from real world systems appears to be part of a general and increasing distaste with laws as the dominant tool of scientific reasoning. Giere's latter work, for example, bears the subtitle *'Science Without laws'* (1995) and a significant focus in philosophy of science is now focused directly on the models themselves rather than the universal lawlike generalisations which may be about them. It has also been argued that models need not be understood as exclusively abstract mathematical devices. Models can include physical objects (Morgan and Morrison 1999), fictions (Frigg 2010), analogies (Hesse 1963) and metaphors (Black 1954). Supposedly it is *these* entities which our scientific generalisations really concern.

But while these posits may help reclaim lawlike generalisations' universality, it comes at the cost of making the generalisations no longer seem relevant to the systems to which they are applied in practice. Those who enquire into the employment of models worry deeply about how they *represent* the parts of the real world to which they are applied (da Costa and French 1998, Van Fraassen 2008, Suárez 2010, Sánchez-Dorado 2017). In the original paper of Giere's, for instance, he suggests that elements of the structure of abstract models can be identified with those in the real world (Giere 1988, 41). But how much structure is required? Is similarity in this respect sufficient for representation (arguably not, see Goodman 1976)? More recently, it became clear to Giere (see his Giere 2010) that representation also requires an agent and a purpose in order to be successful. In general, the discussion of these issues shows little sign of abating.

Whatever the prospects for a solution to issues of models' representation, I don't really see the move of making models the subject of laws to really be worth it. Saving laws' truth by interpreting them to concern something they don't ostensibly seem to is not only in tension with common (and I believe much of scientific) sense, it generates the issue of representation whilst only postponing the problem of non-universality which motivated the move in the first place. In the discussion of representation it seems to me that what is being searched for are rules which tell scientists *when* it is appropriate to apply a model (fiction, analogy, metaphor). If that's the case, then what are needed are conditionals with an antecedent satisfied by all and only those worldly entities which exhibit the right sorts of features to have a model appropriately applied to them and a consequent satisfied by all and only those entities which behave in a way appropriately similar to the behaviour of the model (whatever 'appropriate similarity' amounts to). But this is just the very structure I argued that laws have! Moreover, such conditionals have real-world systems as their subject, plausibly individuated by conditions of assembly divisible into conditions of componency and organisation.

Ultimately, the problem of non-universality isn't really avoided by interpreting laws as concerning abstract models (fictions, metaphors, analogies, etc.) since there still remains the problem of explaining how those models represent the messy world. This will only be solved if we reinstate something that looks very much like what I've been arguing laws look like. As a consequence I think we'd do better to bypass the move to abstraction (or some other kind of model) and stick to the prima facie intuition that laws are about things in the world. Obviously, the problem of non-universality still remains (I discuss this further in §3.3), but the understanding removes much of the mystery surrounding representation by bringing the application of laws and numerical equations right back to things in the real world.

3.3 Laws are universal

In trying to argue against an interpretation of laws as concerning abstract models rather than real-world entities I have, admittedly, ignored the worry which motivated it in the first place: the problem of non-universality. If the domain of laws is to be understood as concerning real-world phenomena (rather than a model) and being universally quantified, as Schema 2 is, then any system satisfying a law's system-predicate which does not conform to the behaviour described in the law is a counterexample of the law rendering the law false and thereby (we might presume) unexplanatory. Here I want to discuss a few approaches to deal with apparent failures of universality which all stem from a common intuition described in §3.3.1 that laws characterise some embedded phenomenon even in the case of their counterexamples. Although I favour none of these approaches to the exclusion of others, in §3.3.2 I will raise some problems for the dominant approach which aims to modify the system-predicate in order to save laws' truth. In §3.3.3, I then offer an alternative approach of modifying the *behaviour-predicate* so that they define special cases of more precise laws. This second approach is rather under-explored but nonetheless seems to be extremely powerful. Adopting it will become of key importance for later chapters.

3.3.1 The Embedded Phenomenon Inuition

As Hüttemann (2007) has pointed out, a law's claim to universality can seem to be undermined in a number of (potentially overlapping) ways. For instance, universality can seem to be undermined by a particular subset of the domain being counterexamples, e.g. the occasional deformity in a species. It can seem to be undermined by a particular space-time region, e.g. if physical constants are thought not to hold in the early universe. Universality of some laws seems undermined by particular environmental circumstances, e.g. when gravitational forces complicate the effects of Lorentz forces. And universality can seem to be undermined by a particular range of values of the included variables, e.g. when growth models break down at extreme values. Another dimension is the *extent* of undermining. Universality can be undermined by extremely rare recalcitrant cases, e.g. in entropy-decreasing closed systems; in a high volume of cases, e.g. in growth models; or in all cases, e.g. for any idealising or 'classical' law.

Despite this range of putative failures, I think there is plausibly something which binds all the mentioned cases together: what I call the 'embedded phenomenon intuition'. The embedded phenomenon intuition suggests that while a strict law connecting a system-type to a behaviour may be rendered false by counterexamples, all instances of the law's system-type *including the counterexamples* provide a consistent 'contribution' to the overall behaviour of whatever situation they are in (Corry 2009). In genuine examples of the law, the contribution is all there is to the overall behaviour. But in the case of a counterexample, the respective system's behaviour is embedded in the behaviour of a more complex scenario. So, for example, a deformity in a species may come about from some mutation, but the underlying mechanism of heredity and chromosomal translation in reproduction remains a constant contribution to the overall behaviour. In the case of Lorenz-force, we expect there to be a constant contribution of magnetic and electric fields on the force experienced by a charge even when gravity contributes a further component of force.

The embedded phenomenon intuition is reflected in a theory of non-universal laws offered by Pietroski and Rey (1995). In the terminology developed in §3.1, the theory claims that for any system which satisfies the system-predicate of a law but fails to satisfy the behviour-predicate, one can infer that there is a (possibly unknown) independently confirmable disturbing factor which explains that failure. In other words, external factors can disturb but do not eradicate the contribution of underlying phenomena attributed to the system.

Now, Pietroski and Rey's account has been objected to on the grounds that satisfaction of these conditions is trivial: with a broad enough conception of what counts as an independent disturbing factor C, any system-type S and behaviour B can form a true statement of the form 'if x is S, then it is B unless factor C is present' (Earman and Roberts 1999, Schurz 2002). The objection indicates the failure of the theory to suitably explain what distinguishes laws with counterexamples from any arbitrarily cobbled together generalisation. But that is not to undermine the general insight about what sort of entity is being described in these non-universal laws' system-predicates. The idea of an independent but disturbing factor indicates that there is some*thing* which can be disturbed, perhaps a natural kind (Harré 1993, Ward 2007, Nickel 2010), some kind of paradigm (Cartwright 1983), a 'normal' system (Pietroski and Rey 1995, Schurz 2002), or an underlying mechanism (Glennan 1996). This is exactly what I intend to refer to as the emebedded phenomenon intuition.

Obviously, the intuition does not in and of itself save the truth of a law. Understanding that the term 'reproducing population' picks out a specific natural kind or mechanism allows us to be confident that disturbing factors may be explained by further independent considerations beyond those in simple growth laws. But for all that the intuition suggests, the Malthusian growth law remains false. Without further qualification the growth law says that *all* populations will develop in a certain way which we know in fact they won't. In what follows I will explore some responses to this general problem each of which tries to say something about how to save the truth of laws whose instances appear to suffer from such disturbances. I will consider two broad approaches: *system-type modification* and *behaviour-modification*.

3.3.2 System-type modification

If all laws are indeed system-laws (see §3.2), it might seem that we have a fairly straightforward response to the problem of non-universality. For it is always possible for us to simply deny that the right system-type has been predicated. This move makes the classification of laws as 'ideal' or 'classical' logically significant. Although gases in general do not satisfy the ideal gas equation, *ideal* gases do; although not all pendulums satisfy the classical pendulum law, *classical* pendulums do.

This response to the apparent non-universality of laws is rightly figured to have an air of triviality about it unless the notions 'ideal' and 'classical' can be given an informative gloss and as it turns out we can get pretty close in these two cases. Ideal gases are defined to comprise particles of negligible size with no interaction forces; classical pendulums operate in a uniform gravitational field with an inextensible cord and frictionless pivot. Hence, the approach in general is to 'modify the system-predicate' to get a law which is no longer subject to counterexample.

This tactic has been endorsed on a number of occasions (although not in such terms). For instance, Fodor (1991) gives the suggestion that we might append a 'completer' C to antecedent conditions in laws to save their truth in many cases (see also Reutlinger 2011). Consequently, idealising laws would have the following more appropriate form.

$$\forall x((Sx \& Cx) \rightarrow Bx) \qquad \text{(Schema 4)}$$

In some cases we may expect to specify enough of the potential disturbing factors in the antecedent to retain a law's truth (hence 'completer'). Perhaps Newton's first law is an example of this. It tells us how massive systems behave in the absence of disturbing factors where we know (*pace* Cartwright 2000) what the likely disturbing factors are: electromagnetic, gravitational, strong or weak forces. Hence, we can introduce a more explicit expression of Newton's first law which retains universality

> **Newton's First law:** All materially closed systems isolated from any electromagnetic, gravitational, strong or weak force will remain at rest or at a constant velocity.

Despite how promising this approach of system-type modification may look, one might be concerned that it will not always be available. Consider, for instance, Malthus's growth law. A first pass at rendering it might be by the following conditional.

> All ecological populations increase in population size according to the equation $P(t) = P_0 e^{rt}$.

In this form it has many counterexamples. Populations will often be under environmental pressure from natural disaster, disease, predation, habitability of geographical area, etc. In response to these putative counterexamples we might attempt a modification of the system-predicate with a further conjunct as with Schema 4. But in this case it is not so easy to come up with anything precise which assures we condition on only those systems which do behave in the way described by the equation other than by saying something trivial such as that the systems 'have no disturbing factors'.

Given that one can't name all the specific kinds of counterexample to many direct inferential associations between system-types and behavioural formulae, an alternative option is to qualify the system-predicate with a less specific completer-clause which holds in general for all laws with problems of universality. There are two options I know of along these lines, both of which I believe to be problematic.

One suggestion is to modify laws so that they concern only *'normal'* instances of the system type (i.e. $Cx \equiv x$ is normal; see Spohn 2002, Nickel 2010) This may seem initially promising, although we may wonder why laws are rarely made explicit in this way. More problematically, the qualification on normality doesn't seem to appropriately capture the fact that scientists generally *know* when and why a system will fail to obey some behaviour. When an unexpected failure occurs, scientists will actively look for a *reason* for it. The suggestion of a normative qualification simply doesn't seem to have the sufficient precision to make sense of this. Moreover, the restriction on normality suggests that genuine instances of the law will actually exist. But in the case of many laws there are *only* counterexamples (e.g. idealising laws). Ultimately, normative system-type modifiers may provide a modicum of understanding, but they do not respect the kind of information scientists are interested in positing when positing a law.

A different suggestion for how to deal with problems of non-universality is provided by Hüttemann (2004). He suggests that in general, laws describe the behaviour systems approach the more *isolated* they become (i.e. $Cx \equiv x$ is isolated). While disturbing factors in the real-world might divert a system's behaviour from what the law predicts of it, the same system in isolation would have no disturbing factors and hence satisfy the behaviour-predicate.

This 'isolation-approach' to otherwise problematic laws seems continuous with the embedded phenomenon intuition, since it suggests that once we take away all the other factors in which a system is embedded, the behaviour will return to that predicted by the law. Moreover, in many cases I think modifying the system-predicate with the further condition of isolation would indeed retain truth—although its application beyond the physical sciences may be hard to conceive. Furthermore, we might think that it is a far more plausible representation of scientific knowledge than the use of a normative modification, since isolation actually aims to specify the kinds of conditions under which a system-type's instances will obey a certain behaviour-predicate.

Nonetheless, the isolation-approach has significant limitations as well. I offer three specific complaints before drawing conclusions on system-type modification in general. A first reason one might be sceptical about the isolation-approach is that the concept of isolation seems causal, i.e. an abbreviation of 'isolated from disturbing causal factors'. If isolation can't be given a non-causal interpretation, then Hüttemann's solution to the problem of non-universality blurs the distinction between laws of association, which aren't supposed to contain any explicit causal language, and causal laws, which are.

A second concern is that despite an improvement on normative criteria, the concept of isolation is not really precise enough. Assume, for instance, that Coloumb's law is an example of the sort Hüttemann describes. It states that if a system comprises two charged objects separated by a non-zero distance *and is isolated from any other influence*, the force experienced will be given by the equation,

$$\mathbf{F} = k\frac{q_1 q_2}{r^2}.$$

But this will only be true if we understand isolation in a specific way. It can't simply be understood as meaning that the two objects in the system are the only two contributing to each other's behaviour, since if they have mass there will still remain the disturbing factor of gravitational force. It can seem, therefore, that we need to understand the requirement on isolation in this case to pick out precisely the components' charges and nothing else about them. But is it legitimate to suppose that the law really concerns only charged objects with no mass? After all, the only charged objects in the standard model without mass are relatively recently discovered W bosons, and they are posited to mediate an entirely different force, viz. weak force. By the time we concede enough physics (or other science) to make coherent a scenario in which two objects are influenced *solely* by each other's charge, we may wonder whether it is reasonable continue to assume the relevance of Coloumb's law in that scenario too.

A third complaint specific to the isolation-approach is that because the expected behaviour is supposed to be fully manifest in isolated conditions, it requires that all laws' behaviour-predicates are those in which the variables for disturbing factors are set to zero. But some idealisations are those in which unwanted variables are fixed to other values. For instance, when we stipulate that a cord is inextensible, we are idealising by assuming its resistance to extension under load (Hooke's spring constant) is infinite; and when we idealise in order to use Newtonian mechanics to predict the path of a projectile, we often assume gravitational acceleration is $\approx 9.81ms^{-1}$. As long as we require that the ideal behaviour is at *some* limiting case we can allow such idealisations which do not condition on isolation.

Drawing on all the proposed system-modification approaches considered, they seem to all suffer from what has become known as 'Lange's dilemma' (after Lange 1993a). Lange pointed out that qualifications of law tend either to be too specific about disturbing factors, in which case they remain false, or else they are too unspecific, in which case they are trivial. Attempting to specify exactly the kinds of counterinstances which might crop up in the antecedent is often impossible, hence any attempt will not save the laws' truth. But attempting to save truth by simply conditioning on 'normal' or 'isolated' instances of the system-type seems either open to such wide interpretation that it is bound to be true or else insufficient to capture the exact cases we want.

But even if some kind of system-type modification can work in certain cases, there is also another problem which befalls any of these approaches which I think highlights best the need for a different approach entirely. Note that any system-predicate modification approach works by *restricting* the class of entities the law is about. For instance, it asks us to consider *only* the ideal cases, the normal cases, or the classical cases. But often laws are used to say things about the very systems we know fail to perfectly obey the laws' behavioural formulae. This is not recalcitrant scientific practice, but exemplary of the process of idealisation: when we come across a complex situation for which a perfectly accurate law is either beyond our knowledge, our ability to solve, or our practical requirements, then we can often benefit from idealising the situation. Crucially, this practice is one in which we use a law with some explicit formula to talk about a system which does not precisely obey it.

Take, for example, Malthus's growth law. Many (perhaps all) real-world populations fail to obey the explicit equation in that law,

$$P_t = P_0 e^{rt}.$$

In response to this, the approach of system-predicate modification suggests that we treat the law as really being about only those cases which do obey the formula, e.g. the ideal, normal, classical, or isolated cases. But even if that succeeds in rendering the Malthusian law true, we now have the problem of saying exactly what relevance those normal cases have to all the non-ideal, abnormal, non-classical, unisolated cases. And there must be *something* which connects them, since often in practice an ecologist will use the equation to describe the behaviour of these latter cases.

The problem described here certainly bears similarities with that of representation mentioned earlier. Just like the problem of saying how analogies, metaphors, fictions and abstract models are able to represent a wider class of system (including real-world cases) those who pursue the system-modification approach to the problem of non-universality have to say how a more restricted class of systems can count as representative of a wider class. But this whole issue could be abandoned if one simply took laws to be about the entities they're used to describe. This is where I think behaviour-predicate modification may have the answer.

3.3.3 Behaviour modification

Most attempts to rescue the truth of laws in the face of the problem of non-universality seem to suggest some modification of what in current context we may call the system-predicate. But what I have just suggested is that we might also consider rescue by qualification of the behaviour-predicate. For instance, we might expect that if a phenomenon underlies a more complex scenario, it will reveal itself at least partially in that complex behaviour. The issue then is how to make such an idea precise. Here I think a passage from Malthus on his growth law can be instructive.

> The passion between the sexes has appeared in every age to be so nearly the same that *it may always be considered, in algebraic language, as a given quantity.* The great law of necessity which prevents a population from increasing in any country beyond the food which it can either produce or acquire, is a law so open to our view [...] that we cannot for a moment doubt it. The different modes which nature takes to prevent or repress a redundant population do not appear, indeed, to us so certain and regular, but though we cannot always predict the mode we may with

> certainty predict the fact (my emphasis Malthus 1798, Ch.4).

Malthus seems here to clearly be endorsing something like the embedded phenomenon intuition. In the first sentence, he suggests that there is some natural or unchangeable mechanism—the passion of the sexes—from which the exponential growth model can be inferred. This is where we get Malthus's law from. In the second sentence, he suggests that despite the regularity of the mechanism, we know (by the 'law of necessity') such mechanisms will always be disturbed, resulting in behaviour other than that predicted by the passion of the sexes. But Malthus also makes the further claim that the passion of the sexes 'may always be considered, in algebraic language, as a given quantity'. This goes beyond the embedded phenomenon intuition and actually suggests something about how we might formalise the contribution a system makes in embedded contexts.

An idea I find plausible, and which complements Malthus, is that the inaccurate formulae used to characterise some laws' behaviour-predicates describe a *special case* of more complex and more accurate formulae involving further variables not inferred from the system-type alone. A more complex formula which corrects that in the law will typically comprise further variables which, when set to some other fixed number, will reduce the complex formula to that of the law's. In the case of growth models we can, for instance, derive the Malthusian equation from the Lotka-Volterra equations:

$$\frac{dx}{dt} = \alpha x - \beta xy \qquad \frac{dy}{dt} = \delta xy - \gamma y$$

These equations are used to describe systems in which different populations compete, e.g. via predation. If we set the population of one species (e.g. y) to zero, we get the differential interpretation of Malthus's equation for the 'passion of the sexes' of the other population (x) as a consequence.

A similar observation works for many other of equations. For example, the applicability of the Rayleigh-Jeans formula to cases of black-body radiation at low frequency can be obtained by putting the Planck constant to zero; the Newtonian constitutive equation for incompressible flow can be obtained from the constitutive equation for compressible flow by putting the term for change in mass-continuity to zero; the dynamical behaviour of an undampened oscillator can be obtained by putting viscosity of non-Hookean springs to zero; the pressure-volume relationship of an ideal gas can by obtained from the Van-der Waals equation by putting mediating forces and molecule-size to zero, etc., etc. In each case, we can think of the law making explicit the less complex formula as derivable from the more complex formula, which goes unspecified. Note, however, that the variable being set to a single value need not be set only to zero. For instance, the Rydberg formula can be obtained from the Bohr model of the atom by putting the atomic mass to infinity and treating the Rydberg constant to be equivalent to the term,

$$\frac{m_e e^4}{8\epsilon_0^2 h^3 c}.$$

In general, I suggest we treat behaviour-predicates of laws as not being fully characterised by their contained formulae, but by the condition that whatever satisfies them is described accurately by some further unspecified formulae which have the contained formulae as a special case. Since the more accurate formula remains unspecified, we can thereby retain the truth of laws despite the existence of many possibly unknown disturbing factors.

Mathematically, there may appear something rather vacuous about this suggestion. After all, isn't it trivial that for any formula there is a more complex formula which has the first as a special case and which will fit any finite set of data? Maybe so. But that doesn't mean the suggestion for preserving the truth of laws renders them trivial. Practitioners don't expect the additional variables, coefficients and functions to come from anywhere but to be grounded in scientifically appropriate and plausible properties of the surrounding environment. Sometimes a property might be introduced in a rather ad hoc basis to retain the underlying phenomena (e.g. dimensional constants) but for the most part, the additional variables will either be 'natural' additions to the numerical explanation of a complex behaviour (e.g. when molecule size is included in real gas equations) or become fundamental constants of a burgeoning scientific research programme (e.g. when the Planck constant was introduced into descriptions of black-body radiation).

Of course, if the modification is to avoid triviality then we certainly need to include in the modification to laws' behaviour-predicates a characterisation of what kind of unspecified formulae are acceptable. Here I think the embedded phenomenon intuition may again be of assistance. For it tells us that instances of the system-types of idealising laws really do exist, but typically in a more complex environment in which those systems are embedded. This suggests that for any complex environment in which an instance of a system-type S is embedded, we have an instance of a different system-type S' of which the instance of S is a part. In this context, parthood can be understood to mean that the conditions of assembly for S are entailed by S' (hence S is a part of itself). So, for example, the conditions of assembly for a population in a habitable environment will be entailed by the conditions of assembly for two populations (one predatory on the other) in that environment.

Obviously, pointing to a more complex system in which some other system is embedded is of no benefit to providing us with a restriction on the relevant unspecified formulae K' unless we can connect the two together in some informative way. The trick is, I think, to connect S' and the unspecified formula K' *by law*. In other words, we preserve the universality of a law relating a system-type S and behaviour-type B by expressing it in the following form.

> For all x, if x is in instance of the system-type S, then there is a system-type S' and behaviour-type B' such that,
>
> 1. $\forall x(S'x \rightarrow B'x)$ is a law, and,
> 2. instances of S' have an instance of S as a part, and,
> 3. $\forall x(B'x$ if and only if x has variables $V_1, V_2, ...$ related by the formula K' which has as a special case the formula, K).
>
> (Schema 5)

In instances of the schema, K is replaced by the laws' explicit formula and K' by a variable ranging over unspecified formulae. Notice that to state a law in this form, the corresponding law $\forall x(S'x \rightarrow B'x)$ referenced in the modified behaviour-predicate need not be stated; the behaviour-modification simply asserts that there is such a law. Consequently, it is not essential to know how the explicit formula K might fail to accurately describe the system it is applied to in order to know that the behaviour-modification is appropriate. So, for example, if we wanted to save the truth of Malthus's law by this behaviour-predicate modification technique, we could propose that it be understood as follows.

> **Malthus's growth law:** For all x, if x is in instance of a reproducing population in a habitable environment, then there is a system-type S' and behaviour-type B' such that,
>
> 1. $\forall x(S'x \rightarrow B'x)$ is a law, and,
> 2. instances of S' have an instance of a reproducing population in a habitable environment as a part, and,
> 3. $\forall x(B'x$ if and only if x has variables $P_t, P_0, r, t, ...$ related by

the K' which has as a special case the formula,

$$P_t = P_0 e^{rt}.)$$

Let me draw attention to a couple of details in this approach to behaviour-predicate modification. First, it's quite possible within the approach that the instances of a law which don't precisely obey the law's explicit formula can appear in a range of different types of system. The Malthusian equation doesn't only fail to describe a population's growth because of predation, but also because of environmental pressures, natural disaster, disease, etc. Plausibly, there is a distinct and more complex system-type for each of these additions. For the behaviour-predicate modification to be correct, it only needs to be the case that for each instance of the Malthusian system-type, there is *some* law which satisfies the criteria 1–3 in Schema 5.

Second, I presume that there is a reasonable amount of vagueness surrounding the 'correct' expression of Malthus's growth law with the prescribed behaviour-modification. In some cases, ecologists might have a much better defined list of criteria for the applicable system-type S specified in premise 2, in other cases, perhaps, fewer. This doesn't seem to me a problem so long as they various expressions bear a significant degree of overlap.

Let me now address a predictable concern. According to the behaviour-modification approach as I've introduced it, laws like Malthus's are acceptable only so long as there exists some other law from which the system-type and formula in Malthus's law can be derived. But what if those further laws describe embedded phenomena as well, i.e. suffer from instances which don't obey their own explicit formulae? This would be a problem, I think, if for any instance of a law with modified behaviour-predicate there were *no* law which existed for which its explicit formula was always obeyed. Behaviour-modification can't go on forever and eventually there must be some laws which just get the situation exactly right.

Maybe this assertion is overly optimistic, but in its defence I should reiterate that these laws need not be known now or ever. So long as for each instance of a reproducing population in a habitable environment there is some law out there—perhaps incredibly complex—for which the corresponding system-predicate at least requires its instances to also be a reproducing population in a habitable environment *and* its behaviour-predicate both describes growth absolutely correctly and whose explicit formula has the Malthusian growth equation as a limiting case, then Malthus's growth law (in the above modified form) is true. Moreover, as it turns out, the account of laws I will develop throughout §5 and §6 makes no requirement that any law have more than one instance (unlike the axioms of the best system, see §4.2). It is therefore consistent with the view that each of the perfectly accurate laws has only one instance (although it will need to have at least one instance).

One further query we might have over the behaviour-modification is whether or not *any* special case of a formula in an accurate law can be used to define a further idealising law. While I think it is beyond the scope of the solution to problem of non-universality to have a ready answer to this question, it will be useful here to speculate. For it seems clear to me that only some special cases are reasonable candidates for defining the behaviour-predicate of a law. For instance, while the equation in the ideal gas law is a legitimate special case of the equation in Van der Waals's law, an equation which instead took the special case in which pressure and volume were held fixed as constants would seem not to be. This, I believe, is due to the fact that the relation of special and general case between laws and their idealisations must preserve a continuity of *causal asymmetry* among the variables. I will return to this idea in §5.4.5.

If the approach of behaviour-predicate modification ('behaviour-modification' for short) can be made to work, it provides a number of benefits over the approach of system-predicate modification discussed in §3.3.2. Unlike a condition of 'isolation' or an explicit prohibition of 'disturbing factors' the approach as I have advocated doesn't make any explicit causal claims. The expression of the idea is purely mathematical. Hence the approach retains the prospect of a Humean account of laws. (Indeed, I will later use this modification to *analyse* the causal disposition of embeddability; see §6.1.) Also, notice that the approach doesn't require from us that we conceive of what a situation would be like in which the disturbing factors are absent. The system-types and behaviour of the more accurate laws go unspecified in behaviour-modification.

Another benefit is that we are not restricted in what amount of influence we take the disturbing factors to have in cases which do obey the explicit formula. In system-predicate modification (and this is most clear in the use of an isolation-modifier) the disturbing factors must be completely non-existent for the ideal behaviour to occur. However, as we saw, some laws don't seem to follow this pattern. The behaviour-predicate modification I have suggested only requires that there are *some* values at which the variables which don't feature in the explicit formula can be fixed such that the explicit formula results. Often this value is zero, but it need not always be (the special case of the Rydberg formula is one such example).

Finally—and I take this to be the approach's most significant merit—the behaviour-predicate modification allows us to say that, for example, Malthus's growth law is about any population for which it might feasibly be used to explain. It is not, therefore, restricted to a subset of populations which behave exactly according to the explicit formula. Consequently, there is no problem of giving a story about analogy, metaphor or abstraction in order to explain why laws like Malthus's might be relevant in all the real-world cases we apply it to.

This last point raises an interesting question over what to make of laws where system-predicate modifications of the sort considered in §3.3.2 actually occur in practice. The ideal gas law, for instance, is often explicitly referred to as a law about ideal gases. This suggests that theorists have opted for a system-predicate modification: the law is not about real-world cases, only some abstract ideal. This is how the law was presented earlier in §3, repeated again here (in slightly different notation).

> **Ideal gas law*:** For all x, if x is an ideal gas in a cavity then x has variables P, V, n, T which satisfy the equation,
>
> $$PV = nRT$$

However, if we wanted, we could offer an alternative interpretation of the ideal gas law as really being about *all* gases in a cavity, represented as follows.

> **Ideal gas law**:** For all x, if x is a gas in a cavity, then there there is a system-type S' and behaviour-type B' such that
>
> 1. $\forall x(S'x \rightarrow B'x)$ is a law,
> 2. instances of S' have an instance of a gas in a cavity as a part, and,
> 3. $\forall x(B'x$ if and only if x has variables P, V, n, T, related by the formula K' which has as a limit case the formula,
>
> $$PV = nRT)$$

So which interpretation should we opt for? It seems to me that both have their merits. The ideal gas law* says explicitly what it would take for a gas in a cavity to obey the explicit formula. Such a formulation of the law is possible because, unlike the Malthusian growth law, we know exactly what the system must be like to exactly obey the formula. But the ideal gas law** says explicitly how the formula might come to be applicable to real-world cases. This is something I have argued that system-predicate modifications don't do so easily. Ultimately, I think it is plausible that both interpretations reflect genuine and distinct scientific knowledge and are worth making explicit. Something similar could be said for many laws which appear to have been given explicit system-predicate modification sufficient for retaining truth (e.g. Newton's law for incompressible flow).

At the very least, I will from hereon assume that the behaviour-modification is a suitable way to solve problems of non-universality in laws.

3.4 Beyond logical form

In this chapter I have argued that all laws are system-laws, specifically with the form of Schema 2 where the antecedent is satisfied by instances of a specific system-type and consequent instances of a specific behaviour, often characterised by an equation. I also just offered in §3.3.3 the more specific Schema 5 for laws whose instances otherwise appear to suffer from failures of universality. We have seen that while we know what sorts of conditions individuate a system-type, we couldn't say that such conditions make system-types in general distinct from mere types of assembly for which there are no laws. I also suggested that the best answer to this particular question would start with an understanding of what makes laws considered as a whole special.

But by focusing on logical form, I haven't been able to say very much about what makes laws special. For presumably mere types of assembly can also feature in the antecedent of conditional generalisations with a behavioural consequent. This is not a new problem. It is arguably *the central issue* of providing a philosophical account for laws to say what makes them distinct from accidental generalisations (see Ayer 1999 [1956], Hempel 1965, Swoyer 1982, Lewis 1983, Armstrong 1983, Carroll 1994, Harré 1993, Maudlin 2007, Carroll 2016). However, what I hope to have achieved in this chapter is to provide a new and precise way of framing the issue. For we know now that whatever it is that makes laws special, they are conditional generalisations which condition a behaviour on instances of a certain system-type. This observation will help us both to critique contemporary accounts of law and also to point us in the direction of a more promising account.

4 The causation-mirroring conception of laws

In this chapter I will discuss and criticise a way in which laws have often been thought about which I call the 'causal-mirroring conception of laws'. This conception has it that laws structurally mirror the relationship of cause-effect relations including those separated in time and space. Drawing heavily on §3 I will show in §4.1 why this conception is misguided. In the first instance this will help to more clearly see what a plausible nomological account of causation will have to look like. But the discussion will also help me in §4.2 to comment on Humeanism's most popular account of laws, the *best system account*, which I argue is implicitly committed to the causal-mirroring conception.

4.1 Against the causation-mirroring conception of laws

This section discusses a particular way in which laws have been been conceived often as a consequence of the conviction that laws bear an inferential connection with causal relations. As will become clear in §5, I am sympathetic to the conviction in general, but not in the specific way this particular conception has it. Briefly, this conception is captured by the idea that for any causal relation, there is a law which 'covers' it, what I call the 'causation-mirroring conception' of laws. Despite its popularity and simplicity, I am going to argue against it. My criticism stems from the central complaint that the conception represents laws with a logical form distinct from their true form, i.e. Schema 2,

$$\forall x(Sx \to Bx).$$

As a result the conception is incompatible with the observations of the foregoing chapter. My plan is first to introduce a generalised logical form for laws as understood by the causation-mirroring conception and then show in §4.1.1 and §4.1.2 how each of the features which distinguishes it from Schema 2 lead it to trouble.

The causal-mirroring conception of laws takes laws to have a similar structure to that of a causal relation. One might think this if, for example, one agreed with Armstrong and Heathcote (1991), Armstrong (1997, 2004a) who argue that causal relations should be understood simply as *instantiations of laws*. Armstrong and Heathcote's particular explication of this comes idea packaged up with various metaphysical commitments to nomological necessitation relations, states of affairs and other dubious entities which we might understandably be weary of. But other less metaphysically-loaded positions have suggested a very similar relationship between laws and causation. Indeed, regularity accounts in the Humean tradition have often made effectively the same judgement (e.g. Ayer 1999 [1956], Goodman 1954 [1983]). For these theorists, a regularity simply generalises over a conjunction of property-instances or events which constantly occurs in nature. Although there are exceptions (see, e.g., Baumgartner 2008, 2013b), typically these accounts of law distinguish the cause-variables from the effect-variable in the generalisation by virtue of the temporal order of their respective instances—causes precede effects—and certainly it seems in general to be assumed that instances in antecedent and consequent of the regularity should be spatiotemporally distinct.

Even when laws are not doing all the analytical work in analyses of causation, it is also often the case that causal relations are treated like instances of laws. For instance, Paul (2000) argues that a necessary condition for a causal relation between c and e is one of lawful entailment, defined as follows.

> For any two distinct, actual events or individuals c and e, and any two logically distinct properties p and q, aspect c_p (c's having p) lawfully entails aspect e_q (e's having q) iff c's exemplification of p is subsumed by the antecedent of the right law or laws that entail a consequent subsuming e's exemplification of q. (2000, 214)

And Schaffer (2001) claims that causes are probability-raisers of chains of 'direct process links' between events where,

> c is directly process linked to e if and only if (i) c and e are actual distinct events, (ii) there is a law L of the form $(\forall R)(\forall R')((P,R) \rightarrow (P',R'))$ such that c is an instance of the antecedent and e of the consequent, (iii) there is no law *L'* more fundamental than L such that actual prior event a is an instance of the antecedent and $c\&e$ is an instance of the consequent, and (iv) $t_c < t_e$. In speaking of laws as having the form $(\forall R)(\forall R')((P,R) \Rightarrow (P',R'))$, I take the quantifiers to range over regions, and the laws to relate the properties of one region to those of another. Thus events, understood as (P,R) pairs, can instantiate the antecedents and consequents of laws. Process linkage in general is then defined from the direct case: c is process linked to e if and only if there is a chain of direct process links between c and e. (2001, 78)

There are many differences between all the above-mentioned accounts. However, they all share the common idea that laws mirror the structure of causal relationships between entities potentially separated in time and space.

This is exactly what I mean by the causation-mirroring conception of laws. Each of the above interpretations of law confers on them a common structure which licenses a cross-temporal and cross-individual implication from any instantiation of some specific property by an object or region at a time to that of another specific property by a different object or region at a different time. Schematically this looks as follows.

$$\forall x, t((Fx \text{ at } t) \rightarrow \exists y, t'(Gy \text{ at } t')) \qquad \text{(Schema 6)}$$

We are encouraged to think of the variables of this implication as ranging over the objects (regions, states of affairs) involved in cause-effect relations where instances of the antecedent assert the occurrence of a cause and instances of the consequent assert the occurrence of an effect, where $t \leq t'$ and x need not be identical with y (see Pietroski and Rey 1995, for an explicit endorsement of this schema).

Schema 6 mirrors the structure of paradigm causal relations such as that involved when someone's throw of a rock breaks a bottle. As with any mirror, the reflection captures important structural features but leaves out some depth of information in what is reflected. Hence, Schema 6 captures features typical of the causal relation: instances of the conditional may relate distinct events comprising distinct objects (a rock-thrower, a bottle) and the possibility of time-directedness (throwing the rock occurring before the bottle's breaking). However, Schema 6 refrains from any explicit causal language. This is why so many Humeans have thought it helpful in *analysing* causal relations.

It is clear from the context in which this conception of laws is often made explicit that when laws are used to analyse causal relations, they are thought to be able to do so in virtue of having this structure. If a causal relationship exists between two events a and b, and there is available some law to analyse it, then the law will do so by 'covering' it, i.e. by showing how b, under some description G, is implied (perhaps to some non-trivial probability) from a, under some description F, via a law of the form of Schema 6 (the term 'covering law' comes initially from Dray 1957, and was popularised by Hempel 1965).

Moreover, the causation-mirroring conception reads into a schema like Schema 6 the scope for chaining together laws' inferential relationships in much the same way we often do with causation. Whilst the relation of token causation is only controversially viewed as a transitive relation (see Lewis 2000, Hall 2004a, and the discussion in §7.2.6), it is invariably thought not to be *atransitive*. If the breaking bottle, caused by Billy's throw, woke the cat, then we can infer that Billy's throw woke the cat. Under the causation-mirroring conception, laws should work much the same way.

So, for instance, Hempel talks of 'genetic explanation' as having the following characteristic:

> In genetic explanation each stage must be shown to "lead to" the next, and thus to be linked to its successor by virtue of some general principles which make the occurrence of the latter at least reasonably probable, given the former. But in this sense, even successive stages in a physical phenomenon such as the free fall of a stone may be regarded as forming a genetic sequence whose different stages—characterised, let us say, by the position and the velocity of the stone at different times—are interconnected by strictly universal laws; and the successive stages in the movement of a steel ball bounding its ziggyzaggy way down a Galton Board may be regarded as forming a genetic sequence with probabilistic connections. (1965, 449)

As with chains of causal relationships, for such genetic sequences of law-based explanation to be possible they must include properties in the antecedent and consequent which are of the same broad ontological category such that the consequent of one law can feature as the antecedent of another. What this individuation amounts to exactly is hard to say. But it should be clear enough to critique the general idea in what follows.

In summary, the causation-mirroring conception of laws can be captured by the following principles:

Causation-mirroring conception of laws

A generalisation G is a law if it is a conditional for which the following holds,

1. G relates instances of two types of property by conditional inference,
2. G analyses causal relationships by covering them inferentially,
3. G's antecedent predicate can be the consequent predicate of another law.

And if we assume that some causal relations are separated in time or space, the conception also implies that,

4. G can support a logical or statistical inference of the occurrence of later events from the occurrence of earlier events.
5. G can support a logical or statistical inference of the occurrence of an event involving one object/region from the occurrence of an event involving a non-spatially overlapping object/region.

So expressed, the conception only aims to provide sufficient criteria for laws. I see no reason why one under the sway of the conception couldn't consider there to be other kinds of law which support other kinds of inferences. However, it seems nonetheless plausible that the conception could be put to work in philosophical analysis. If one takes causation to be a primitive, or at least prior to laws with regards to metaphysical analysis, the causation-mirroring conception could be interpreted as a partial *account* of laws. Such an account would consider some generalisations' worthiness of lawhood to be *constituted* by satisfaction of the criteria specified in the conception. Alternatively, many of the other philosophers considered above, particularly those with Humean intentions, have considered it more plausible that the conception be rather treated as something to help analyse *causation in terms of laws* (e.g. Paul 2000, Schaffer 2001). Either way, if one accepts the truth of the conception, it's no wonder that Schema 6 is the relative schema of choice for laws. In what remains of this section (§4.1) I will not assume that the conception is to be employed in either direction of analysis. I simply want to critique it on its own merits.

There is much in the causation-mirroring conception of laws which is in agreement with the discussion on laws' logical form in §3.1, specifically, Schema 6 has a universal quantifier and is a conditional. But Schema 6 is not the particular schema I argued for in §3, viz., Schema 2. Indeed, we might count Schema 2 as capturing limit cases of the more general Schema 6, in which $x = y$, $t = t'$ and in which the antecedent and consequent predicates are restricted to system-predicates and behaviour-predicates respectively. Given the lack of restriction, one might think Schema 6 is the more appropriate general schema for laws rather than Schema 2—a view to which we might attribute Hempel, who claimed that,

> lawlike sentences can have many different logical forms. Some paradigms of nomic sentences, such as 'All gases expand when heated under constant pressure' may be construed as having the simple universal conditional '$\forall x(Fx \rightarrow Gx)$'; others involve universal as well as existential generalisation as does the sentence 'For every chemical compound there exists a range of

> temperatures and pressures at which the compound is a liquid' (Hempel 1965, 338-9).

In response to Hempel, we might defend Schema 2 by noting that even if his second example doesn't have the schema's form, Hempel has yet to justify the claim that it is a *law*. After all, it is not explicitly referred to as a law in science (although it may be explanatory and counterfactually robust). But a more informative line of response returns to the lessons learned in the discussion in §3.1. Laws which describe functional relationships between variable properties, as both of Hempel's examples do, are invariably descriptions of the behaviour of instances of a certain type of system. In Hempel's second example, the type of system is *chemical compound* and the functional relationship is that between a range of temperatures and a particular phase state. But the overall functional behaviour is attributed in the generalisation to the system-type *chemical compound*. Hence, the putative law tells us that there is a range of temperatures and pressures at which *the compounds denoted in the antecedent* are a liquid. It is possible then to render this generalisation in the form of Schema 2 where the consequent behaviour-predicate is understood as satisfied by an entity which *is such that there exists a range of temperatures at which it is a liquid*.

That such reformulations are often possible is unlikely to support a general preference for Schema 2 over Schema 6. In what remains of §4.1 I provide arguments against the three features which give Schema 6 its wider class of instances, viz. the ability to license cross-saptio-temporal implication, genetic implication and covering implications. If these features are to be done away with, we are left with a schema which looks to all intents and purposes, just like Schema 2.

4.1.1 Cross-temporal and cross-individual inference and laws

Schema 6 doesn't prohibit instances which associate instantaneous and co-located property-instantiations, but it is certainly formulated with a mind to avoid any limitation to such generalisations. So much is clear from the use of distinct times t and t' and distinct object-variables x and y. As far as the conception is concerned, we should expect there to be laws which license an inference from causes to effects which happen far off in time and location.

The problem with this is that there just don't seem to exist any such laws. First, look at the behaviour-predicates. In many cases what we have is a behaviour described by a formula in which the quantities are, from the perspective of the relevant theory, *simultaneously related*. The equations in Newton's law of gravitation, in Ohm's law, in Gauss's law for electricity are all portrayed as simultaneous relationships from within the classical framework they were developed. The same goes for many equations in physics.[1] Of course, we suspect now from the perspective of special relativity that gravitational force, voltage potential and charge are mediated across space over a non-zero time. But Newton, Ohm and Gauss never supposed this when proposing their formulae, and when working within their respective theories, one treats these differences as negligible.

Sometimes a time-difference is expected to exist between the correlated variables. For example, when the ideal gas law was proposed, it was already known that the temperature of a gas can (in theory) increase without an immediate change in pressure or volume. But this possibility is in most applications practically irrelevant. And even in examples where it is not irrelevant, it's not clear that we ever get a structure as that suggested by Schema 6. Consider, for example, the equation in Yoda's law. Here the

[1]This is even true of the second law of thermodynamics when expressed in its more correct differential form.

dry-weight of the crop W at harvest is expected to correlate with a value for the density ρ of seeds sewn earlier in the season. But even then, the equation is far from specifying anything like the specific times at which one variable is expected to change given the time of change in another. If the causation-mirroring conception implies precise temporal procession, then it's not clear how such equations could support it.

Regardless of the presence of temporal delays between correlated variables in behavioural formulae, the key argument in §3.1 is that laws describe how things will behave so long as they count as an instance of a particular system-type. A system's behaviour is, therefore, implied by law throughout a time only if it remains an instance of the system-type *for that time*. Yoda's power law does not say that if something starts off as an agricultural system then some outcome will appear somewhere some time later. Rather, it relates variables in agricultural systems of a certain sort *as they remain instances of those systems*.

Notice that the fact that these laws themselves aren't sufficient to license cross-temporal inference is not to suggest that causation itself can't be cross-temporal, nor that laws can't be used with sufficient other posits to help support cross-temporal inference. The point is merely that the laws don't 'cover' such relationships in the way they are often portrayed to. If laws are to be a central component in an analysis of causation, they will have to do so in a rather different way (see §7).

A little thought reveals why this is in general the case with laws of association. A system may be established at one point in time and so allow us, by law, to deduce its behaviour at that time. But simply given this fact alone, one cannot deduce how the system or indeed anything else will be behaving at any later time. After all, something could always interfere with the system in the intervening time. Just knowing that we have a system to which Yoda's law adequately applies at the beginning of a season is not alone sufficient to license an inference of the fact there will be a relationship described by the behavioural equation $W = C\rho^{-3/2}$ at the end of the season. In the first place, we need assurance that the system remains an instance of the same system-type throughout the season and is not disrupted or broken in the intervening time (e.g by the building of a motorway through the field or an earthquake, etc.).

Curiously, this issue seemed to be something Hempel was already aware of, since he admitted that,

> the additional premises required for [DN prediction/explanation] must provide not only a specification of the state of the system at some time t_0 earlier than the time t_1 for which the state of the system is to be inferred, but also a statement of the boundary conditions prevailing between t_0 and t_1 [...] The assertion therefore that laws of theories of deterministic form enable us to predict certain aspects of the future from information about the present has to be taken with a grain of salt. (1965, 366).

Hempel here seems to be making a point we might interpret in the current context as concerning the instantaneous correlation between system-type and behaviour-type in instances of laws.

Of course, we shouldn't think that problems of non-universality only exist as a result of cross-temporal interference. In §3.3, I discussed cases of interference as a result of the relevant system being continuously embedded in a more complex environment throughout its existence. But in this sort of case we might hold out the hope of providing a behaviour-modification of the law thereby saving universality (see §3.3.3). The problem with cross-temporal inference is that the possibility of defining all the conditions which need to be in place at some earlier time such that at a later time the consequence can be lawfully implied seems vanishingly small. The reason for this is arguably that the embedded phenomenon intuition (see §3.3.1) does not capture the right kind of interference. When a cross-temporal interference of a system occurs, it is possible that the influence of prior behaviour on later behaviour is completely removed. In such cases instances of the antecedent system-type are in no way 'embedded' in instances of the consequent system-type and so behaviour-modification is unlikely to be sufficient to save universality.

In response to these observations, a quick fix to the conception of laws might be to simply reject Principle 4; or more precisely, to uphold a schema for laws in the form of Schema 6 in which $t = t'$, viz.,

$$\forall x, t((Fx \text{ at } t) \rightarrow \exists y((Gy \text{ at } t)) \qquad \text{(Schema 7)}$$

The corresponding conception of laws mirrors *instantaneous* causal relationships, and clearly can't be faulted for an inadequate view about cross-temporal implication with laws. However, it is certainly far from clear that such a conception entirely avoids the problem of interference. For so long as a law licenses an inference from some property instance F to another property instance G in a different spatial location, there is again always the possibility of interference which behaviour-modification will be insufficient to deal with.

Consider another example. Often it is the case that we can explain the volume V of a cavity (e.g. a balloon) by referring to its temperature T. But we could not hope to ground such an explanation in a law of the following form.

> For all x, if x is a gas with temperature T, then (given values for P, n) there is a cavity which has a volume V

Such a conditional is simply not true—not even for ideal gases! For the gas could be out in space, i.e. not subject to any cavity, or mixed with other gases, such that the volume of any containing cavity is a function of the temperatures (and other values) of multiple gases.

This indicates the importance of being able to specify the entire setup in laws' antecedent conditions. The ideal gas law, for example, conditions behaviour on a system comprising a single ideal gas in a closed cavity, the behaviour says what those components of the system do. Likewise, the law of gravitation relates the properties of very distant objects, but only due to the fact that those relations exist among the components of a single system with which the law is concerned, viz. a two-mass system at nonzero distance. In general, laws provide inferences from one property to another belonging to the same individual system.

4.1.2 Genetic inference, covering and laws

If the foregoing arguments are correct, then we should be understanding laws to at least have the *syntax* of Schema 2. We can imagine a corresponding 'reduced-causation-mirroring conception' of laws in which Principles 4 and 5 of the original causation-mirroring conception are given up and the schema for laws (leaving time implicit) is taken to be the philosopher's caricature, Schema 1 (p.67),

$$\forall x(Fx \rightarrow Gx)$$

At this stage it is important to remember that despite syntactical equivalence, even Schema 1 is not identical to Schema 2. First, the reduced-causation-mirroring conception of laws still holds, according to Principle 2, that when a law can be used to analyse causal relations, this is because the law mirrors the structure of those relations. In other words, whenever a causal relationship '*a caused b*' is analysable by law, there is some law (or combination of laws) of the form of Schema 1 which allows an inference from a, under some description F, to b, under some description G. Second, according to Principle 3, the properties denoted in Schema 1 (i.e. instances of F and G) can be of the same ontological category so that a predicate appearing in the consequent of one law may appear in the antecedent of another. This is crucial if laws are to be to chained together in the manner Hempel suggested in order to facilitate genetic-type causal explanations.

Notice that if we adopt Schema 2 as our schema of choice for laws, one of these features may hold only for very specific causal relationships and the other not at all. Regarding Principle 2, the conditional-form of Schema 2 is at best limited to mirroring causal relationships which occur between systems and their behaviour. I think it is plausible that there are such causal relationships (in §5.2.4 I will discuss these relationships under the title 'law-level causal asymmetries'), but it is clear that the restriction excludes an abundance other causal relationships from being covered by law which we might have thought plausible under the reduced-causation-mirroring conception. For example, although the behavioural formulae of instances of Schema 2 describe the numerical relationships between variable-properties, the instances themselves do not in any clear sense mirror the asymmetrical causal relationships which exist between them, e.g. the influence of force on acceleration or of temperature on pressure (in §5.2.3 I will discuss these relationships under the title 'variable-level causal asymmetries').

Regarding Principle 3, Schema 2 is simply inconsistent with it. By definition, the antecedent properties of instances of this schema are system-types, and the consequents are behaviour-types. We could not, therefore, hope to provide chains of law-based explanation with these kinds of generalisation, since that would require there to be laws in which these property-types appear in the wrong side of the conditional.

The limitations of Schema 2 might appear undermotivated. But they seem exactly right when we consider the sorts of inferences laws actually allow us to make. The reason can be expressed as a dilemma for the choice of Schema

1 (in comparison with Schema 2). On one horn, if a conditional can be employed in genetic inference, then we have to admit that the antecedent and consequent properties are of the same type. But by admitting they are of the same type, we loose the ability to locate an asymmetry adequate for covering causal relations. On the other horn, if we restrict a conditional to having a distinction in ontological category of antecedent and consequent properties, we retain an adequate asymmetry, but loose the ability to employ the conditional in genetic inference.

Let's consider this dilemma more slowly. Unlike the contast between laws' system-types and consequent the behavioural properties, the variable-properties which feature in laws' behavioural formulae (e.g. force, acceleration, charge, flux) are in some relevant sense of the same ontological category. This means that the explanatory relationships (arguably causal, see 5.2.3) which exist between these properties are just the sort that can be chained together in genetic inferences. For instance, we might use a law to explain why something experienced a force on an object by reference to the electrical field at its location, then we might use a different law to explain the change in momentum of the object by reference to the force it experienced. We can chain together the nomological explanations so that we can explain the change in momentum by reference to the field.

According to the reduced-causation-mirroring conception, we can perform such genetic explanations because there are laws which cover each of the stages. There are laws of the following form:

> If the combined field of force at an object's location is X then the total force it experiences will be Y.
>
> If the total force experienced by an object is Y then it will have a change in momentum Z.

These two conditionals can then be chained together to form the following genetic inference:

> If the combined field of force at an object's location is X then it will have a change in momentum Z.

The problem is, however, that we seem to know by law that the reverse inferences hold too. If something has a certain change in momentum Z, then we also know it will be experiencing a total force Y and likewise, if something is experiencing a total force Y, then there will be a combined field of force X at its location. The problem harks back to the same old problem of the flagpole and its shadow raised against the D-N model of explanation (see Bromberger 1966). Ultimately, variables related by equations in laws can be manipulated to allow us to infer any value given the values of the others—this is what makes them *laws of association*. There is, therefore, no asymmetry in the explicit logical inferences licensed by laws among their variables.

This is an issue for Principle 2, since there appear to be explanatory relationships (e.g. between field of force and total force, or that between force and change in momentum) which are not covered by laws in an unambiguous way. There may exist conditionals following from known laws which mirror the explanatory direction (like those above), but there will also be conditionals which go in the exact opposite direction of inference. Given this, it is simply misleading to say that the laws 'cover' such explanatory relationships, at least if covering is supposed to have any kind of analytical power when it comes to justifying inference to an explanatory asymmetry.

We might try to avoid this problem by looking for a further asymmetry (beyond causation and conditional inference) which exists between the two relata, but it is not easy to come up with anything we could feasibly take laws themselves to be representing. An interpretation of laws as stating conditional probabilistic relations will make no difference; for example, total force is dependent on change in momentum just as change in momentum is dependent on total force. In general, when we know the direction of inference between variable properties goes one way, we will have no reason *not* to posit further conditionals going the other way as well. Supporters of Schema 1 might suggest that while force and change in momentum are truth-functionally symmetrical, they may not be subjunctively symmetrical, i.e., while we should accept inferences such as,

> if total force been $1N$ then change in momentum would have been $1kgms^{-2}$,

we should not accept inferences of the form,

> if change in momentum had been $1kgms^{-2}$ then total force would have been $1N$.

But this is not a promising suggestion. Notice first that whether one accepts an asymmetry in truth of the above subjunctive conditionals will come down to what one's views are concerning nomic preservation. In §2.2 I suggested we take counterfactual preservation as a useful diagnosis of laws: laws remain laws under any counterfactual supposition consistent with all the laws. If that's correct, then we should not expect a difference in truth-value between the two subjunctive conditionals above, since the antecedent of both subjunctives is consistent with the total force law and the consequents follow by that very law. But regardless of one's views about nomic preservation, it is almost unanimously agreed upon that the truth of subjunctives is something to be explained by laws rather than something which explains them. Hence, if there is an asymmetry here to be found, we should be able to use the laws themselves absent of explicit subjunctive language to justify it. Consequently, the recourse to subjunctive conditionals is a non-starter, unless one is willing to treat such conditionals as primitive (or, at least, more fundamental than laws), and this is something few are willing to admit (*pace* Lange 2009) and decidedly unHumean (see §2.2).

To reestablish the asymmetry, a different tactic entirely—this is the move that shifts us to the dilemma's second horn—is to look for intrinsic differences among the relevant relata. Given such intrinsic differences, it would not matter if our knowledge of the laws provides us with symmetrical truth-preserving inferences, since there is a clear asymmetry in the ontological category of property ascribed, just as there is between system-type and behaviour. But if we restrict ourselves to laws which have antecedent properties of one intrinsic characteristic and consequents with another (as instances of Schema 2 seem to), we thereby prohibit genetic explanation, since antecedent properties will be intrinsically incapable of being consequent properties and vice versa.

It should now be clearer where the dilemma for supporters for the reduced-causation-mirroring conception emerges. On the one hand, if laws relate properties with no intrinsic differences so that laws can be chained to form inferences of one property-instance with another then there will be no apparent asymmetry to establish that the inference is explanatory. On the other hand, if laws relate properties with intrinsic differences, then the explanatory asymmetry is reinstated, but there is no possibility of chaining. Genetic explanation in the way Hempel described and in the way that seems necessary to cover indirect causal relations does not, therefore, seem possible with laws. This is exactly as Schema 2 predicts.

4.1.3 On causation-mirroring

If we are to get a better conception of the relationship between causation and laws of nature, I suggest we hold on to Schema 2 ($\forall x(Sx \rightarrow Bx)$), which has a single object variable, lacks a temporal variable and captures an intrinsic asymmetry between antecedent system-types and consequent behavioural properties. In looking ahead to the prospects for a nomological analysis of causation, I draw the following morals from the foregoing discussion.

First, the asymmetry of many causally explanatory relationships won't be reflected in the logical structure of laws, i.e. not all causal relationships are *covered* by laws. Although the asymmetry between system-type and behaviour may be determined by intrinsic differences in ontological category (I will, in fact, suggest a more precise distinction in §6.2.5), the asymmetry of explanation of certain variables in terms of others will have to determined in some other manner. In §6.2 I will suggest a probabilistic analysis to do just that. Under this analysis, it will not be the laws themselves which bear the structure of explanatory asymmetries between variable-properties, but certain causal and probabilistic information with which they are inferentially connected.

Second, any inferences laws license from causes to effects won't be one which extends beyond the systems in which the causes occur. This will significantly limit the extent to which a single instance of a law can connect a cause and effect. In §7.2.1 I refer to this connection as 'intra-system causation'. It will form the core relation with which the analysis of token causation is developed.

Third, given the issues concerning the need for behaviour-modification of many laws, we should not expect laws to support a logical entailment or even a precisely statistical inference of effects (under some description) from their causes (under some description). Again, the analysis will have to proceed in a different way. Accordingly, the core concept of 'intra-system causation' will pair events within systems which instantiate the system-type of a law in way which doesn't rely on such entailment or inference but rather on the type-level causal information associated with the laws and the particular systems instantiated.

Fourth, if indirect token causal relations are to be analysed by laws and the systems described by them, we should not hope to do so by chaining laws together as with genetic inference. Instead, in §7.2.4, I will suggest we chain *systems* together. Again, this will not be a method of showing how the cause entails or statistically implies the effect, but is rather 'connected' to it due to the systems present in the local environment.

In general, the observations made here show us that whatever procedure we use to extract an account of causation from laws will have to be rather more subtle than that presupposed by the causation-mirroring conception of laws. That is because laws in fact have a logical form which precludes certain inferences required by that conception. In §4.2 I want to show why this causes problems for Humeanism's most popular account: the *best system account* (BSA) of laws. For I want to show that BSA is implicitly committed to something like the causation-mirroring conception of laws.

4.2 The best system account of laws (BSA)

The Humean should not offer an account which contradicts their own methodology. Consequently, it seems inevitable that they propose regularities in the mosaic as the determinants of laws. For although it is coherent to hold onto a regularity-based account of laws of nature if one doesn't restrict the regularities to those in the mosaic (see, e.g., Tahko 2015, Demarest 2017), if one does make such a restriction then regularities seem to be about the only worldly resource one can go on.

Granting this assumption, the obvious follow-up is 'which regularities?'. To answer this, the current trend is a so-called 'best system account' (BSA), which determines the laws as those generalisations which are theorems in that deductive system which best captures the world's history according to an overall balance of simplicity and comprehensiveness—or 'strength' (Braddon-Mitchell 2001, Cohen and Callender 2009, Earman 1986, Lewis 1973b, 1983, 1994, Loewer 1996, Schrenk 2007, 2014). (In order to accommodate probabilistic laws a further parameter of 'fit' is sometimes included see especially Lewis 1994, but nothing I say in this chapter will hang on its inclusion.)

One thing to note immediately: the notion of 'system' at play here is not that invoked in §3 to refer to the instances of laws' antecedents. The former is a 'deductively closed, axiomatisable set of true sentences' (Lewis 1973b, 74), not to be confused with the latter, a set of physical assemblies. As Lewis has pointed out, BSA certainly seems to have an answer to some crucial characteristics desirable for an account of laws. For instance, it shows how laws are not individuated by logic alone, it allows for laws to be contingent and mind-independent, it suggests why lawhood might seem a difficult and vague concept (given the vagueness of the parameters of strength and simplicity), and it suggests why many theorems of scientifically confirmed theories are reasonably held to be laws also.

BSA has been incredibly popular among Humeans, although it has not been without its critiques. For instance, Roberts (2008, §1.6), and more recently Woodward (2013), have raised a number of queries about the scientific legitimacy of a balance between the parameters of strength and simplicity in determining the laws. One issue is that judging by much of scientific endeavour, it can seem that strength has a priority over simplicity. The following quote from Einstein provides an example.

> It can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience. (1934, 165; quoted in Woodward 2013, 12)

This would make it seem that strength acts more as a threshold, and simplicity only as an afterthought. Indeed, it is argued that simplicity only really enters into practice to determine *whether* a generalisation is true, for instance, in utilising the results of Akaike (1973) in curve-fitting problems, rather than in order to settle which among the true generalisations deserves the title 'law'. In the terms of Reichenbach (1958), it is 'inductive simplicity' which scientists make use of, not 'descriptive simplicity'.

It seems to me that the supporter of BSA has responses to these concerns. While theoretical physics aims to concern a very broad range of things, the scientific enterprises from which biological, economical, ecological, chemical laws emerge appear to consider strength much less of a lexical priority over whatever other parameters are at play. The fact that strength plays such a big role in physics might be put down to the fact that the world has been kind enough to let it do so.[2] As Woodward himself points out, our best candidates for strict physical laws (e.g. those represented by Maxwell's equations, Einstein's field equations, Schrödinger's wave equation) are both incredibly simple and strong.

On the issue of simplicity, I think it is open to a supporter of BSA to accept the 'inductive' use of simplicity in curve-fitting problems and to suggest that descriptive simplicity plays a role (if not explicit) as well. Indeed, one might think the two uses are closely related. When a choice of curve through some data is made on the grounds of simplicity, we may believe (perhaps justified by Akaike's result) that this will lead to a theory or hypothesis more likely to be predictively accurate in the future. This is an example of inductive simplicity playing a vital theoretical role. But in dispensing with the less simple curves we are also making choice which has repercussions for the descriptive simplicity of our adopted theories. After all, we may know that many (all?) theories rejected on inductive grounds could be revised in various ad hoc ways to accurately predict any future data. So the choice of a simpler curve for inductive reasons is by default a choice to pursue theories which are *descriptively* simpler too. In this way the complaint with BSA that scientific practice is concerned exclusively with inductive simplicity loses its force, since inductive simplicity cannot play a role in science without the latter doing so too.

Despite the intuitions that foregoing criticisms of BSA are inadequate, I do

[2]Although, if idealisation in physics is as widespread as Cartwright (1983) has suggested, we might have to recognise this strength as attained only through some universality-saving device, e.g. behaviour-modification (see §3.3.3).

believe that it is a seriously flawed account of laws. Given what appears to be a necessary adjustment to the account recommended by Woodward, it becomes clear that BSA is implicitly committed to the highly problematic causation-mirroring conception of laws. Once this conception of laws is removed, the parameters provided by BSA seem woefully inadequate to give us the laws. As a consequence, I think Humeans have a duty to look for an alternative account.

4.2.1 BSA and causation-mirroring

Consider again the notion of strength. At first glance it might be thought best cashed out in terms of deducibility of content about the Humean mosaic: a system is stronger if it entails more information about the Humean mosaic. This interpretation can seem inevitable from Ramsey's thought—a precursor to BSA—that laws are,

> consequences of those propositions which we should take as axioms if we knew *everything* and organised it as simply as possible in a deductive system (1978 [1928], 138, my emphasis).

This understanding of strength turns out to be problematic. If BSA is to be extensionally adequate, it should assure that non-lawlike generalisations, such as that all planets orbit in the same direction or that all gold spheres are smaller than a mile in diameter, do not come out as theorems. But in order to remove all possibility of such derivation, a deductive system will have to omit a significant proportion of initial conditions from its axioms. Otherwise it is plausible that in conjunction with uncontroversial laws, non-lawlike generalisations may be inferred. However, if we remove *all* initial conditions, we lose entirely the possibility of *any* deducibility of content. After all, if our laws have the form of conditionals then they say nothing unconditional about what happens and the strength of a system entirely comprised of them is zero.

As Woodward suggests,

> one way out of this difficulty is to give up on the idea that 'strength' has to do just with what can be deduced about the [mosaic] from the [best system] by itself. Here is an alternative proposal: the strength of a systemization is (among other things but perhaps primarily) a function of the extent to which claims in the systemization can be combined with some facts reported by the [mosaic] to deduce other facts in the [mosaic]; the more such facts in the [mosaic] that can be so deduced from a 'small number' of other facts in the [mosaic], along with premises in the [best system], the stronger this systemization. (2013, 9-10)

Perhaps Woodward's proposal would avoid the problem of adequate strength. But whether it is accepted by the BSA theorist or not, the suggestion points towards a conception of laws which seems to have been implicit all along: that they have the sort of structural form which will license inferences from facts about one part of the mosaic to facts about another, i.e. across time and space. That this was Ramsey's idea is indicative from his explicit pursuit of an account of laws which can explain our causal inferences. For instance, he remarks that,

> the world [...] exhibits as we must all agree a good deal of regularity of succession. I contend that over and above that it exhibits no feature called causal necessity, but that we make sentences called causal laws from which (i.e. having made which) we proceed to actions and prospositions connected with them in a certain way, and say that a fact asserted in a proposition which is an instance of causal law is a case of causal necessity.' (1978 [1928], 148)

Ramsey is here suggesting that our idea of causal necessity comes from what he calls 'causal laws', i.e. generalisations of regular succession of properties from which causality is inferred (not to be confused with causal laws in Cartwright's sense). But by invoking the notion of succession, Ramsey is clearly assuming that laws license cross-temporal and cross-individual inferences. At no point did Lewis try to distance himself from *this* idea of Ramsey's and no one I know of since has tried to offer any alternative.

But now we must recall that laws simply don't have the form of conditionals which 'can be combined with some facts reported in the mosaic to deduce other facts in the mosaic' *where those facts concern temporally and spatially distinct regions.* Such inferences are only licensed given boundary conditions *at* and *in between* both the region denoted in the antecedent and consequent of the inference. Why? Because laws don't mirror causal relations in that way. Rather, laws are conditional generalisations of the form of Schema 2, i.e. $\forall x(Sx \rightarrow Bx)$, and so the available deductions will in each case be restricted to facts about a single type of system.

Again, the point is not that laws can't be used to perform such inferences, but that they alone are insufficient to support them. For instance, the position of a projectile can be calculated knowing initial trajectories plus mechanical laws of motion which describe its dynamics at every point. But this is only possible under the assumption that, for example, the gravitational forces remain constant, that there is no unexpected interference which destroys the accelerating system, and that space-time structure remains conserved.

The kinds of inferences laws are sufficient for supporting are merely *local.* In the simplest sort of deduction, we learn that something is an instance of a certain system-type S and deduce that it behaves in manner B. Plausibly, the comprehensiveness of a system-type and behaviour would be valuable to a deductive system of laws, but this comprehensiveness is only tendentiously associated with a notion of strength conceived in terms of deducibility of content on the basis of initial conditions. The laws can't in general 'fill out' the details of the mosaic on the basis of conditions which specify *only* what the mosaic is like in some far off time or place.

This issue with how BSA treats laws has been noted before by Ismael (though not in exactly the same terms) in discussion of its compatibility with quantum statistical laws.

> Quantum mechanics, our central example of a theory of chance from physics, assigns probabilities to results of possible measurements, on the basis of the present state of a system and whatever external influences it is subject to, but on the standard interpretation one cannot at any given time assign probabilities to events that occur in the non-immediate future without knowing the results of all intervening measurements. (2008, 301; see also Glynn 2015)

Ismael is here aiming to criticise the BSA's notion of 'fit', but part of her point in so doing is exactly that BSA's expectation that the laws it outputs will license inferences from present (or past) to non-immediate futures is not consistent with the laws as we know them, specifically, those of quantum mechanics. As was predicted by the arguments in §3.1, the laws of quantum mechanics are temporally and spatially *local* to the systems they concern.

If BSA is to be plausible it must, therefore, include a conception of strength which is liable to return laws of the form of Schema 2 rather than Schema 6 (that of the causation-mirroring conception). There are at least two problems with this which seem to me so problematic that Humeans should look for further parameters beyond strength and simplicity (and fit).

First, it is unclear how can a best deductive system balancing simplicity with strength interpreted as one of deducibility of content can be assured to avoid axioms and theorems which do not the form of Schema 6. While no law that we know of actually has this form nothing in the way BSA is described would seem to prevent one appearing in the best system. One predictable response is that the world just turned out to be such that our laws are all localised in the way that they are (having no distinct time-index for antecedent and consequent and only one object-variable). But this can seem incredible. For doesn't it seem plausible that once the history of the world is up, the deductive system which best compromises strength, simplicity and fit is bound to return some generalisations which are sufficient to license inferences across time and space? If so, BSA is committed to such generalisations being laws.

A supporter of BSA might consistently accept that there will be such laws, admitting that we haven't discovered them yet. But I do not think this is a reasonable stance. For it encourages the supporter of BSA to strike a division between what the laws are and the kinds of generalisations we are liable to posit as laws. Something about our epistemic capabilities means that the generalisations we induce on the basis of evidence are of the form of Schema 2. What good is it to us if the 'true laws' turn out to have a very different form? Does this mean that our inferential practices, our modal reasoning, our justifications for our interventions are misguided? Such an outcome appears beyond what a Humean should want to admit. Therefore, there must either be an additional parameter to compete with simplicity and strength (and fit) for best system which limits the system's generalisations to same-time, same-individual inferences, or else an additional constraint on which generalisations count as laws from the system determined by strength and simplicity (and fit) alone. Both options take us beyond BSA.

Second, even granting that the BSA can return us a set of generalisations limited to same-time, same-individual inferences, it is another further requirement that the generalisations be of the form Schema 2. For inferences of this latter form have extra constraints on the kinds of properties which feature in their antecedent and consequent. Given the meagre set of parameters recommended under BSA (even including whatever further criterion might limit the returned generalisations to same-time, same-individual) it seems implausible that the outputted generalisations will inevitably denote *system-types* in their antecedent and *behaviours* in their consequent.

In fact, the problem is even worse. We know that laws are supposed to contribute to the justification of our goal-directed actions (see §2.4). Hence, the outputted laws should not only have the form of Schema 2, but have a system-type and behaviour which are actually *useful* for such purposes. Although I have not gone into the details of what this might amount to, I think it is clear that BSA will need some significant further qualification before it is able to assure us of *this*.

In sum, it is very difficult to make sense of the trade-off between simplicity and strength as playing the central role in accounting for what makes laws important. BSA has typically been presented as an account in which strong laws license inferences across time and individuals. Once we reject this causation-mirroring conception of laws, the interpretation of BSA must be rejected too. But it is entirely unclear how the notion of strength, simplicity and fit can alone provide the required parameters to give us laws in the form they really have, i.e. of the form of Schema 2. While there is scope for a trade-off between simplicity and strength to play some role in determining the laws, this is certainly not all there is to scientific reasoning about laws. Indeed, there clearly seem to be other assumptions in play when a generalisation is in the running to be deemed a law, such as the fact that laws have a system-predicate in the antecedent, a behaviour-predicate in the consequent, and a form utilisable for justifying goal-directed actions.

It is curious that the diagnosis of BSA should put so much stress on a misconceived causal conception of laws. After all, many Humeans have been fairly dismissive of the importance of causation in physics and certainly in the formulation of their laws. But if I am right, the long shadow cast by the causation-mirroring conception has found its way into the very essence of

BSA. Once removed, the account is unrecoverable. In §5, I propose an account of laws which assures us of the kinds of aspects of lawhood we are after by explicitly adopting causal parameters.

4.2.2 Diagnosis of strength and simplicity

If BSA is really committed to the causation-mirroring conception of laws, and that conception must be rejected, then we must reject BSA along with it. Or at least we must reject the sparse form it is typically presented in, since the parameters of strength and simplicity (and fit) cannot alone do the work of determining the laws.

But does that mean there is no role to play for strength and simplicity at all? The account of laws to be developed in §5 and §6 will not make use of them. Indeed, according to the behaviour-modification of idealising laws (see §3.3.3), it is possible that there may be—though this is not essential—incredibly complex laws which have only a few instances from which many more abundantly instanced idealising laws are derived. This might seem strange given the enormous plausibility that somehow strength and simplicity shape our concept of law. To placate these worries it can be pointed out that perhaps what we have in mind when we put such importance in simplicity and strength are the properties of idealising laws. Given the behaviour-modification approach to universality (see §3.3.3) we can expect of any law whose behaviour-predicate has few contexts of precise application that there will be idealising laws whose behaviour-predicates are defined in terms of special cases of that complex behaviour. Such modified predicates plausibly have a much wider application and will be mathematically simpler than the complex laws from which they derive their lawhood. The more variables are held fixed in providing a special case, the more widespread the idealisation becomes. So we might expect that a significant proportion of the laws are both strong and fairly simple, as a corollary of behaviour-modification. Nonetheless, the account I will propose does not use these features as *definitive* characteristics of law.

In denying a central role for strength, simplicity or the causation-mirroring conception I do not assume that we should reject entirely the assumption that laws have *some* valid causal conception. Indeed, as I remarked in §1, I do think laws are causally conceived somehow. §5 is all about getting that conception right.

5 The causal-junctions conception of laws

In §4 I argued against a certain causal conception of laws, viz. the causation-mirroring conception. Nonetheless, like Helmholtz, I submit that laws are causally conceived *somehow* and in this chapter I will argue for this in two ways.

The first is the subject of §5.1 and involves rebutting certain classic concerns with the relevance of the concept causation in physics. These concerns might equally be levelled against the causation-mirroring conception of laws, but now I have presented different reasons to reject that view, the need to discuss these concerns becomes relevant to my hope of *reinstating* a causal conception of laws. In responding to these issues I will draw attention, specifically, to how the notion of experimental intervention can lend support to a causal conception of laws.

The second way in which I provide support for the Helmholzian thought will take up the majority of this chapter and involves the development of a new 'causal-junctions conception' of laws which I believe to be both independently plausible and also to provide a key insight into a promising new Humean account of laws. I begin in §5.2 by proposing four causal features which I believe to be associated with many laws. In §5.3 I then define a particular class of laws, the *robust causal junction laws*, which exhibit each of the four causal features in a particular way. In §5.4 I then work through some examples of laws which do not fall into this class and show how they are inferentially connected to those which do. I end this section, in §5.4.5, by defining the causal-junctions conception explicitly. In §5.5 I then propose the conception be modified into a causal junctions *account* of laws, drawing attention to a number of its merits. While the account may be adopted by causal primitivists as it is, a Humean who accepts it will have the duty of analysing the causal features which characterise it; this is the task of §6.

Although a causal-conception implies some kind of conceptual priority of causation over laws, by saying that laws are causally conceived I certainly don't wish to claim anything about the *meaning* of lawhood or individual laws. After all, with the end goal of providing a non-causal analysis of causation in terms of laws, the laws I am interested in must be contrasted with what Cartwright (1979) has called 'causal laws', i.e. laws which have the word 'cause' or some surrogate right in them. Moreover, as I use the term, a 'conception' should not entail any metaphysical or epistemic priority of causation over laws or vice versa nor any implication that the two cannot be provided a Humean account.

What I do intend by the idea that laws are causally conceived is that knowledge of laws is *inferentially connected* with causal knowledge. There are systematic and truth-preserving ways in which we can and in fact *do* move from the understanding that a law applies in some context to the understanding that there exist certain causal relationships. Moreover, and this is where the basis for the new account of laws comes in, there are systematic and truth-preserving ways in which we can and do move from the understanding that there are certain causal relationships in a context to the understanding that a law is applicable. Hence, drawing attention to this causal conception will give us an answer to the question raised at the end of §3 concerning what makes laws special (as opposed to accidental generalisations and generalisations about mere assemblies of objects).

Obviously, establishing this may seem bound to have metaphysical implications. For instance, it can't be that laws are causally conceived and yet the correct analysis for causation (reductive or otherwise) be completely independent of the correct account of laws (reductive or otherwise). In §5.1 I will rebut some of the classic complaints that certain areas of physics and physical laws are devoid of causal import. But none of this should imply anything about the direction of metaphysical priority between causation and laws. Indeed, by framing the discussion in terms of 'systematic and truth-preserving inferential ties' from understanding of one to the other I am consciously trying to avoid committing to a direction of metaphysical priority. The purpose of this is to show how much a common basis for laws can be achieved without assuming Humeanism or otherwise. Even the hardened Humean should appreciate that, since any Humean account of laws built on such a conception would inherit much of its plausibility.

5.1 Causation in physics

Since the late 19^{th} century, there has been a trend of dissatisfaction with the concept of causation as a useful concept, particularly within physics. Physics is certainly not the only science whose laws I think should be understood as causally conceived (though, of course, not in the way suggested by the causation-*mirroring* conception). However, enough of the phenomena I ultimately want draw upon in causal analysis are from physics that I think failure to accommodate it within my overall claims about causation and laws of nature would severely undermine the plausibility of my arguments.

Many of the early complaints with causation as a useful concept in physics were addressed to very particular conceptions of causation (I take Frisch 2014, for my lead in this discussion). For instance, Kirchhoff (1876) complained of the inherent vagueness of metaphorical associations with the concept of causation in analysis of force. Mach (1900, 1905) argued (and the same point appears in Russell 1912) that the idea of 'similar causes' being related to 'similar effects' is either too indeterminate, imprecise and perspective-relative or else so specific that the putative lawlike regularity could not occur more than once. As a result, Mach concluded that the concept of causation should be replaced in science by the more precise idea of functional dependence. Russell (1912) argued furthermore that the inherent asymmetry of the causal relation has no place in physics, which is predominantly time-symmetric (see Norton 2009 for a more recent version of this complaint). Another common issue which emerged with the dawn of quantum mechanics concerned the plausiblity of the so-called 'principle' or 'law of causality' (e.g. Schrödinger 1951) that "Everything that occurs, has one or more causes, which together necessarily lead to the event in question" (Planck 1937, 83; translation in Frisch 2014, 13). Under its usual interpretation, the indeterminism of quantum processes seem to undermine this principle.

While these complaints may well serve to undermine certain interpretations of causation, I do not think they serve to reject entirely the relevance of the causal concept from physics. We now have at our disposal a plethora of theories of causation which render it precise and unmetaphorical (to cite a few: Lewis 1973a, Salmon 1977, Fair 1979, Mackie 1980, Glennan 1996, Dowe 2000, Pearl 2000, Schaffer 2001, Woodward 2003). All of these theories appear to accommodate the distinction between a total cause and the particular causes picked out for perspectival reasons, and they can each define conditions for causation which do not suffer from the failure of basic regularity accounts in being caught between uninformative vagueness and too much specificity. Indeed, it might be that behind each theory is a method for interpreting the functional dependencies specified in physics as causal dependencies in their own terms (see Frisch 2014, 112). Furthermore, none of these theories demands that causation be deterministic—the principle of causality is not intrinsic to causation according to them (see, especially Anscombe 1971).

Admittedly, it is typical for theories of causation to make use of a criterion of asymmetry (not necessarily temporal) in order to distinguish causal attribution from effect attribution, but this need not conflict with the fact that the basic laws of physics are represented by formulae which themselves provide no grounds for inferring asymmetry. Responding to Van Fraassen (1993), Frisch points out that 'we cannot conclude from the fact that an *uninterpreted* formula $\mathbf{F} = m\mathbf{a}$ does not on its own mark $\mathbf{F}$ as a cause and $\mathbf{a}$ as effect—that the causal "distinction is made outside the theory"' (2014, 114). Indeed, it can seem that the attribution of cause to force and effect to acceleration is exactly the interpretation we should be providing as part of the physical theory incorporating that formula. As physicist Gustav Fechner pointed out 'physicists often speak of force simply as the cause of motion' (1864, 126, translation by Frisch 2014, 2; see also Suppes 1970 and Hitchcock 2007a).

Moreover, the fact that laws might allow the past to be *inferable* from the future in the same way they allow the future to be inferred from the past (given other assumptions; see §4.1) is not enough, in itself, to prevent a meaningful *causal* asymmetry playing a role in physical theory. Nor is it enough to show that dynamical equations *void of causal information* are physically indistinguishable from their equivalent but time-reversed analogues (as opposed to what the arguments in Norton 2009, seem to suggest). What would be needed to show *this* is an argument to the effect that positing causal asymmetries are always unjustified when considering time-reversal invariant laws.

Although there do exist such arguments (see, e.g., Price and Weslake 2009) to my mind, (Frisch 2014, especially ch.6) has successfully defended the importance of causal knowledge in physics. Two key points lead him to this conclusion. First, many conclusions reached in experiment rely heavily on the principle of common-cause: that when correlated variables are not related as cause to effect, there will be a common causal variable of both. While this principle may, on rare occasions, give unsound advice (see Arntzenius 2010), it appears in many cases to be this very principle which plays a crucial role in the formation of successful hypotheses when the available evidence provides no reason to prefer one dynamical model (void of causal information) from another.

As an example, Frisch considers the inferences we would make based on a series of observations of points of light at the same observed location in the night-sky. The non-causal dynamical models derived from Maxwell and Lorentz combined with the wave equations and the available evidence are insufficient to confirm that the various points of light can be attributed to a single source, viz. a particular star, rather than, say, multiple light-sources coming in from infinity. Nonetheless, we typically do make such an attribution. Frisch argues, I think successfully, that the only way we can plausibly infer the existence of such common-causes is with an active principle like the principle of common-cause, thereby justifying the role of causal knowledge in physics.

The principle of common-cause is shown to be particularly useful in cases where the available evidence is limited (in the above case, we do not have access to the field values across the entire future-surface of each light-source's light-cone). But Frisch's second point is that employment of the principle of common-cause itself serves to support an explanatory assumption that is recognisable even when data is not lacking. For a strictly non-causal view cannot support the intuition that, for example, while processes in which broadcast waves emit from antennas can seem perfectly normal, a time-reversal of this process in the form of 'anti-broadcast waves coming in from spatial infinity and collapsing on the antenna' would seem 'nearly miraculous' (the example is from Earman 2011). This intuition—shared by both supporters and naysayers of causation in physics alike—suggests that while we have reason to think the future will bring ever more dependence among events, our past recedes into occurrences that are ever more independent of each other. The principle of common-cause may seem to explain such an asymmetry of dependence, since it implies that dependencies between variables can be screened off by their (past) causes rather than anything futurewards. Here again causal knowledge is playing a crucial role in making sense of physical phenomena.

Frisch has shown that just because an uninterpreted formula doesn't mark some variable, e.g. of force, as causally relevant to acceleration this doesn't entail that the causal distinction is made outside theory. However, accepting this alone will not be enough to persuade everyone that causal distinctions are part of the very conception of physical laws—we are still a long way off being able to justify the claim that laws *themselves* are causally conceived. One may, for instance, be reminded of Russell's sceptical observation that,

> in the motions of mutually gravitating bodies, there is nothing that can be called a cause, and nothing that can be called an effect; there is merely a formula. Certain differential equations can be found, which hold at every instant for every particle of the system, and which, given the configuration and velocities at one instant, or the configurations at two instants, render the configuration at any other earlier or later instant theoretically calculable. (1912, 14)

In §5.2 I will show explicitly how, on a plausible interpretation of Russell, he was wrong to say this. There are, in fact, at least four distinct causal features of many laws including the law for mutually gravitating bodies. But we need not wait until then to make it plausible that laws are often conceived of causally in some sense. One very clear way to see how laws come to be conceived causally is in the process of intervention. Consider, for instance, the interventions possible with an air-pump of the sort used by Boyle. The gas in a sealed chamber can be diminished in volume under a load or its temperature increased by a heat source and its pressure measured with a gauge. By experimentally intervening in this way the causal relationships between volume, temperature and pressure are revealed, since a variable intervened on (if done so correctly) will be a cause of any other variable which changes with it. The interventions of this sort are just those used by Boyle to infer his law. By carefully measuring the volume and pressures as they changed under various loads, Boyle could infer that volume and pressure were inversely proportional. It seems inevitable, then, that this law would be causally conceived as a consequence.

Admittedly, the ideal interventions of the sort recommended by Woodward (2003) may not always be possible in order to prove the existence of a causal relationship between the variables in laws' behaviour-predicates. In particular, it has been questioned whether interventions should be conceived of as having to be ideal according to the standard Woodward originally suggested (e.g. by Eberhardt and Scheines 2007, Frisch 2014 and even by Woodward himself see Woodward 2014, 2015, 3583-3584).[1] But it is not my concern to validate Woodward's—or, indeed, any—account of causal manipulation via intervention here. Rather, my concern is to make plausible the fact that knowledge of many laws is closely associated with knowledge of causal information. Once we help ourselves to the notion of an intervention, understood in some intuitive sense, the close association between causal and law-based knowledge is easily stated: that we know the discovery of many laws including Boyle's and Charles's laws, Galileo's laws of freefall and Coloumb's law of electrostatics were the result of careful experimental interventions. (Although I do not say this is *all* that was involved). Since successful experimental interventions are a paradigm way to reveal relationships of causal influence, it is hard to conceive how the derivation of such laws by experimental interventions could avoid the derivation of causal information as well. When interventions are used to discover laws, causal knowledge comes *as part of the very method* of learning the laws; the association of laws with causal information is, in this sense, unavoidable.

Some of the laws we know, however, were not inferred from experimental

[1] For example, interventions may need to be, and in some cases are to be preferred when, 'soft' in the sense that they do not break the connection between original causal influences and the variable intervened on (see Eberhardt and Scheines 2007). Also, it may be that in practice, a change in the values of variables cannot be performed by manipulating some single component, but rather by replacing it with another, as when Galileo observed the effect of mass on velocity with his inclined-plane experiment by exchanging one steel ball for another.

interventions at all; instead they may have been derived from purely observational data (e.g. Kepler's laws), from further principles (e.g. symmetry principles), or have been introduced as starting assumptions (e.g. Newton's first law). However, laws of this sort will still need testing. And where possible, the kinds of tests scientists will want to carry out will be interventions. So, for example, inspired by Newton's inverse-square law of gravitation, Coloumb (as did others) assumed that electrostatic forces would obey a similar relationship to distance. He did not infer the relationship from intervention, but he certainly tested it this way, by manipulating the distances between loadstones he was able to measure the torsion induced in his torsion-balances and so confirm a inverse-square relationship with force. Coloumb obviously felt able to infer what experimental interventions would enable him to confirm the relationship. The hypothesis therefore had associated with it certain inferences about what the causal relationships among the variables would be if the hypothesis were true. Without such an association, the experiment would never have been performed.

I take these observations to show that at least from the perspective of how we think about discovering and testing laws, they are intimately bound up with causal inferences. This lends support to the idea that there is *some* kind of causal conception going on with laws. The problem is to say what that is.

5.2 Four causal features of laws

In developing the new conception of laws, I will begin here by drawing attention to four causal features which I believe to be associated with many laws: two features of *dispositionality* and two features of *causal asymmetry*. Each of the four causal features of laws which contribute to laws' causal conception have in fact been touched upon at some stage in the foregoing sections. Dispositionality is divided into the dispositional properties of components (noted in §3.2.1) and the disposition of systems to *embed* their behaviour in larger more complex systems (noted in §3.3). Causal asymmetry is divided into the variable-level asymmetries among the variables of laws' behavioural formulae and the law-level asymmetries between system-type and behavioural properties themselves. Both asymmetries were mentioned in §4.1.2 in discussion of the possibility of genetic explanations with laws. Also, both forms of causal asymmetry are, I believe, implicit in the idea that laws can justify our goal-directed actions, first noted in §2.4 and just brought again to our attention in discussion of experimental interventions, §5.1.

The following discussion of each causal feature should serve to make the particular kind of inferential connection with laws clearer. In each case I will draw attention to the particularly valuable and unique form of utility each feature provides us with. This should serve to strengthen support for the idea that laws are systematically related to causal information. Consider an analogy: I show you instances of the same kind of device one after another and draw attention to the fact that they have features which make them particularly useful for banging sharp pieces of metal into wood. But unless you think there's a general reason for wanting to do the latter, you may remain unconvinced that all instances of such devices are able to perform this task. Then I point out that banging sharp pieces of metal into wood is peculiarly useful for building furniture. Wouldn't that then lend credence to the idea that instances of the device I showed you were in general *made* or *sought out* for that purpose? The particular value of the device is justified by the uniqueness of the utility of the activity it is employed to enable. Something similar goes for laws and causation: the idea that laws are inferentially connected with certain kinds of causal information is supported by the fact that that causal information is a peculiarly useful thing to know, e.g. for navigating, predicting, controlling and manipulating the world around us. If knowing certain laws is part of what it is to have that causal knowledge (as I will suggest) then we can see why those laws have the features they do.

5.2.1 Component-level dispositionality

The first causal feature of laws I will discuss is component-level dispositionality. This refers to the dispositional requirements components have in order that assemblies in which they feature satisfy the system-types laws are about.

As I will understand it, a dispositional property is one defined by its causal potential (although, that is not to say the property is *irreducibly* causal). For example, fragility is defined as a potential for being caused to break, irrascibility as a potential for being caused to get angry. These dispositions are typically (though there may be exceptions, see Vetter 2015) understood in terms not only of the behaviour they 'manifest', but also in terms of the stimulus (or 'trigger') required to cause that behaviour. So, for instance, something isn't fragile just because it has the potential to break under a hydraulic press nor is person irascible just because they have the potential to get angry from catastrophic events. In each case the stimulus must be correctly specified.[2]

I pointed out in §3.2.1 that laws often concern entities defined dispositionally. The classical pendulum law, for example, has among the conditions of componency which constitute its system-type, the requirement that the string be inextensible, the bob have mass and the pivot be frictionless. Arguably, the gravitational law applies to systems made up of components with the causal disposition of *mass* where to have mass is to the causal disposition to exhibit inertia under force and to accelerate proportionally with such forces. Recently Higgs' hypothesis that mass can be understood in terms of interaction with a certain kind of field—the 'Higgs field'—confirmed the attribution to mass of dispositionality. We might plausibly take to be dispositional many of the other variable-properties required for the construction of the systems with which laws (often idealising) are concerned,

[2]By treating dispositions as having stimulus and manifestation conditions I assume it safe to ignore the subtle translational issues which exist between so-called 'conventional dispositions' (e.g. being fragile) and 'canonical dispositions' (e.g. being disposed to break when struck; see Choi and Fara 2016).

e.g. being charged, being electrically conductive, being frictionless, being inextensible, being a predator, being an ideal gas (consult Table 3.1, p.92 for more example dispositions and their corresponding laws).

Perhaps someone might object to such interpretations and rather opt for an interpretation of these properties as categorical, i.e. having unconditional non-causal conditions of individuation (what are sometimes referred to as 'quiddities' see, e.g., Black 2000, Schaffer 2005a). However, I think such an interpretation would be wrong. Conductivity is essentially the property of being able to conduct electricity; mass is essentially the property of being able to interact with the Higgs field; being inextensible is essentially the property of being unable to be extended. These properties' dispositionality is reflected in the cross-cutting taxonomy of what we have found to have them. Conductivity is exhibited by metals, and in certain states also carbons, silicons, and many other 'doped' materials. Mass is exhibited by leptons, quarks, animals, planets. Inextensibility is exhibited by nothing at all, but many disparate materials come close under proportionate loads. Insofar as we recognise the world as constituted ultimately by natural categorical kinds, the dispositional properties which feature in system-types' conditions of assembly cut across the taxonomic structure of these categories in ways which suggest that what is crucial to dispositional roles is not the categorical nature of the substances which fulfill them but something consistent in the way they enter into causal relations. It is arguably for this reason that science remains open to new substances being found to behave in familiar ways: consider, for example, our discovery that gases can be conductive when ionised, the hypothesis that photons have mass, that dark matter or neutron stars are inextensible. If the properties I have suggested are dispositional are in fact categorical properties then this kind of openness would be hard to rationalise.

The observation of dispositions' cross-cutting nature resonates with some of Quine's remarks.

> The general [dispositional] idiom is programmatic; it plays a regulative rather than constitutive role. It forms families of terms on the basis not of structural or causal affinities, among the physical states or mechanisms that the terms refer to, but on the basis only of a sameness of style on our own part in earmarking those states or mechanisms. (1973, 11)

In Quine's terms, dispositional characterisations capture the phenomenal affordances of things rather than the deep physical explanations for them, that is why dispositions can be multiply realised by different substances and so be instantiated in ways which cut across categorical taxonomy. But Quine was also sceptical about the lasting utility of dispositions. He remarked that despite the dispositional idiom being practically 'indispensible',

> if I were trying to devise an ideal language for a finished theory of reality, or of any part of it, I would make no place in it for the general dispositional idiom. (Ibid.)

Pace Quine, I would argue for quite the opposite view.[3] Other than the pursuit of knowledge, scientific discovery is supposed to help us take control of our environment. What we want to know is how to make use of the natural world in order to do that. Dispositional knowledge is, therefore, enormously useful if, for instance, a material with the right disposition is proven to be undesireable for other purposes, e.g. if it is too expensive, difficult to source, is too heavy, brittle, has poor thermal properties, etc. It is not the underlying mechanism that we care about in such cases, but the disposition it affords us. Therefore, component-level dispositionality is both a widespread and an important feature of laws we should be happy to embrace.

[3] Quine's distaste for dispositions partly stemmed from his belief that an extensional analysis of them was impossible (see the discussion in Quine 1973, §3). I present my own extensional analysis of dispositions in §6.1.3.

5.2.2 Law-level dispositionality

Whereas component-level dispositionality concerns the dispositions of system-types' components, a kind of dispositionality involved with laws which enters at a different level concerns the *embeddability* of the instances of entire system-types. The embedded phenomenon intuition described in §3.3.1 was the idea that apparent failures of laws' universality can be understood as a result of systems being embedded in more complex environments. This very idea seems to me to be a dispositional one.

When we come across a complex situation for which a perfectly accurate law is either beyond our knowledge, our ability to solve, or our practical requirements, then we can often idealise the situation. We do this by abstracting away certain features in the circumstances for which we have a law (Weisberg 2007). The whole practice of idealisation therefore relies on the understanding that the behavioural relationships described by an idealising law *causally contribute* to the more complex phenomena going on in the real world in which the simple system-type described by the law is embedded (Corry 2009, Cartwright 2017).

Arguably, Newton's law of mutually gravitating bodies precisely describes the behaviour of no system in the world. All systems of mutually gravitating bodies experience gravitational force from all other such systems, and assuming (as seems plausible) that all large masses have some degree of non-neutral charge, there are further contributions to total force which means that the formula does not precisely apply. Nonetheless, we still want to say that instances of the simple two-body system-type mentioned in the gravitational law are embedded in those real-world cases such that they contributes to the overall behaviour. For how else, could it be so appropriate to use this formula as an idealisation? Similarly, the ideal gas law precisely describes the behaviour of no system in the world, since all real gases are non-ideal. Nonetheless, we can say that real-world gases in some sense have instances of the ideal gas system-type embedded in them, contributing to the overall real-world behaviour.

This common thought seems to be exactly the embedded phenomenon intuition. Instances of system-types mentioned by laws underlie and hence causally contribute the behaviour associated with them by law to whatever is the more complex behaviour in which they are found. In §3.3.3 I suggested that the truth of such laws might be retained by modifying their behaviour-predicates to the claim that the relevant formulae are special cases of more complex formulae attributable to instances of more complex system-types of which instances of the law's system-type are a part. My belief in that suggestion still stands. The new suggestion I'm making *here* is simply that the contribution that an underlying phenomenon provides as a special case is a result of a causal disposition: *embeddability*. Since behaviour-modification concerns both the behaviour of entire systems, embeddability is a *law-level disposition*.

If we understand the behaviour-modification to formally capture the intuition behind the disposition of embeddability, the criterion that an embedded system be a part of a larger system can be understood as stimulus for the disposition and the criterion that a formula be a special case of whatever behaviour accurately describes the system can be understood as its manifestation. The details of this analysis will be made more precise in §6.1.3 when I come to provide a Humean analysis of laws' dispositionality.

As with many dispositions, embeddability seems to be defined in terms of a behaviour which can be exhibited in degrees. Hüttemann (2004) usefully distinguished these 'continuously manifesting dispositions' from 'discretely manifesting dispositions'.[4] For instance, whereas a disposition like irrascibility can be thought to manifest in degrees depending on the context, the manifest behaviour of the disposition of fragility seems to be all or nothing, since something is either broken or it isn't. Embeddability seems to be of the first sort. When an embedded system only experiences a modicum

[4] Cartwright calls dispostions which have this continuous feature 'capacities' (Cartwright 1999).

of external disturbance, we can expect the idealising behaviour to be approximately accurate, when there is a lot of disturbance, the idealising behaviour may be almost entirely hidden.

As a consequence of this feature, knowledge of embeddability is incredibly useful to us. When we know the instances of some system-type are embeddable we can make certain crude predictions and interventions *despite* external interference. For instance, we know that despite interference from other bodies' masses and charges, a decrease in the distance between two bodies will greatly increase the force they experience, and that if we could increase the mass of one body, the force will be increased also—or, at least, such an intervention will contribute positively to any change in force. While this knowledge may not always help us make perfectly accurate predictions and interventions, it is often suitable enough (for engineering, navigation, forecasting, etc.). It can also be the groundwork for positing more accurate laws which take into account more phenomena.

5.2.3 Variable-level causal asymmetry

In the foregoing two sections, I have discussed two causal features of laws which both count as kinds of dispositionality. Now I want to consider a different kind of causal feature: causal asymmetry. As with dispositionality, I divide the types of causal asymmetry into two. This section deals with that which appears at the variable-level.

Symmetry is considered a paradigm feature of many laws, particularly in physics, and a big motivation behind many philosophers and physicists distaste for causation in physics. As far as the non-statistical laws of mechanics go processes are at least in some sense 'reversible' in time, meaning that the exact reverse of a physically possible temporally evolving process is also perfectly compatible with the mechanics.[5] Insofar as such

[5]There is a need for some delicacy surrounding what exactly is reversible; see, Albert (2000), who argues that we can only make sense of this in the language of 'states' rather

mechanics can be used to accurately predict temporal evolution forwards in time, they can be used to predict the 'backwards' direction as well.

In contrast with this, causation is famously asymmetrical: its instances comprise a cause and an effect, and at least typically, we don't suspect effects to cause their own causes. It can't be doubted, however, that in practice, *explanation* with laws only goes in certain directions as well. As supporters of the D-N model of explanation have been repeatedly reminded, we may predict the height of the sun in the sky by reference to the length of the shadows it casts, but we would not *explain* it that way—a lack of symmetry is reflected in the interventional possibilities that we know laws afford us. Laws themselves may look symmetrical, but when we learn to apply them, we do so in a way that reveals our knowledge that the height of the shadow can be manipulated by intervention on the height of the flagpole and not vice versa.

It's fairly obvious why asymmetry is useful. If an explanatory relation weren't asymmetrical then it would have symmetrical instances, and since the relata of explanation are distinct, this result in a circularity. Circularity is typically looked on as anathema to explanation. Specifically, problematic circularities develop when there is no way to break down the circle into multiple independent and asymmetrical explanatory generalisations, i.e. the circle of causally relevant variables is 'closed to intervention'. This would happen if, for example, we took the length of objects' shadows to always and exclusively explain and be explained by those objects' heights. If this were the case, it would seem impossible to have knowledge of how to use one of these phenomena to influence the other. We could not, for example, hope to change something's height by any means other than via its shadow-length, but similarly we could not hope to change something's shadow-length other than by its height. Such explanatory circles, if they exist at all, would be entirely useless for these purposes.

---

than dynamical conditions.

It is typical, though by no means ubiquitous, to think that the causal asymmetry can be mapped to temporal order, since causes always (or at least typically) precede their effects. But while the kind of asymmetry drawn attention to in the famous flagpole case is temporal, there are plenty of non-temporal kinds of asymmetry which allow us to reach similar conclusions. This seems to be especially the case with those causal asymmetries associated with the variables in laws' behavioural formulae. Take, for instance, the asymmetry which exists between force and acceleration in the total force law. If the mass of an object is known, then we can either use the total force to predict its acceleration or acceleration to predict total force. But we would only ever think to *explain* acceleration by citing the force, not vice versa. That this asymmetry is causal seems to be largely accepted among physicists (though perhaps not among philosophers). Recall Fechner's claim cited in §5.1 that 'physicists often speak of force simply as the cause of motion'. An example of this comes from the following statement from Albert Einstein and Leopold Infeld, quoted in (Cartwright 1979, 419).

> The action of an external force *changes* the velocity [...] such a force either increases or decreases the velocity according to whether it *acts* in the direction of motion or in the opposite direction.[6] (Einstein and Infeld 1971, 9, my emphasis)

The explanatory asymmetries in the total force law is instantaneous, so it is hard to see how an analyses in terms of a temporal (or entropic) order could hope to explain it. And it is certainly not the only law which exhibits such asymmetries. For instance, we will explain the current through a component in terms of its voltage and its resistance, but not typically the other way around.

[6] Curiously, Cartwright uses this to provide an example of a causal law which contrasts with her example law of association that force, mass and acceleration are functionally related. One might interpret the case I am making for variable-level asymmetry as an argument that laws of association are often conceived of *in terms of* causal laws. But I will not press this point in the main text.

Again, this asymmetry seems naturally captured as causal: we say that current is 'induced' by voltage potential and that resistors *resist* current. A typical physics textbook claims that,

> a potential difference $\Delta V = V_b - V_a$ maintained across the conductor *sets up* an electric field $\overrightarrow{\mathbf{E}}$, and this field *produces* a current I that is proportional to the potential difference.

And also that,

> most electric circuits use circuit elements called resistors *to control* the current in the various parts of the circuit. (Serway and Jewett 2006, 813, my emphasis)

If we take 'control', 'sets up' and 'produces' to indicate causal relationships, as seems natural, this would highlight why applications of Ohm's law exhibit the explanatory asymmetries that they do.

In the way considered in §5.1, our intuitions in this regard can be supported by what manipulations via explerimental interventions we know to be possible. Or we might simply be able to reflect on the how instances of the law are built to see the causal asymmetries. For instance, consider an electrical component c for which the equation in Ohm's law accurately applies. It is plausible that the following two equations also hold,

$$R = \frac{l}{\sigma A} \qquad\qquad \Delta V_c = \mathcal{E} - Ir$$

where l is the length of the component, σ the conductance of the component's material, A the component's cross-sectional area, $\mathcal{E}$ is the electromotive force produced across the terminals by the rest of the circuit and Ir is the internal resistance of the rest of the circuit. That these further equations also hold serves to justify the claim that current is reasonably treated as an effect of resistance and voltage-drop, since the causal relations between the values like l_c, σ_c, A_c, E_d and Ir on the one hand and current through the component on the other seem even more pronounced. Surely we would not explain the length of an electrical component by reference to the current flow through it! But we may indeed explain the latter by reference to the former: length explains resistance and resistance explains current. Hence, by mediating the explanation here, resistance shows itself more plausibly to be an explainer of current and not vice versa (something similar can also be said concerning voltage potential).

At least within the context of classical electronics, the relationships between voltage, resistance and current are all instantaneous. Now, it may be objected whether any causal relationships are truly instantaneous. After all, the edicts of Einstein's special relativity appear to prohibit causal signals passing faster than the speed of light. It is unclear whether this point alone is sufficient to prohibit instantaneous inter-spatial causality in *general* (mightn't we say that the next in line to the throne is instantaneously caused to become king or queen if the incumbent dies, regardless of whether they are on the other side of the universe or not?) But we needn't put any argumentative weight on this. It is more important to point out that relativity seems if anything to *imply* instantaneous causation if it is direct and *maximally proximate*, i.e. in the same place as. Huemer and Kovitz (2003) have argued for this point with respect to the Lorentz equation.

> a body with charge q moving at velocity v through electric and magnetic fields experiences a force given by
>
> $$\mathbf{F} = q\mathbf{E} + q\mathbf{v} \times \mathbf{B}$$
>
> where $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic field vectors at the body's current location in space. The vectors $\mathbf{E}, \mathbf{v}$ and $\mathbf{B}$ will typically vary over time, along with F, and the value of $\mathbf{F}$ at any

> given time is determined by the values of $\mathbf{E}, \mathbf{v}$ and $\mathbf{B}$ *at that time*. Bodies always experience the effect of the electromagnetic field *at their location in space and time*. (2003, 559, my emphasis)

Indeed, with many laws, the variables are assumed to be correlated instantaneously because they describe variable properties *at the same location*. Any law relating the force experienced in a field is a paradigm example. Furthermore, there are just the same asymmetries in explanation between such properties. For instance, the Lorentz force law is used to explain why certain forces are experienced due to the strength of electric and magnetic fields, velocity and strength of charge. In contrast, Lorentz-forces do not explain any of these latter variables. It is exactly these sorts of asymmetry I believe we have good reason to treat as causal.

Since the explanatory asymmetries discussed so far have been between variables which appear in the behaviour-predicate of laws (e.g. voltage, resistance, current in Ohm's law; force, charge, electric and magnetic field in the Lorentz-force law). I'll refer to these from now on as *variable-level asymmetries*. Notice that these asymmetries are just the sort we considered in discussion of the genetic principle in §4.1.2. There I said that the logical structure of the causation-mirroring conception could not cover explanatory relationships between variables because the variables appear in manipulable formulae. It is typically possible to isolate any variable we choose to determine one on the basis of the others. However, in pointing towards the systematic variable-level causal asymmetries in laws we seem to admit that the intuition that such explanatory relationships exist is a sound one. Indeed I think it is. But unlike causation-mirroring, the causal conception being developed in this chapter will not aim to have this asymmetry reflected in laws' logical form. There is, however, a different causal asymmetry which the conception will take laws' logical form to reflect. This is the subject of §5.2.4, to which I now turn.

5.2.4 Law-level causal asymmetry

Another kind of asymmetry associated with laws which is plausibly looked on as causal was intimated in §3.2.1 and also discussed in §4.1.2. This is the asymmetry which exists between a system and the behaviour it exhibits. We would not, for instance, expect to explain that something is an electrical conductor by the fact that the voltage across it, the current through it and it's resistance are related by the equation $V = IR$. Rather, the direction of explanation seems to go the other way around: being an electrical conductor explains why it obeys the formula. The relationship here is, like some of the examples considered above, a simultaneous one, since laws are conditionals stating that if some system-type is instantiated by something for some time then it behaves according to some behaviour-type for *that very time*. However, this asymmetry is worth distinguishing from that at the variable-level since it relates entities at the level of system-types and overal behaviour. I will therefore refer to it as *law-level asymmetry*.

In order to motivate further the existence of law-level asymmetry we can move from considering it as a single relationship between two relata (system-type and behaviour) as is reflected in the logical structure of laws according to Schema 2, and instead consider it as an an asymmetry between three relata: the behaviour and the two conditions of assembly (conditions of componency and conditions of organisation). Consider an example. If a heating element is underneath a hot-air balloon's opening and the balloon is rising, we have details of two components (burner and balloon), an organisation (one underneath the other) and a behaviour (rising). I take it that it would be wrong to say that the organisation and the behaviour explain why the components are as they are: the fact that one object is below another, and that the whole assembly is rising does not explain why one is a burner, the other a balloon. I also take it that it would be wrong to say that the components and behaviour explain the organisation: the fact that one object is a balloon and the other a heater and that the whole assembly is rising does not explain why one component is underneath the other. However, notice what happens when we look at the final combination. For it seems entirely plausible that the components and the organisation explain the behaviour: the fact that one component is burner, the other a balloon, and that the former is underneath the latter does very much seem to explain why the whole assembly is rising.

What this example effectively draws attention to is an explanatory asymmetry between each law's system-type (with individuating conditions of componency and organisation) and the corresponding behavioural property. So what, then, is the source of this asymmetry? I believe that there is good reason to treat such an explanatory asymmetry as causal. *Unlike* other forms of explanation, it isn't abstract or logically deductive, as mathematical explanation often is. Nor does the explanation seem adequately captured by the notion of constitutive explanation; the system doesn't constitute the behaviour but rather *performs* it. *Like* paradigm causal relations, the asymmetry can be used to assign responsibility (to broken components or unexpected systems), to answer what-if-things-had-been-different type questions, to develop strategies for intervention and control, to establish successful predictions and to raise the probability of certain outcomes.

Despite similarities with various contemporary analyses of causal explanation, this explanatory asymmetry seems to be rather underrepresented in philosophy of science. As I pointed out in §3.2.1, the notion of an assembly or system bears similarities with that of a mechanism and nomological machine. But the kind of explanation between conditions of assembly and a system's behaviour is not so obviously available to approaches in terms of the latter two phenomena (since behaviour of these entities is part of their identity criteria). Perhaps the main sticking point in regarding the relationship as causal seems to be that there is a salient difference in the ontological category between the relata. While typical causal relationships are taken to be that between events or property-instances what we have with conditions of assembly are the presence of types of components and the way they are organised.

In response to this sort of concern, three things should be noted. First, we can, if we so choose, represent the satisfaction of componency and organisation conditions as a value of a variable with a particular time-index. Therefore, there is no *formal reason* why we cannot treat an assembly of objects' satisfaction of these conditions as the right category of entity to enter into causal relations. Second, when it comes to conditions of organisation especially, there does not appear to be any deep difference with paradigm causal relata in the kind of property being instantiated. Organisation conditions (I have suggested) typically concern spatial arrangement, orientation and relative position. But such properties also feature in the behavioural formulae of laws. If we accept that variable-properties in laws' behavioural formulae can be causal relata, then we should accept the satisfaction of conditions of organisations can be also. Third, conditions of componency are typically dispositional. While dispositional conditions only require entities to exhibit some *potential* for causal relations, it seems reasonable to suspect that the conditions are part of those laws' system-types because they are required to manifest in order to bring about the relevant behaviour. For example, a pivot's frictionlessness must be manifest in any classical pendulum in order for the pendulum equation to apply, an object's mass must be manifest in any gravitational system in order for the gravitational equation to apply, and a species' predatory inclinations must be manifest in any habitat where there exists prey in order for the Lotka-Volterra equations to apply. Given this is the case in general, it seems entirely reasonable to assume that the satisfaction of conditions of componency play a causal role in bringing about systems' behaviour.

If law-level asymmetry is causal then the arrow in laws of nature can be interpreted causally *salva veritate*. Hence laws do cover one very specific form of causal relationship mentioned in §4.1: that between system-type instances and their behaviour. But to admit this is not to presume that laws explicitly state causal facts—the arrow in laws isn't to be read causally. Nor is it to presume that laws *analyse* the type-level causal relationship between system-type and behaviour (cf. the analysis proposed in §6.2.5). However, I do think that the existence of law-level asymmetry lends credence to the claim that causal relationships are part of how we conceive of laws.

5.2.5 Four ways Russell was wrong

Given the forgoing sections, let's return now to a point made by Russell quoted in §5.1.

> [I]n the motions of mutually gravitating bodies, there is nothing that can be called a cause, and nothing that can be called an effect; there is merely a formula. Certain differential equations can be found, which hold at every instant for every particle of the system, and which, given the configuration and velocities at one instant, or the configurations at two instants, render the configuration at any other earlier or later instant theoretically calculable. (1912, 14)

In a sense we must agree with Russell. Take one relevant component of the formula in Russell's example, Newton's gravitational equation:

$$\mathbf{F} = G\frac{m_1 m_2}{r^2}$$

This clearly makes no explicit causal claims. The values taken by the variables are simultaneous (at least within Newtonian theory) and can be manipulated to isolate any variable we so choose. We know the formula systems applies to systems exactly when those systems comprise two massive objects separated by a non-zero distance, i.e. the following conditional is an (unmodified) law.

> **Gravitational law:** For all x, if x is a two-mass system separated by a non-zero distance, then the force F experienced by each mass,

the respective masses m_1 and m_2 and the distance r which separates them are related by the formula,

$$\mathbf{F} = G\frac{m_1 m_2}{r^2}$$

However, I believe that Russell nonetheless overlooked four causal features of this formula and the gravitational law. For the gravitational law exhibits all four causal features discussed above. First, the law's system-type has conditions of componency which specify that it its instances are made of two objects with *mass*. As I have already suggested, I think we should understand having mass as a *dispositional* property, i.e. the disposition to interact with the Higgs field and impede acceleration under force. In the terminology of §5.2.1, the law exhibits *component-dispositionality*.

Second, instances of the gravitational laws' system-type have the ability to contribute behaviour in more complex systems. The gravitational law does not perfectly predict the net forces experienced by any two masses. This is partly because any gravitational component of force on a particular mass is generated by that mass's role in many more than a single two-mass system. *Every other distinct mass* will be a component in some two-mass system of which the first mass is the other component. But the law also does not perfectly predict because there may be other components of force present too. For instance, in Millikan's oil-drop experiment, the gravitational force on charged drops of oil which resulted from being components in a two-mass system with the earth was counterbalanced by the electrostatic force of a charged plate attracting the oil-drops in the opposite direction. The set-up of the experiment is such that the net force on the oil-drop can be zero, or even negative with respect to the gravitational field. Nonetheless, the embedded phenomenon intuition (see §3.3.1) tells us that instances of the system-type featuring in the Gravitational law are embedded (somehow) within more complex experimental systems such that they contribute to the overall behaviour of those more complex systems. In §6.1.3 I discuss various ways this embedding might be analysed. But for now it is enough to realise that the very notion of an embeddable system is a dispositional—hence *causal*—feature of the law. This is what in §5.2.2 I called a kind of *law-level dispositionality*.

Third, the formula $\mathbf{F} = G\frac{m_1 m_2}{r^2}$ exhibits what I referred to in §5.2.3 as *variable-level asymmetry*, i.e. a causal asymmetry among the variables. As in all special force laws, it is natural to think of the variables other than force playing the role of *causes* of force. After all, we wouldn't explain or manipulate any of the other values by reference to or an intervention on force, but we do explain and manipulate force by reference to and intervention on the other variables. Despite being a numerically manipulable equation, there is a causal asymmetry present among the variables. In §6.2 I will discuss how we might go about analysing such asymmetry in terms of *structural interpretations* of these equations which are not manipulable.

Fourth, the law itself exhibits what I referred to in §5.2.4 as a *law-level asymmetry* between system-type (divisible into conditions of componency and conditions of organisation) and behaviour. It is natural to think of the presence of two-masses oriented at distance r explaining their satisfaction of the formula $F = G\frac{m_1 m_2}{r^2}$ and not vice versa. This explanation is, I have argued, plausibly viewed as causal.

One might wonder why there are specifically *four* causal features of laws. Although I'm not in principle closed to the discovery of further features, I think something can be said about the relationship between the four mentioned which gives their individuation a certain legitimacy. The four features fall into the quadrants of Table 5.1 (p.184).

The horizontal axis of Table 5.1 divides causal features into levels of focus. While the higher level is concerned with causation concerning the law-instantiating system considered as whole, the lower level concerns causation at the level of parts of the system, specifically either in terms of the parts' variable-properties or the parts' dispositional properties.

Table 5.1: Two dimensions of causal conception

	High-level causation	Low-level causation
Potential causation	Law-level dispositionality	Component-level dispositionality
Actual causation	Law-level causal asymmetry	Variable-level causal asymmetry

The vertical axis of Table 5.1 divides causal features according to a different criterion. It has been made clear already that the features divide into pairs of dispositionality and causal asymmetry. But what could be made clearer is that the distinction charts the complemantary aspects of causal *potential* and causal *actuality* in lawful systems: dispositions characterise an object's potential for causal relations and causal asymmetries characterise the structures of actual causal relationships.

I take it that the tabulation of features in the above way lends some support to the existence, distinctness and mutually complementary characteristics of each of the four features to which I have been drawing attention. And in turn, these four causal features lend support to a causal conception of laws (although not a causation-mirroring conception!). I do not suppose that every law exhibits every kind of causal feature (except, perhaps, law-level causal asymmetry). However, these features are pervasive enough that I think we should see whether they can be put to use in helping us provide a better account of laws in general. The first step towards doing so is to isolate that class of laws which exhibit these causal features in a particularly important way. These are the *'robust causal junction laws'* to be dealt with in §5.3.

In closing this section, I suppose that even by this stage it is likely that some will remain unconvinced that the causal intuitions I have drawn on in motivating the four causal features of laws are really reflective of anything objective. The Russellian intuition is a strong one, particularly in philosophy of physics, and it is unlikely to be completely overturned by drawing on common turns of phrase found in textbooks or drawing on intuitions about explanation and intervention motivated outside rigorous experiment. Nonetheless, I think I've done enough in the foregoing sections to motivate the fact that we *do have* those causal intuitions about laws. Given that, it seems like a legitimate project to attempt to see if an analysis of the putative relations is forthcoming. This will be the project of §6. However, I also take it as implicit in what follows that the development of a coherent account of laws which not only incorporates the causal intuitions motivated above but also analyses them in ways typically associated with causal analysis (e.g. in terms of probabilistic dependency) would lend even more support to the causal features motivated above. I would ask, then, that those who remain unconvinced by the causal conception as it has so far been developed in this chapter to hold off their judgement until they have taken into consideration the full account developed in the rest of this chapter and in §6.

5.3 Robust causal junction laws (RCJLs)

Here I want to introduce the idea of a particular kind of system which exhibits all the causal features described in §5.2.5 in a particularly important and prevalent way. I call such systems 'robust causal junctions' and laws which concern them 'robust causal junction laws' (RCJLs). The aim of this section is to explain exactly what a robust causal junction is and why RCJLs are so useful to us. While RCJLs form a significant proportion of the laws, they do not account for all laws. Nonetheless, it is my understanding that RCJLs form a central class of laws from which the rest can be understood as playing a derivative or supporting role. Showing how is the project of §5.4.5.

The concept of *robustness* in scientific practice often indicates that perturbations of a certain type do not influence the overall behaviour of a system. In the case of component-level dispositionality, this robustness comes from the fact that changes in the material or construction of a component in a system need not affect the overall type of system if its disposition remains constant. As I pointed out in §5.2.1, this is a very useful feature of system-types in general. For instance, electrical conductivity can be achieved by a number of different materials, hence the systems which electrical laws concern are robust to the extent that the conductivity of their components is not affected by changes in their material. The fact that systems can be robust in this way means that we can look elsewhere for more suitable materials or designs to play the same role in a system.

A different kind of robustness in a system comes from the fact that changes in the complexity of an environment does not affect the qualitative causal relations it exhibits. As mentioned in §5.2.2, the phenomenon of system-types with the law-level disposition of embeddability enables us to idealise away from the real-world complexities to more easily perform rough predictions. An extra benefit arises when it is not only the numerical relations which are preserved (as a special case) in embedded contexts, but when the variable-level causal asymmetries are preserved as well. As has been noted before, though in different terms, such idealisation can help us draw clearer causal analogies across phenomena of varying degrees and types of complexity (Hesse 1963), allow us to set up more stable interventions and controls (Elgin 2004) and allow us to better conceptualise the complex theoretical body of causal knowledge about environments in which embedded systems appear Woody (2015).

Systems which are robust in these two ways are clearly very useful to us, and it is plausible that many of the laws we care about will be those which concern them. RCJLs are just of this sort. But of course, neither form of robustness is particularly desirable unless the behaviour associated with the system is of interest. This is where the idea of a causal junction comes in.

A causal junction is any set of multiple causal relations which all share the same effect—what I will call the 'focus' of the causal junction. Many of the laws we have considered so far which exhibit a variable-level causal asymmetry are ones in which there is clearly a single variable which is the effect of all other

Table 5.2: Laws' focal-variables and their variable-level causes

Law	**Focal-variable**	**Variable-level causes**
Ohm's law	I	V, R
Total force law	$\mathbf{a}$	$\mathbf{F}, m$
Lorentz-force law	$\mathbf{F}$	$q, \mathbf{E}, \mathbf{v}, \mathbf{B}$
Law of mutually gravitating bodies	$\mathbf{F}$	m_1, m_2, r
Law of pendulum motion	T	l, g
Snell's law	θ_2	θ_1, v_2, v_1
Compton shift law	λ'	λ, m, c, θ
Archard's wear law	Q	K, W, L, H
First law of thermodynamics	ΔU	Q, W

variables. This variable is the focus of causal junctions at the variable-level, what I will call the 'focal-variable' (we shortly consider causal-junctions at the law-level). So, for example, in Ohm's law the focal-variable is current, in the gravitational law it is force. In general, it seems to me that laws which exhibit such variable-level asymmetries are very likely to be structured in this way. Table 5.2 (p.187) provides a representative sample.

The tabulated causal junctions all exist at the variable-level. But we can also draw attention to another kind causal junction present in most laws, which exists at the law-level. Recall that in §5.2.4 I pointed out that if we break down the criteria of identity for a system-type into its conditions of componency and organisation, we can observe a causal relationship both between the satisfaction of the conditions of componency and behaviour *and* the conditions of organisation and behaviour. Hence, we have a causal junction at the law-level in which the focus is the overall behaviour referenced in the law. Arguably, such 'meta-junctions' are even more ubiquitous than variable-level causal junctions and are equally important to the definition of a robust causal junction, not least because they help us better analyse law-level asymmetry (see §6.2.5).

One might reasonably enquire whether there are causal relations *between* the variable-level and law-level causal junctions. The connection is certainly not straightforward. Since the foci are different in each case, we cannot simply add to one junction the causes of the other. Moreover, it would seem wrong to restrict the causes in the law-level junciton to a causal influence only over the focus at the variable-level junction. In many cases, law-level causes (i.e. conditions of componency and organisation) have causal influence over the variable-level causes as well. For example, in the case of the causal junction in Ohm's law, the value of voltage potential V (a cause in the variable-level junction) owes its existence to the resistor and voltage potential being appropriately organised (a cause in the law-level junction).

In general, I think we can draw at least the following point concerning the causal relations between law-level causal junctions and their corresponding variable-level junctions: that the causes in the junctions at either variable-level or law-level are always causal influences, though perhaps only indirectly, of the focal-variable (i.e. the focus of the variable-level causal junction). Even if the law-level causal junction has the satisfaction of the whole variable-level model as its immediate focus, the causes in that junction will have the focal-variable as a 'derivative' effect. For example, in the specific case of Ohm's law this amounts to the claim that the causes (derivative or otherwise) of current I are the cause-variables from the variable-level causal junction (i.e. R and V) and also the cause-variables from the law-level junction (i.e. the conditions of componency and organisation for Ohm's law's system-type; see Table §3.1).

I believe that many laws can be naturally understood as concerning both variable-level and law-level causal junctions. Observing this trend one might wonder why so many laws are like that. Continuing with the theme of explaining features of laws by exposing their utility, consider a simple game in which players roll a die d R_d times and calculate scores by summing the values of each roll, where V_d^i represents the number of times value i comes up on d. The numerical relationships in the scenario can be represented in the following equation.

$$R_d = V_d^1 + V_d^2 + V_d^3 + V_d^4 + V_d^5 + V_d^6$$

This equation tells us that the number of rolls is equal to the sum of the total number of times each value of a die comes up. If the number of rolls increases, so will $\sum_{i=1}^{6} V_d^i$. Arguably, the equation exhibits variable-level causal asymmetries. After all, we know that it is the roll of a die which causally explains the number which comes up and not vice versa (see Jeffrey 1971, for some indirect support of this). However, the asymmetries do not have the form of a causal junction, since there are multiple effect-variables $(V_d^1, V_d^2, ...)$ and only one cause-variable (R_d). For that reason, it is of limited value for certain practices.

To see why, notice that we can't ascertain for any instance of an increase in rolls with any certainty which specific effect-variables will increase. Now, when we construct gambling games we often want a repeatable element of randomness beyond the powers of the players to straightforwardly predict. This is what the roll of a die provides us with. But for that very same reason, a die would make a relatively poor device with which to perform other goals, such as *controlling* the rate of increase in score as one increases the rolls, or finding a way to achieve specific scores given that one knows how many rolls will take place. In general, the roll of a die is a poor tool to use in the service of in performing the effective strategies we often expect causal knowledge to provide us with.

One way to improve our predictive abilities for an indeterministic process is to *reduce the number of effect-variables.* Obviously, the most desirable structure for these purposes is the limit case in which there is only one effect-variable. If only rolls which resulted in a 'four' were counted then one could very easily predict the score one would achieve with a certain number or countable rolls, and one could very easily control the score by controlling the number of rolls. This, in effect, is an expression of the utility of causal scenarios in which there is a causal focus.

But there is also another way to improve our chances: one could *increase the number of cause-variables.* If the game described above instead allowed players to choose the number of dice D_d they threw each roll, where the score for each roll is the sum of all the values on each die rolled in that throw, then they would be able to exercise more predictable control over their scores. If someone's current score is 93 and their main priority is to end up as close to 100 as possible, they can work out that their best bet is to have one more roll with two die, or two more rolls with one die. Furthermore, if some technicality limits that player to a certain number of dice D_d they can compensate by choosing an appropriate number of rolls R_d and vice versa. This, in effect is an expression of the utility of the presence of multiple cause-variables for any effect. More cause variables often means more options for control, prediction and intervention (at least, where the variables can be intervened on).

In effect, what this reasoning shows us is the utility of causal junction structures in which there is only one effect—a focus—and multiple causal variables. Moreover, while we have reasoned this through considering variable-level asymmetries, it also carries over to junctions at the law-level. In the dice-game case, we know that substituting a 6-sided die for a 4-sided die (a change in the components) will affect the overall relationship between number of rolls and potential increases in value. We also know that sticking all the dice together so that the values on each die are no longer independent (a change in organisation) will similarly have an effect on overall behaviour.

We might also imagine that we could compensate for a change in the number of sides on a die by sticking it to another die to control for certain outcomes.

In combination with the two kinds of robustness described earlier the presence of two levels of causal junction in systems licenses a particularly useful kind of generalisation about them. These generalisations, I propose, are the *robust causal junction laws* (RCJLs), and the systems they describe *robust causal junctions*. More specifically:

Robust Causal Junction Laws

A generalisation of the form $\forall x(Sx \rightarrow Bx)$ is a *robust causal junction law* if and only if:

- S is a system-type with dispositional conditions of componency,
- B is modified (see §3.3.3)
- There are at least three variables in B exactly one of which is the effect of all the others (i.e. a focal-variable),
- B is the effect of the conditions of componency and conditions of organisation for S (i.e. a focus in the law-level causal junction).

RCJLs will play a central role in defining the soon-to-be-defined causal-junctions conception of laws (see §5.4.5), the Humean account of laws to be provided in §6 *and* the analysis of causation to be provided in §7. They are, therefore, perhaps the key concept of this entire thesis. But despite their pervasiveness, RCJLs do not account for every law, hence it is incumbent to say how they are related with other laws.

5.4 Non-robust causal junction laws (Non-RCJLs)

Not all laws are RCJLs. Some are almost RCJLs, others are not even close. This section works through a few examples which show how a law might fail to be an RCJL. My aim is to show, in each case, why these 'non-RCJLs' are nonetheless closely related to causal junctions and RCJLs. I will take this as an indication that the notion of lawhood in general is dependent on the existence of and close connection with the characteristics which define RCJLs.

5.4.1 Example 1: the ideal gas law

The behavioural equation of the ideal gas law in its standard arrangement is,

$$PV = nRT.$$

Plausibly (and I take this to be supported by the standard arrangement) the variable-level asymmetries are such that temperature T and quantity of substance n are cause-variables and pressure P and volume V are effect-variables. After all, if we intervene on either T or n (say by increasing them), we will expect the function $P \times V$ to increase. However, if the system is materially and mechanically isolated, an intervention on either pressure or volume would not result in change to quantity of substance or temperature.[7]

[7] Sometimes in practice when we intervene on volume, we can affect temperature too—as in diesel engines when the gases contained in the piston-chambers ignite as a result of an increase in temperature as the piston compresses the gas. (Thanks to Inge de Bal for suggesting this example to me and pressing me on the issues discussed in this section.) However, we must be careful not to intuit a causal symmetry between volume and temperature. Strictly speaking, the temperature of a gas increases in a diesel engine because work is supplied to the system by imparting momentum to the gas-particles and not because the volume the particles have to move around in diminishes. It is, thereby, possible (though not easy) to intervene on volume without supplying work to the system such that temperature does not change, and it is possible to intervene on work without changing the volume such that temperature does increase. Changes in volume and work of a diesel engine in such a process are effects of a common cause, viz. compression of the piston. Temperature increases are a result of the change in work only.

Table 5.3: Contexts of application for the ideal gas law

<table>
<tr><th>Law</th><th>System-type</th><th>Further system specification</th><th>Focal-variable</th></tr>
<tr><td rowspan="3">Ideal gas law</td><td rowspan="3">comprises an ideal gas in a cavity</td><td>...and is isobaric</td><td>V</td></tr>
<tr><td>...and is isometric</td><td rowspan="2">P</td></tr>
<tr><td>...and is isothermal</td></tr>
</table>

Granting that the variable-level asymmetries associated with the equation should be interpreted this way, it is clear that they do not capture that of a causal junction, since there are two effect-variables and hence no focal-variable. If we knew nothing more about the systems which obey the equation than that they are ideal gases in a closed cavity, we would not be able to tell, for some adjustment in temperature, say, whether or not pressure, volume or both would increase.

I think we should concede that in this under-specified state, the law does not have an unambiguous interpretation in terms of a causal junction.[8] But we can still claim that it has interpretations as an RCJL *in specific contexts of application*. For instance, when we turn up the heat under a sealed pressure cooker, we know that the pressure of the gas inside will increase. When we we put a blown up balloon in the fridge, we know that its volume will decrease. We know these things because we know something extra about the respective systems than is required by what it takes to be an instance of the system-type in the ideal gas law. Specifically, we know whether or not the systems are (nearly) isometric or isobaric. Table 5.3 (p.193) represents the process of disambiguation.

So, while the variable-level asymmetries of the ideal gas law does not have an

[8] I suppose we could posit some new variable 'X' where $X = P \times V$ as the focal variable. This might render the variable-level asymmetries of the modified equation that of a causal junction. Although I take it that replacing $P \times V$ with X in the ideal gas law would (at best) change the law itself. We are interested here specifically in the ideal gas law.

unambiguous causal junction associated with it, there are further contextual specifications commonly applied in practice which are clearly associated with a single kind of causal junction.[9] In general, we might reasonably expect other laws with ambiguous focal-variables to be associated with further contextual disambiguating procedures which render them able to be treated as RCJLs in those contexts.

5.4.2 Example 2: Gauss's law

The second law whose variable-level asymmetries fail to have a straightforward association with a causal junction is Gauss's law. The behavioural equation relates the displacement electrical field **D** normal to a closed surface S to the charge ρ contained within the volume V enclosed by the surface:

$$\oiint_S \mathbf{D}.dA = \iiint_V \rho.dv$$

Similarly to gravitational fields, which can be understood as the result of the presence of mass, the displacement field in Gauss's law is typically considered the result of the presence of free and bound charges. It can therefore seem natural to consider the variable-level causal asymmetries associated with the equation in Gauss's law to be such that the charge p contained within the volume is a cause and the displacement **D** normal to the closed surface its effect.

The choice of Gaussian surface also has an influence on the values we obtain for contained charge and displacement. This might lead us to posit the surface as another cause of displacement thereby allowing for an

[9]From a historical perspective, this is just what we would expect, given that the ideal gas law was, in fact, discovered via a precisification of Avagadro's law and three further previously known laws concerning the dependence of volume on temperature at constant pressure (Charles' law), of pressure on temperature at constant volume (Guy-Lussac law), and of pressure on volume at constant temperature (Boyle's law). All of these latter laws are RCJLs.

interpretation of the law as an RCJL (with displacement as the focal-variable). However, I think it would reasonably be objected that S is not a genuine cause. For the displacement field itself is not causally influenced by our choice of surface, only the value of field displacement across the surface. Moreover, since Gaussian surfaces are generally acknowledged to be imaginary mathematical devices 'which need not coincide with any real physical surface' (Serway and Jewett 2006, 626), a causal effect of the surface on anything can seem even less plausible.

Nonetheless, knowledge of the change in field across different Gaussian surfaces can help us make practical decisions. For instance, it can help us decide how best to develop the geometry of a capacitor. When we build a capacitor we will want to build conductors which can contain a charge-distribution across surfaces which can be brought close together to create a large uniform electric field. Knowledge of Gaussian surfaces in the context of Gauss's law tells us that flat plates will serve well since they have a high surface to volume ratio and can be uniformly separated by a small distance. The three variables of Gauss's law thus serve to provide *knowledge of* causal junctions where charge and conductor-geometry causally influence the electric field.

So, even if Gauss's law does not have a straightforward interpretation as an RCJL, the law seems to be conceived of in such a way that it can be used in contexts which enable it to be treated as one.

5.4.3 Example 3: the Planck-Einstein law

The third example of a non-RCJL I will consider is the Planck-Einstein law. This relates photons' energy E to their frequency ν according to the following equation (where h is the Planck constant).

$$E = h\nu$$

There are at least two reasons the variables in this equation cannot have asymmetries of the form of a causal junction. First, there are only two variables. So even if there were any causal influence at all between them, there is no third variable with which to form a causal junction. Second, it would be grossly misleading to say that physicists thought of frequency as a cause of a photon's energy or vice versa. The relationship, instead, seems to capture something like a *conversion* from one set of units to another. Frequency, we might say, is the way a photon's energy manifests.

Unlike the previous non-RCJLs, I don't believe there is a legitimate way to conceive of the variables in this law as causally related. However, I think we can show how the Planck-Einstein law is informatively associated with a law which does exhibit variable-level causal asymmetry and more specifically is an RCJL: *Planck's law*.

Planck's law, relates the system-type *black-body* with a behavioural property defined by the following equation (where $B_\nu(\nu, T)$ is the spectral radiance of a body in the frequency ν at temperature T; c is the constant speed of light in a vacuum and k_B the Boltzmann constant).

$$B_\nu(\nu, T) = \frac{2h\nu^3}{c^2(e^{h\nu/k_BT} - 1)}$$

In contrast with the Planck-Einstein relation, the Planck equation more plausibly has variable-level asymmetries in the form of a causal junction, since the temperature and frequency of radiation may be reasonably considered cause-variables and spectral radiance a focal-variable. This interpretation is supported by the fact that spectral radiance is an emission of energy which has been absorbed as a result of the heat transfer process which confers the black-body with heat, and further that the quantity of radiance is explained (in quantum physics) by the distance between energy levels at different frequencies.

But Planck was only able to explain the correctness of the Planck law under the assumption that the numerical relationship between the frequency and potential energy-changes of his 'atomic oscillators' in black-bodies' cavity walls was $E = h\nu$. Therefore, despite itself failing to exhibit the features characteristics of an RCJL, the Planck-Einstein law bears an important theoretical role supporting the truth of a law which does.

5.4.4 Example 4: Single-value laws

The final non-RCJL I will consider is rather different from any law so far considered. All the previously considered laws are *multi-variable laws* (MVLs): their behaviour-predicates are characterised by a formulae in which there appears more than one variable. However, I think it is plausible that we should treat a whole further range of conditionals as laws which do not have this feature. Table 5.4 (p.198) displays some examples.

The tabulated cases may not always be explicitly referenced as being laws, but they certainly seem sufficiently *lawlike* for them to be titled so. After all, they concern the behaviour of certain kinds of systems whose criteria of identity can be divided into conditions of componency and (admittedly trivial) conditions of organisation. These conditionals are also plausibly taken to support certain counterfactuals, can be incorporated into causal explanations and so on.

However, these laws clearly do not have a natural conception in terms of causal junctions. First, the behaviour-predicates only incorporate reference to one property. Since causal junctions require at least *three* property-terms (one for the focus and at least two for causes), there is no way that laws can exhibit the necessary variable-level asymmetries to form a causal junction. Second, the system-types' conditions of assembly are too simple. In most cases of MVLs the system-types' conditions of assembly comprise both conditions of componency *and* conditions of organisation. The laws just tabulated, however, only require of systems that they be an entity of a certain monadic property

Table 5.4: Table of single value laws

Law	S	conditions of assembly		B
		Comp'ts	Organis'n	
electron-charge	electron	x: electron	none	-1eV(x)
proton-charge	proton	x: proton	none	1eV(x)
metal-conductance	metal	x: metal	none	thermally-conductive(x)
copper-ductility	copper	x: copper	none	ductile(x)
Diamond-transparency	Diamond	x: diamond	none	Transparent(x)
Water's heat-capacity	water	x: water	none	4.186 J/gmK(x) (at 15°C, 101.325kPa)
Electrical field	electrical field	x: electrical field	none	x exerts force on charged particles.
Gene-heritability	gene	x: gene	none	heritable(x)
Lion-nutrition	lion	x: lion	none	Carnivorous(x)
Gold value	gold	x: gold	none	x is good money.

with trivial conditions of organisation.

For these reasons, I call the above laws ‘single-value laws’ (SVLs) and distinguish them from MVLs. SVLs are a special case of counterexample to the claim that laws can be interpreted in terms of causal junctions which are worth focused consideration. For as well as failing to count as laws concerning causal junctions, SVLs also fail to exhibit component-level dispositionality. Consequently, they are not robust in the way that RCJLs are. (This is also plausibly true of some MVLs, such as the Planck-Einstein law.)

However, whilst there are no dispositional criteria in SVLs’ system-predicates, there are dispositional requirements in their behaviour-predicates. This feature makes them pertinent to the application of many MVLs. For we can see how knowledge of an SVL can help us determine whether a particular object is a suitable component for an instance of some multi-variable system-type. An SVL tells us, for instance, that,

All copper is electrically conductive,

where *being electrically conductive* is a disposition to conduct electricity. Yet another SVL tells us that,

All polystyrene is electrically insulating.

where *being electrically insulating* is a disposition to insulate (rather than conduct) electricity. Clearly such SVLs are invaluable if, for instance, we want to build an electrical circuit which obeys Kirchhoff’s or Ohm’s laws, since to be an instance of those laws’ system-types, the respective system must be built from components which are electrically conductive. So, although SVLs do not themselves exhibit the variable-level or law-level causal asymmetries of a causal junction, they nonetheless often play a vital role in helping us make use of laws which do.

5.4.5 The causal-junctions conception of laws

We've now looked at four kinds of law which are not RCJLs. Certainly this is enough evidence to conclude that not all laws are RCJLs; there exist *non-RCJLs*. However, in each case, I have tried to show how the non-RCJLs are inferentially connected to RCJLs in some crucial manner. Some laws can be treated as *derivative* of RCJLs which have further contextually variable conditions of assembly concerning the context in which they are applied; some can be treated as *derivative* of RCJLs used in construction or design of certain systems; some can be treated as *supportive* of RCJLs by allowing for certain substitutions of variables; some can be treated as *supportive* of RCJLs for the sake of construction and design of their system-types. In general, non-RCJLs appear to be inferentially connected with RCJLs in what I will call from hereon a 'supportive or derivative way'. Although I understand that the above examples of various supportive or derivative ways may not be exhaustive, I will take them as evidence that *all* laws are either RCJLs or connected in some supportive or derivative way to laws which are RCJLs.

This insight forms the the basis of the causal-junctions conception of laws which can be summarised with a single slogan: *all laws concern robust causal junctions*. But more precision is provided by the following definition.

> **The causal-junctions conception of laws**
>
> A generalisation G is a law if and only if it has the form $\forall x(Sx \rightarrow Bx)$, and either:
>
> - G is an RCJL (see p.191), or,
> - there is some other generalisation G' which is an RCJL, *and* G is inferentially connected to G' in some supportive or derivative way.

Despite the open-endedness and vagueness of the notion of a 'supportive or derivative way' I think the conception is suitably defined to provide us with a fairly deep understanding of the relationship between laws and causal relations. The conception reveals that RCJLs are key to our general conception of lawhood, and since causation is central to the classification of RCJLs, this puts causation right back into our conception of laws.

5.5 Proposal for a causal-junctions *account* of laws

Recall that to say there is a causal conception of laws is to imply that there are certain truth-preserving inferences between laws and causal facts. We saw how this worked in the case of the causation-mirroring conception of laws. Under that view, laws mirrored the structure of causal relations such that for any causal relation, there would be a law which 'covered' it, and for any law, there would be causal occurrences which could count as an instance of it. Crucially, granting the conception does not imply any direction of metaphysical priority. While some who upheld the causation-mirroring conception would follow the traditional Humean approach of grounding causation in facts about laws (e.g. Paul 2000, Schaffer 2001, Baumgartner 2008), others (e.g. Armstrong 2004a) take there to be no direction of priority and perhaps others still, a priority in the opposite direction.

Having said that—and as we saw with the causation-mirroring conception—if we take causation to be a primitive in our analysis, then the causal-junctions conception can provide us with a fairly substantial non-Humean *account* of laws (due to the open-endedness of the conception I stop short of calling such an explication of laws an analysis). The account can be expressed as follows.

The causal-junctions account of laws

What *constitutes* a generalisation having the status of a law is its satisfaction of the criteria described by the causal-junctions conception of laws.

By saying that the conception *constitutes* lawhood, the account makes the inferential connections defined by the conception part of what underlies the very character of laws, explaining what makes them special.

But observe that in fact, there is nothing in the causal-junctions account as it is stated above which a Humean could not *in principle* agree with. Under a Humean perspective, the causal-junctions account indicates how certain type-level causal relations can be employed to account for laws thereby implying that previous Humean attempts to analyse all causation only once the laws have been established (as those who uphold BSA do) was wrong. Obviously it is then incumbent on the Humean to show how causal features which contribute to the definition of RCJLs can be analysed in terms of the mosaic.

This Humean strategy is precisely what I will be proposing in §6. But before moving on to that specifically Humean goal, it will be informative to consider briefly the merits of the causal junctions account of laws regardless of whether it is to be adopted by the primitivist about causation or the Humean.

As a first remark, the causal-junctions account of laws provides an indication of how we might account for those aspects of law mentioned in §4.2 from Lewis. For example, the connection with robust causal junctions shows that laws cannot be individuated by logic alone and it seems to be consistent with the idea that laws are contingent and mind-independent. The account also gives us a suggestion as to why many theorems of scientifically confirmed theories are reasonably held to be laws. For if those theorems concern similar causal relationships, or else are inferentially related to such relationships in appropriately derivative or supportive ways, then those theorems will also be laws according to the account. Notice, however, that without any further criteria, the account is not committed to *every* theorem of a confirmed law coming out as a law. This, I take it, is a good thing; as I pointed out in §2.2 with regard to NP, the commitment to an unrestricted deductive closure seems to render accounts of laws subject to counterexample.

Although it beyond the scope of the thesis to explore the idea fully, we might also be hopeful that the connection made in the account between laws and causation will help us to see more clearly the way in which counterfactuals come to be supported by laws (a feature noted by many, e.g. Goodman 1947, Hempel 1965, Swoyer 1982, Fales 1990, Loewer 1996). Moreover, it should be considered a significant boon of the account that it is so clearly inferentially connected with the causal knowledge. Throughout §5.2 and in §5.3 I have made an effort to show how useful all the causal features of RCJLs are and so, derivatively, how useful the non-RCJLs inferentially connected to them. If we think it is a feature of the laws we posit that they bear a significant degree of practical and epistemic benefit (as seems especially plausible from a Humean perspective Hall 2015) then any account which stays true to the causal-junctions conception would seem to offer a particularly comprehensive explanation of what this benefit is.

As a last point in support a causal-junctions account of laws, I want to return to a point briefly touched on in §3.3.3 concerning the legitimacy of some special case formulae over others. Consider, for instance, the equation which features in the behaviour-predicate of the Newtonian Consitutive law for compressible fluids:

$$\tau = -\mu\frac{\partial \mathbf{u}}{\partial y} + \Big(\frac{2}{3}\mu - \kappa\Big)(\nabla \cdot \mathbf{v})I$$

For this equation, τ is shear stress of the fluid, μ is the steady-shear viscosity, $\frac{\partial \mathbf{u}}{\partial y}$ is the velocity variation across the y direction (transverse to direction of flow), κ is bulk viscosity and $\mathbf{v}$ the bulk flow velocity (I is the identity tensor). There are many variables we could fix at constants to get a special case equation. One notable example is the equation we get by putting $\nabla \cdot v = 0$. For this gives us the equation which describes the behaviour in Newton's constitutive law for *incompressible* fluids. However, some fixings of variables are certainly off the table. Perhaps most clearly, we would not get a legitimate special case by taking τ as constant.

In §3.3.3 I had no explanation for this distinction, but now it seems like there might be one available: *that special cases of a formula are only legitimate for reference in laws' behaviour-predicates if they do not treat the focal-variable of the formula as a single value.* Since τ is the focal-variable of the constitutive law for compressible flow, we get a legitimate special case when fixing other variables (such as $\nabla \cdot \mathbf{v}$), but not when fixing τ itself.

It's not hard to see why this would be a general, if implicit, rule in our practice of formulating laws. Laws, and RCJLs in particular, are constructed around causal junctions which allow us to control or predict a particular variable of interest: the focal-variable. Since interventions or changes in variables in such junctions typically lead to a change in the focal-variable, special cases in which the focal-variable is stable will be relatively rare. Indeed, it is this feature of causal junctions which forms the basis of the probabilistic analysis for them in §6.2. But already we can see why special cases of formulae in which the focal-variable is assumed to remain a single value are unlikely to be of much use to us.

Given the foregoing, an account of laws which adopts the causal-junctions conception as the central insight clearly has many merits. Were we to take causation as a primitive, we might end there. But my project aims to satisfy the Humean methodology. Consequently, there is further work required to show how the relevant causal features of RCJLs (two of dispositionality, two of causal asymmetry) can be given a Humean analysis.

6 A Humean analysis of robust causal junction laws

At the end of the previous chapter I proposed a 'causal-junctions account of laws'. In effect, this suggests that all laws can be understood derivatively of a special class of laws: the robust causal junction laws (RCJLs). These laws in turn are to be accounted for in terms of their causal features, viz. the component-level and law-level dispositionality, which makes them robust by stabalising their instances against certain kinds of interference, and the structure of variable-level and law-level causal asymmetries, which take the form of causal junctions each centering on a single respective effect or 'focus'. In proposing a causal-junctions account I have made free use of causal terms, thereby offering it as a complete account for the primitivist about causation or a partial account for the Humean. In this chapter I aim to make it respectable from the perspective of a Humean methodology by analysing the causal features of RCJLs.

Analysing the kinds of causal feature which go into characterising RCJLs in Humean terms has proven tricky in the past, and the task at hand might plausibly be thought even tougher given the causal-junctions account of laws. Typically Humean treatments of causation have assumed the laws are decidable prior to any account of causation (e.g. Reichenbach 1956, Lewis 1973a, Dowe 2000, Baumgartner 2008, Fenton-Glynn 2011). This way the laws can be used to then provide an account of causation. However, if type-level causal phenomena such as those concerning dispositionality and causal asymmetry are to be used to provide an account of laws in the way this current proposal has it, we certainly can't expect to analyse *those* causal phenomena with laws.

Nonetheless, I believe the task of giving a Humean analysis of those causal features which enter into the causal-junctions account of laws can be provided without circularity. In §6.1 I focus on the features of RCJLs which contribute to their robustness, i.e. component-level and law-level dispositonality. I then move on, in §6.2, to give analysis of the causal asymmetry in RCJLs by providing a probabilistic analysis of causal junctions. §6.3 concludes.

6.1 Analysis of RCJLs' dispositionality

Here I propose to analyse the dispositional features present in RCJLs. I begin, in §6.1.1 with a defence of an extensional conditional analysis of dispositions 'StoCA' inspired by that given by Storer (1951). I then discuss, in §6.1.2, the problem of *masks*, which I take to be the central obstacle to developing successful conditional analyses. As I understand it, the problem of masks is closely connected with issues of universality and in §6.1.3 I show how the solution of 'behaviour-modification' proposed in §3.3.3 can also be put to work in giving us an extensional analysis of both law-level and component-level dispositions in laws without the threat of masks. The outcome of this analysis is that we can confidently assert the dispositional features of RCJLs without concern that this threatens the Humean methodology.

6.1.1 In defence of an extensional analysis of dispositions

It might not immediately seem obvious why a Humean is threatened by dispositional properties. After all, the Humean methodology was defined in §2 as a position which requires no reference to primitive causal *connections* in the final analysis, rather than causal properties. Still, dispositions were introduced in terms of a *potential* to cause or be caused to do things. At best, this can sound like shifting causal talk away from the actual mosaic to some possible causal events, thereby invoking both subjunctive causal talk. I will assume from hereon that any honest Humean would aim to find an analysis of dispositions in terms of the mosaic.

The project of analysing dispositions in non-causal (specifically, *Humean*) terms has been long and, arguably, unresolved. A good place to start recounting its development is with the early and quickly dismissed suggestion that dispositions should be understood as having the form of an extensional conditional analysis (ExCA) of the following form.

$$\textbf{(ExCA)}\ \forall x(Dx \leftrightarrow \forall t(T(x,t) \rightarrow M(x,t)))$$

In ExCA, x ranges over objects and t times, 'D' is replaced by some dispositional property term, T by a categorical stimulus or 'trigger' condition and M by a categorical manifesting behaviour of the disposition. So, for instance, the disposition of fragility might be analysed,

$$\forall x(\text{Fragile}(x) \leftrightarrow \forall t(\text{Stress}(x,t) \rightarrow \text{Breaks}(x,t)))$$

Despite all its problems (to be discussed shortly), ExCA has a number of notable merits. First, it has the form of a genuine *analysis*. This is because its instances, if true, would provide conditions that *all and only* objects with the corresponding disposition replacing D have. Second, it is a *conditional* analysis drawing on stimulus and manifestation conditions. This is invariably the way in which philosophers have thought an analysis of dispositions should proceed and I will assume the same. Third, it is something the Humean can accept, since the stimulus and manifestation conditions (the properties T and M) are defined to be *categorical* (as opposed to dispositional) with no non-trivial causal, subjunctive or higher-order characteristics. Fourth, it is *extensional*; that is, the truth-values of instances of ExCA are determined by the distribution of the extension of their terms. This is to be contrasted with opaque truth-conditions (defined in terms of beliefs or desires) and *intensional* truth-conditions (which make reference to possibilia). While I suppose the latter could be entertained under a Humean approach which openly embraced modal realism (see Lewis 1986f) I will assume this would constitute a concession on the part of the Humean. An extensional analysis restricted to the actual is just the sort of thing a Humean would aim for were it considered workable.

So much for the merits of ExCA. One issue with it concerns the inclusion of an explicit time-variable. Presumably, in propositions which do quantify over times, the absence of a time variable from a predicate-clause is supposed to indicate the permanence of the property. But though D lacks a time variable in ExCA, I see no reason in general to suspect that dispositions are essentially always had by their bearers (see McKitrick 2005). We should assume, I think, only that the disposition needs to be instantiated at least for the time in which it is stimulated and manifests. But if stimulus and manifestation are at the same time (as indicated above) then why include an explicit time variable in the dispositional predication at all? Just as we did with Schema 2 for laws, it may seem better to simply leave time out of the explicit formulation.

But is it in fact true that all dispositions manifest simultaneously with their stimulation? Although this conflicts with some analyses (e.g. Malzkorn 2000), it seems to me that as long as we treat the relevant predications in the present-continuous, we will retain the simultaneity in all the cases we're interested in (viz. those dispositions which feature in laws). Indeed, it is plausible that this will work for dispositions in general. For example, a cube of sugar may dissolve after it is placed in water, but it will be dissolving as soon as, and throughout the time that, it is immersed in the water. A vase may be breaking still after the striking is over, just as temazepam makes one sleep after one has ingested it. But at least it seems correct to say that a vase is breaking as soon as it is being struck, and a person is growing sleepy as soon as the temazepam starts to be ingested. In general dispositions are manifesting as soon as, and throughout the time that, they are being stimulated.[1]

Insofar as ExCA is plausible at all, I will assume it is harmless to leave out the time-variables. This leaves us with the following.

$$\textbf{(ExCA*)}\ \forall x(Dx \leftrightarrow (Tx \rightarrow Mx))$$

A harder problem to solve both for ExCA and ExCA* was pointed out by Carnap (1936). If something is never struck or it breaks for any reason whatsoever, then the above analysis deems the object fragile. But this is surely too inclusive. Plenty of fragile things never break and plenty of non-fragile things break, e.g. under extreme stress.

In response, Carnap suggested that the best we can hope for are 'reduction sentences' for dispositions of the following schematic form.

$$\textbf{(CRS)}\ \forall x \forall t(T(x,t) \rightarrow (D(x,t) \leftrightarrow M(x,t))).$$

The corresponding reduction sentence for fragility does not appear to make the unwanted ascription of fragility to non-fragile objects (as ExCA* does). But this comes at the cost of only *partially* defining the disposition, since a reduction sentence does not give a necessary and sufficient condition for having a disposition but only the condition for having the disposition D *if it is subject to stimulus* T. Hence, reduction sentences will not help us *analyse* dispositional ascriptions by substituting them with any expression in terms of the mosaic's categorical properties.

[1] Ultimately, the move to get rid of time-indexes is one of simplification. If it turns out that they are after all required to define some dispositions (e.g. in relativistic contexts) I will assume they could be put back in.

In response to this failure of extensional analysis, a subsequent approach was to introduce subjunctive conditionals into the analysis. What has become known as the 'standard conditional analysis' (SCA) of dispositions has the following schema.

$$\textbf{(SCA)} \quad \forall x(Dx \leftrightarrow (Tx \succ Mx))$$

'$\succ$' symbolises the subjunctive conditional so that the right hand side of SCA is to be read 'had it been the case that Tx then it would have been the case that Mx'. Under this schema, fragility gets defined informally,

Anything is fragile if it would break were it struck.

As Choi and Fara (2016) point out 'SCA has been explicitly endorsed by Ryle (1949), Goodman (1954), and Quine (1960) and implicitly by countless others'.

Despite its plausibility, there are many reasons not to accept SCA. One issue with it is the plethora of counterexamples to many of its putative instances. Consider the above analysis of fragility in its terms. A moment's thought shows that it is false for at least the following four reasons. First, if a something is wrapped in enough bubble-wrap it may not break despite being struck even if it were fragile. In such cases we say the disposition has been 'masked' (or there exists an 'antidote', Bird 1998). Second, if something were protected by a force-field whenever anything hard comes near it above a certain momentum, it may not break if it were struck even if it were otherwise fragile; in such cases we say the disposition is subject to 'reverse-cycle finks' (Choi and Fara 2016). Third, if something is liable to be subjected to some further destructive mechanism while being struck, e.g. a dissolving agent, then it might be the case that it *would* break were it struck even if it were *not* fragile; in such cases we say the disposition has been 'mimicked' (Johnston 1992). Fourth, if something were subject to a significant restructuring mechanism whenever struck it might break were it struck even if it were not otherwise fragile; in such cases we say that the disposition is subject to a 'fink' (Martin 1994).

There have been a number of responses to one or other of these counterexamples, some more plausible than others (Bird 2007a, Choi 2008b, Lewis 1997, Malzkorn 2000, Manley and Wasserman 2011). But while I will come to consider some of these shortly, there is a more immediate concern with SCA worth mentioning for the particular project at hand. As we have seen, SCA proposes to analyse dispositions in terms of subjunctive conditionals and hence is no longer extensional. Now, there are Humeans who have accepted the employment of subjunctives in their analysis of dispositions (e.g. Lewis 1997), but the duty then must be to analyse the subjunctive claims themselves in Humean terms. While such analyses do exist, they will not be appropriate for our project, since almost every plausible reductive semantics for subjunctive conditionals (including those explicitly Humean) uses *laws* to analyse them (see, e.g. Stalnaker 1968, Lewis 1973b, Jackson 1977, Maudlin 2007). Remember that the purpose of providing an analysis of dispositions is so that the causal-junctions account of laws can be rendered in a Humean light, thereby giving us a Humean treatment of the causal-junctions account of laws. The laws themselves cannot, therefore, be employed in an analysis of those very dispositions which help to analyse them.[2]

So, SCA is not easily adopted under the current project. However, it appears that there is, anyway, an extensional analysis which both improves on ExCA *and* is not subject to so many of the problems listed for SCA two paragraphs back. This is the analysis proposed by Storer (1951). Admitting that there

[2]Perhaps there are more nuanced options which might seem available to the Humean who wants to maintain both a subjunctive conditional analysis of dispositions and propose a causal-junctions account of laws. For instance, one could hold that there are 'fundamental' laws which have no dispositionality whatsoever and are sufficient to get the subjunctive facts from where we get a causal analysis and finally the remaining dispositional laws. However, I will not pursue this thought. I'm neither confident about the claim that there are fundamental laws, nor about the claim that they would be without dispositionality were there any.

may be objects with a disposition which never get to manifest it because they are never appropriately stimulated, Storer suggested that such objects will nonetheless share some (categorical) property with a class of objects to which the disposition can be attributed at least one of which is stimulated and none of which fail to manifest the disposition when stimulated. In Storer's original (but expanded) formulation, this looks as follows.

$$\begin{aligned}\forall x(Dx \leftrightarrow\ & (\exists t(T(x,t)\ \&\ M(x,t)) \vee \\ & \exists F((Fx\ \&\ \exists y(Fy\ \&\ \exists t'(T(y,t')\ \&\ M(y,t')))\ \& \\ & \neg\exists z(F(z)\ \&\ \exists t''(T(z,t'')\ \&\ M(z,t'')))\end{aligned}$$

It says that for anything x, it has a disposition D if and only if there is a time t at which x is subject to stimulus T and manifests property M or else, there is a property F which x and another object y has where y is stimulated by T and manifests M at some time t' and nothing z which has property F and is stimulated in way T at a time t'' is such that it manifests M at t''.

Storer's origninal formulation seems to me a bit long-winded. Following the reasoning concerning ExCA, I think we can remove most of the time-variables. Moreover, it can be reasonably reformulated in the following almost logically equivalent but even more abbreviated way.

$$\textbf{(StoCA)}\ \forall x(Dx \leftrightarrow \exists F(Fx\ \&\ \exists y(Fy\ \&\ Ty)\ \&\ \forall z(Fz \rightarrow (Tz \rightarrow Mz))))$$

I say 'almost logically equivalent', since StoCA makes use of my assumption that all dipositions manifest whilst being stimulated and unlike Storer's original precludes something having a disposition without also having a categorical property F to which it can be attributed. Insofar as we expect to explain dispositions via reference to a categorical property at all (something made more plausible given consideration of SVLs; see §5.4.4), this latter preclusion seems like a reasonable further assumption.

Whilst being extensional, StoCA quantifies over properties as well as particulars. I don't suppose there is anything specifically problematic about this. But if the analysis is to avoid triviality, it must be that the domain of properties is restricted to some degree—the domain obviously shouldn't include dispositional properties! But it must also be restricted further to avoid gerrymandered grue-like properties or properties defined by a list. Storer's rather brief suggestion is to restrict the properties to those denoted by "'natural' predicates; *i.e.* undefined predicates, or predicates whose definitions do not involve explicit reference to an individual' (1951, fn.1). Hence, the corresponding non-formalised expression of an analysis of fragility might look something like the following.

> For all things, it is fragile if and only if it has *some natural categorical property* which something which is struck has such that if anything has that property and is struck, it breaks.

It seems to me that there are many properties to which we might both plausibly refer to as natural in some sense and to which we attribute objects' fragility, e.g. being made of thin glass, being made of thin porcelain, etc.

One might understandably be concerned with the brief definition for naturalness given by Storer. But there is more to say. For the Humean, what counts as natural is likely to have some connection with the basic constituents of the mosaic. So, for Lewis, naturalness is a matter of degree, with a set of perfectly natural properties populating the mosaic and less natural entities being defined derivatively in terms of them; the more natural, the simpler the definition (Lewis 1983, 1984). Others have proposed more specific claims about what count as the fundamental properties and consequently what counts as the most natural. Earman (1986), for instance, counts space-time magnitudes as the central class of natural properties whereas Loewer (1996) considers them to be 'changes of the locations and motions of macroscopic spatio-temporal entities' as well as 'changes in properties that account for locations and motions and so on' (p.110).

A more radical Humean approach to naturalness has been advocated by Callender and Cohen (2009; see also Schrenk 2007). Partly in an effort to allow for the basic entities of special sciences to have a place among the natural, Cohen and Callender (2009) suggest we advocate a world comprising a plurality of mosaics layered over each other and individuated by the properties relevant to existing and potential scientific fields. According to such a picture the natural properties and kinds are relativised to the particular granularity of the mosaic in which they are found. The various mosaics have their own taxonomic structures of properties either basic to or compounded from properties which are basic to that specific mosaic. Zoology has a structure in which the basic properties are presented by genus-species relations among animals; particle physics has a structure in which the basic properties are presented by the standard model; microeconomics has a structure in which the basic properties are presented by the model of exchange of commodities between individuals and firms; etc. The naturalness of a property is, therefore, relative to the mosaic in which it appears.

Whatever our answer concerning what defines the natural in the mosaic, I take it that the analyses we provide for dispositions of the form StoCA should quantify over properties which are at least as natural as the dispositions' stimulus and manifestation properties. Otherwise we once again risk trivialising the analysis by allowing 'grue'-like properties to count as suitably natural. Of course, if it turns out that we must commit to a particular view of naturalness, then we may have to deny that certain dispositions can be analysed in the manner of StoCA. Here it's worth recalling that the project at hand is to analyse the dispositional properties *which feature in laws* rather than all dispositions whatsoever, and I take it that the dispositions which appear in nature's laws will be at least as natural as the laws are.

A different concern we might have is that in contrast with SCA, StoCA requires that for a disposition to be analysable, it has manifested sometime and someplace in the mosaic. This may seem more of a problem than it is. For example, Storer considers 'a class of things none of which would ever be put into water,' and asks 'would these things be soluble or not?' I think Storer's own response is just right.

> On our definition [i.e. StoCA], they would be soluble only if they belonged to some larger class such that all the members of this class that had been tested in water did dissolve. Thus we may say of some hypothetical chemical substance that *will never be tested* for solubility that it is soluble if, for example, it belongs to the class of salts of which all the members that are ever tested do, in fact, dissolve. (1951; see also Lewis's 1994 discussion of 'unobtanium'.)

Perhaps this response will not eradicate all concern. Couldn't it have been the case that *no* soluble thing ever dissolves? Or at least, couldn't it have been the case that no thing which dissolves shares a natural property with some class of objects all of whose instances dissolve when placed if water? To some degree I think we Humeans can simply be hard-nosed about this. At least, everything in the actual world which has a disposition shares some natural property with a class whose members all manifest in the appropriate way if they are stimulated in the appropriate way (whatever those ways are) and where at least some member in that class is appropriately stimulated. After all, it seems in keeping with the Humean-empiricist approach that we don't posit the existence of properties which have no unexplained manifestation in the mosaic at all.

Things get more nuanced when we move to modal contexts. For it is plausible that it *could have been* that no soluble thing ever dissolves in water, and if we accept this then we'll have to drop any expectation that dispositions are determined by the content of mosaic in which they appear. But this is anyway something which must be given up upon accepting NP (something I have already voiced support for; see §2.2). The repercussions of the various decisions we make with respect to how dispositions and laws behave under counterfactual suppositions are worth thinking about. But to discuss them here would take us beyond the problem at hand of getting the extensional analysis right.

In sum, I think there is a lot going for StoCA as a schema for dispositional analysis. By virtue of remaining extensional it is more in keeping with a Humean methodology which accepts that dispositionality is an explanatory feature of laws. But notice also that by remaining extensional, StoCA also appears to avoid many of the previously mentioned problems which beset SCA. The problem of finks and reverse-cycle finks stems from the fact that a stimulus condition can at the same time introduce or remove (respectively) the disposition which is otherwise present by default. In order to avoid such problems Lewis (1997) supplemented his subjunctive analysis with the qualification that some relevant intrinsic property be retained throughout the stimulus and manifestation process. But because StoCA is extensional, and implicitly attributes properties simultaneously, it already entails that the disposition is had at the time of that process. StoCA also does not suffer from the existence of mimickers. On the one hand, so long as there isn't a class of entities which fall under the same natural property *all of which* mimic some genuine disposition, a single case is no threat. On the other hand, if there is such a class, there would be good (Humean) support for the idea that its instances were not mimics at all, but genuine instances of the disposition.

StoCA therefore appears superior to SCA not just for the specific project at hand, but also for more general considerations. However, we cannot ignore entirely the counterexamples posed to SCA, since StoCA does suffer from the problem of *masks*. A moment's reflection should reveal why. The existence of masks means that it is simply not the case that for all instances of the disposition, if they are stimulated then they will manifest the right property. If a vase is wrapped in bubble-wrap then the vase will not break if struck, despite having the disposition *and* having a relevant natural property (e.g. being made of thin porcelain). The next section considers what to do about this.

6.1.2 The problem of masks

In contrast with some, I assume that the problem-cases of masks are indeed genuine and problematic. Choi (2008a), for instance, has suggested that we are interpreting masking cases wrongly: a Ming vase wrapped in bubble-wrap is *no longer* fragile and a plastic cup which melts when struck (because of some peculiar mechanism involved) *would be* fragile. I do not think this can be right. A Ming vase remains wrapped *because* it is fragile, not that it would be fragile were it not wrapped. More importantly, dispositional ascriptions are useful to us because we know what sorts of systems an object with the property can contribute to. We know that copper is conductive, so we put it in electrical circuits. We desire a way to recognise the utility a feature of an object can supply us with even when it is not supplying it; that is, the dispositional property 'carries over' from context to context (Cartwright 1999, Hüttemann 2004).

The problem of masks means that, in general, dispositions cannot be analysed in the form of StoCA. Here I take inspiration from the great many attempted solutions to the problem of masks on behalf of SCA and consider whether an extensional reinterpretation of any of them would be a suitable response on behalf of StoCA. The resounding answer will be 'no', leading me, in the next section, to propose an alternative solution.

One response to masks on behalf of SCA, which is considered by Bird (2007a, §3.3) allows that while masks (Bird calls these 'antidotes') are a threat to SCA for non-fundamental dispositions, they may not be for the most fundamental dispositions.

> Since we are dealing with the fundamental level, and have already removed the problem of multiple realizability, it might be reasonable to expect that any dispositions of this sort will suffer from relatively few antidotes. [...] [T]he direction of the development of physics with ever fewer fundamental properties and corresponding forces indicates that the prospects for antidote-free fundamental properties and thus strict laws only at the fundamental level are promising. (2007a, 63)

Taking inspiration, we might hope an analogous argument could be made in defence of StoCA. However, there are problems. First, Bird seems to be working under the widespread assumption that the physical laws (for Bird, ultimately grounded in fundamental powers) constrain what systems cannot exist (see also Fine 2002). But if we take seriously the claim that laws are conditionals and concern properties essentially local to the systems they are about (see §3.1 and §4.1) it can seem entirely unjustified to infer from the fact physics is positing fewer and fewer properties and forces that this implies that there *can be no* masks at that level. The fact that there don't happen to be any systems in existence whose lawlike behaviour involves masking the effects of all four forces by some unknown further kind of influence is nothing the laws of physics can tell us. A strict law may be strict because nothing in fact interferes with its instances, but in developing an analysis of a disposition we are typically after something modally stronger.[3]

[3]This is even the case in the context of our preference for an extensional analysis, since Humeans may still want to draw inferences based on the extensional analysis about how dispositional properties behave in other worlds. Under the assumption of NP, I suggest we are safe to infer the instances of the following schema.

$$\Box_{Phys.}\ \forall x(Dx \leftrightarrow \exists F(Fx \,\&\, \exists_{@}y(Fy \,\&\, Ty) \,\&\, \forall z(Fz \rightarrow (Tz \rightarrow Mz)))).$$

Instances of this schema remain true in physically possible worlds (the range of $\Box_{Phys.}$) where a '@' suffix restricts a quantifier to the domain of the actual.

Second, even if all physically possible objects were constrained by the strict fundamental physical laws which concern what actually happens, I simply disagree that there would be no masks at the fundamental level. Presumably mass and charge are pretty fundamental properties if any are, and yet they appear to be able to mask each other—making use of this fact seems to me precisely the ingenuity of Millikan's oil drop experiment. Gravitational mass is the disposition to accelerate under the presence of another gravitational mass, electrostatic charge is the disposition to accelerate under the presence of another electrostatic charge. If we set up the situation correctly an object with mass and charge can do neither of these things in the presence of other gravitational masses and electrostatic charges. The fact that we know how to combine strict (though potentially behaviour-modified) laws in order to predict such behaviour does not, I take it, contradict the fact that the case is one of masking.

Third, and regardless of whether the previous points are correct, the project I am engaged in here is to say something general about very many more dispositions than those which exist at the so-called fundamental level. For example, frictionlessness can be masked by drag forces, elasticity can be masked by distribution of loads, predatorial dispositions can be masked by environmental factors (e.g. disease, geographic distribution), the value of a currency can be masked by political and social factors, the fertility of soil can be masked by weather conditions, etc., etc. Even if Bird is right about the fundamental level, its worth pursuing an analysis which doesn't limit itself to getting things right only at that level.

A different approach to the problem of masks is to develop a more sophisticated analyses (e.g. Lewis 1997, Malzkorn 2000). Perhaps the most sophisticated of which is that provided by Manley and Wasserman (2007, 2008, 2011). They suggest the following conditional analysis.

$$\textbf{(MWCA)} \quad \forall x(Dx \leftrightarrow (Tx \succ (P(Mx) > Threshold_C)))$$

This is to be read 'for all x, x has the disposition D if and only if were it subject to stimulus T it would have a probability of being M, $P(Mx)$, greater than some threshold (on the number line between 0 and 1) defined by the context of utterance C'. Taking inspiriation, an extensional analogue of MWCA in the style of Storer's proposal would then be the following.

$$\textbf{(MWCA-Ex)}\ \forall x(Dx \leftrightarrow \exists F(Fx \mathbin{\&} \exists y(Fy \mathbin{\&} Ty) \mathbin{\&} \forall y(Fy \rightarrow (Ty \rightarrow P(My) > \mathit{Threshold}_C))))$$

Under this definition, it need not be the case that *every* instance of a disposition manifests when stimulated, only that there is a suitable probability (above the defined context-sensitive threshold) that each does.

By only requiring a non-trivial probability of manifesting M in response to T, MWCA-Ex (like MWCA) is consistent with the existence of recalcitrant cases as a result of masks.[4] Nonetheless, I do not think MWCA-Ex (or, indeed, MWCA) is a very helpful analysis.

First, as I pointed out in §3.1.4, laws are the sorts of thing which should not be subject to so much variation from context. Laws are written down in textbooks, taught in lectures, posited in international journals, etc. If laws are to be characterised partly in terms of dispositions, as the causal-junctions account of laws would have it, then those dispositions had best not be overly context sensitive. But the inclusion of a parameter of context seems necessary for Manley and Wasserman's analysis to work. Otherwise, there would have to be implicit in every dispositional ascription an exact probability for manifestation upon stimulus for each dispositional property (or at least, those which feature

[4] Manley and Wasserman's account is not only motivated with the intention of solving the problem of masks, but also a number of other problems which they raise for the traditional SCA of so-called 'achilles heels' and 'reverse achilles heels'. While I don't doubt such problems are logically possible, I think they are certainly beyond the scope of the sort of problems which face component dispositions, and so will not take into account the merits of MCWA and MCWA-Ex in avoiding them.

in laws), which can seem implausible without further motivation.

Second, MWCA-Ex and MWCA both seem to imply that masking and manifestation is an all-or-nothing affair. The probabilities in question generate an expected frequency of cases which are completely unmasked or completely masked. However, we saw in the discussion of embeddability §5.2.4 that some dispositions may be 'continuously manifesting'. The embeddability of a system is like this because it can manifest certain behaviour completely in the isolated case or partially under some interference. Similarly, it seems that electrostatic charge will be continuously manifesting depending on the extent to which non-electrostatic forces interfere with objects which have it. Indeed, given the heavy emphasis on quantification in scientific laws, it seems that it will often be the case that the kinds of disposition which feature in system-types' conditions of componency can be masked *to some degree* rather than completely or not at all. Therefore, we need an analysis of dispositions which makes sense of the continuous nature of masking and manifestation.

While I will not suppose that analyses of the form Manley and Wasserman present cannot be made to work, I want to now explore a different option which does not rely on context-sensitive probabilities of manifestation.

6.1.3 A systems analysis of component-level dispositions

In proposing a new solution to the problem of masks it is important to remember that the task at hand is not to analyse *all* dispositions, but only those which feature in laws, viz. the various component-level dispositions and the law-level disposition of embeddability. So while the solution I propose should work for those dispositions which feature in laws, I will not attempt to argue that it works for all dispositions. The solution I propose stems from two basic insights which I discuss in turn. First, that the analysis of dispositions can be rephrased in terms of the system-types and behavioural properties found in laws. This will help us propose an new schema for analysis specifically of component-level dispositions, *SysCA*. The second insight is that the embeddability of systems is not a maskable disposition. Indeed, I will show how an analysis of it of the form of StoCA is not subject to masks. Putting the two insights together I will show how instances of SysCA will not be subject to masks either.

Insight #1: component-level dispositions can be analysed in terms of laws

To show why the first insight is true, we begin by noticing that laws' component-dispositions have 'correspondent variable assignments' (schematically: V_D for disposition D). This means component-dispositions can be defined in terms of a permitted range of values of specific variable properties which appear in the behavioural formulae of certain laws in which they feature. So, for instance, some object has mass if and only if $m > 0$, where m is the object's mass; some object is electrically conductive if and only if $0 \leq R < \infty$, where R is the object's resistance; some object is frictionless if and only $\mu = 0$, where μ is the object's coefficient of friction; and some object is inextensible if and only if $k = \infty$, where k is the object's Hooke constant.

That a disposition has a correspondent variable assignment is extremely helpful, for it suggests we might be able to define it by drawing on the formula in which its correspondent variable appears. For instance, we might think that the manifestation of inertial mass is defined by having impedance to acceleration a under the stimulus of a net force F; a relationship represented in the total force equation,

$$\mathbf{F} = m\mathbf{a}$$

for $m > 0$. Similarly, the manifestation of electrical conductivity is an inversely proportional resistance to electrical current I under the stimulus of voltage potential V; a relationship represented in the behavioural equation of Ohm's law,

$$V = IR$$

for $0 \leq R < \infty$. The manifestation of frictionlessness is an absence of contribution to frictional force F_F under the stimulus of a reactive force R perpendicular to the plane of movement; a relationship represented in the behavioural equation for the dynamic frictional force law,

$$F_F = \mu R$$

for $\mu = 0$. The manifestation for inextensibility is an infinite resistance to extension x under the stimulus of an extensional force F_E; a relationship represented in the behavioural equation for Hooke's law,

$$F_E = kx,$$

for $k = \infty$.

In each of these cases, the formula in which the disposition's correspondent variable appears can be thought of as defining the manifestation property of the disposition and the remaining conditions of assembly for the law's system-type as defining the stimulus.[5] So, an analysis of component-level dispositionaliy

[5] Note that not *every* law's behavioural formula in which a disposition's correspondent variable appears can be considered as defining it. For instance the variable R for resistance appears not only in the equation for Ohm's law, but also in the equation for Pouillet's law of electrical resistivity $R = \rho\frac{l}{A}$, which relates resistance to length l, cross-sectional area A and specific electrical resistivity of the material ρ. While resistivity can be calculated with this equation, one might think that it doesn't describe the manifestation conditions which identify resistance. Resistance is a *resistance of flowing current*, not of length or area.

can be provided by looking for those laws in which they and their correspondent variables feature (ignoring further restrictions, see fn.5) and rearranging the conditions of assembly appropriately.

For instance, an analysis of conductivity would proceed by identifying a law in which it features as a condition of componency and its correspondent variable appears in the behavioural formulae. An appropriate law would be Ohm's law, which may be presented in unmodified form as follows (with conditions of assembly made explicit).

> **Ohm's law (unmodified):** For all x, if x is a system S comprising a conductive component, a source of charge in the component and a voltage potential across it, then the resistance of the conductive component R, the voltage potential V and the flow of charge through the resistor I are related by the formula,
>
> $$V = IR$$
>
> .

Here we clearly see how the dispositional property of being conductive and the manifestation behaviour of taking a certain value in systems behaving according to $V{=}IR$ are displayed in Ohm's law. By pulling out the system-type's condition of componency that some component be conductive we establish the left hand side of our biconditional analysis of conductivity. The right hand side is given by following the structure of StoCA, using the further conditions of assembly to define the stimulus property and the behavioural formula of Ohm's law with a specified range for the disposition's correspondent variable as the manifestation property.

The fact we know this suggests there is some causal directionality in dispositonal ascription which is lost when we analyse manifestation in terms of a mathematical equation. I assume that such directionality is something to be addressed once we have properly considered whether the correspondent variables are focal-variables or causal variables—something to be determined by the laws' variable-level causal asymmetries.

Analysis of conductivity (first pass): For any x, x is electrically conductive if and only if it has a natural property F which if had by the component of a system (of which there exist instances) which also comprises a source of charge in the component and voltage potential across it, then the voltage potential V and flow of charge through the component I are related to resistance R by the formula,

$$V = IR,$$

. where $0 \leq R < \infty$.

Examples of F in such a case are copper, carbon, and liquid helium, and indeed the antecedent of any SVL which has conductivity as a consequent behavioural property (see §5.4.4).

Analysing dispositions in this way shows the close analytical connection between laws and dispositions. In general, I suggest we approach analysis of component-level dispositions by drawing on the very laws in which they feature. One way to conceive of this approach is that we are treating the analysis of a component disposition as in fact concerning the system-types in which components feature. Begin with the form of a law:

$$\forall x(Sx \rightarrow Bx)$$

where S is a type of system and B a type of behaviour. Now recall Schema 3 which indicates that the criterion of identity for a system-type S is given by a set of conditions of assembly divisible into conditions of componency $C_1x, C_2x, ..., C_nx$ and conditions of organisation $O_1x, O_2x, ..., O_nx$, i.e.,

$$\forall x\Big(Sx \leftrightarrow \Big(\sum_1^n C_i x \;\&\; \sum_1^n O_i x\Big)\Big)$$

where $\sum_1^n X_i$ to indicates a conjunction X_1 & X_2 & ... & X_n. For many laws some condition of componency $C_j x$ will be that the system x has a component part y with some dispositional property D. Hence, the following translation schema will hold.

$$\forall x(C_j x \leftrightarrow \exists y \exists j(P(y,x) \ \& \ Dy))$$

where j ranges over numbers, x over potential components, y over potential systems, and '$P(x,y)$' says that x is a part of y. Since conditions C_i can be arbitrarily ordered, we can stipulate that for some disposition D of interest, this translates the n^{th} condition of componency (i.e. C_n) for systems in which the disposition is required. Hence we get the following variation on StoCA: a *systems conditional analysis*.

$$\textbf{(SysCA)} \quad \forall x\Big(Dx \leftrightarrow \ \exists F\Big(F(x) \ \& \ \exists y \exists z\Big(Fy \ \& \ P(y,z) \ \& \ \textstyle\sum_1^{n-1} C_i z \ \& \ \sum_1^n O_i z\Big) \ \& \ \forall y \forall z\Big(Fy \ \& \ P(y,z) \rightarrow \Big(\Big(\sum_1^{n-1} C_i z \ \& \ \sum_1^n O_i z\Big) \rightarrow B^{V_D} z\Big)\Big)\Big)\Big)$$

The 'B^{V_D}' indicates that the behaviour-predicate includes an assignment of some range to the disposition's correspondent variable V_D.

I admit SysCA looks monstrous, but the basic idea is straightforwardly built from ideas already developed in foregoing sections and chapters: any object x has a disposition D if and only if there is some natural property F which x has such that if had by any object y (of which there exist instances) which is a component part of a system z which itself satisfies further conditions of componency $C_1, ..., C_{n-1}$ and conditions of organisation $O_1, ..., O_n$, then that system z will behave in the way described by the behaviour-predicate B with a restriction on values of the disposition's correspondent variable.[6]

[6] It may seem plausible that in some cases, the natural property F will simply be captured

Of course, by formulating analyses of dispositions in this way we don't see immediately how to deal with the problem of masks. But it is the start of a solution which is completed once we make use of the behaviour-modification approach to the embedded phenomenon intuition.

Insight #2: embeddability is not a maskable disposition of systems

I said that the solution to the problem of masks (at least for those dispositions which feature in laws) requires two basic insights. The first is noticing that all component-level dispositions can be given an analysis in terms of the very laws in which they feature. The analysis involved reinterpreting the dispositional property as one ascribable to *system-types* in which the component disposition features. But what does all this formal machinery for reinterpretation and system-based analysis give us? For one thing, it may not appear enough on its own to solve the problem currently under investigation, viz. the problem of masks. For as we have seen in §3.3, in its unmodified form Ohm's law is subject to failures of universality as a result of interference, effectively masking the conductivity of components. But here is where we can combine the first with the second insight: that embeddability is not a maskable disposition.

Sometimes when a law makes a generalisation of the form $\forall x(Sx \rightarrow Bx)$, it seems to be rendered false by the existence of counterinstances (Ss which are not Bs). This would be a failure of the the law's *universality*. But for most theorists of laws, laws are supposed to be true (Cartwright 1983, Lange 2000, are two rare examples which suggest otherwise due to this very issue). So some sort of 'fix' is needed. Solutions to apparent failures of universality in laws have invariably assumed what I called the 'embedded phenomenon intuition', the idea that even in the case of counterinstance, there is some underlying phenomenon described by the law which contributes to the overall

by the system-type's conditions of componency, in which case the condition on F would be redundant. However, as I have tried to make clear elsewhere (see §5.2.1), conditions of componency are often dispositional rather than categorical, thereby establishing the need for Storer's extra condition on a natural categorical property.

behaviour. This suggests in the first place that we might render the false law true by a modification of it in some manner.

My specific proposal in §3.3.3 of modifying the behaviour-predicate, captured in Schema 5, in effect displays the manifestation conditions for the law-level disposition of embeddability of instances of system-types with respect to certain formulae. Extracting the essential features, we get the following analysis of embeddability in a form consistent with StoCA.

Analysis of embeddability:

For all assemblies x, x is embeddable with respect to some formula K if and only if x has a natural property F which if had by something y (of which there exist instances) such that y is part of an assembly z, then the overall behaviour of z has K as a special case.[7]

Notice, there's no mention of any system-type property here. After all, while we have no reason to expect all assemblies to do so, it is an open question for each assembly whether it might be embeddable with respect to some formula in the sense that the formula is a special case of the behaviour of any larger assembly in which it is a part. Nonetheless, it *is* a feature of the system-types which feature in laws, especially RCJLs, that their instances are embeddable with respect to certain formulae featuring in those laws' behaviour-predicates.[8] It was this which made the behaviour-modification a plausible solution to the problem of non-universality (§3.3).

As was noted in §3.3.3, what makes behaviour-modification a general solution to the problem of non-universality is the assertion that any external interference with the instance x of a system-type with respect to the formula K associated with it by law will be captured under a (larger) system-type for a non-idealising law whose behavioural formula has K as a special case. This

[8]Again, it might be tempting to identify the natural property F with system-types S in some cases. I refrain from making that suggestion here for reasons noted in fn.6.

is made plausible by acknowledging the clear mathematical relationships between known idealising and less idealising laws (e.g. in electronics, fluid dynamics, and thermodynamics).

If this assertion concerning the interference of instances of laws' system-types is correct, then it is clear that the embeddability of systems with respect to the formulae with which they are lawfully associated cannot be masked. To see this, first assume otherwise. In that case, there will be instances of a system-type which are part of larger assemblies such that the overall behaviour does not have as a special case the formula associated with that system-type by law. This means there can be no law $\forall x(S'x \to B'x)$ whose system-type S' has the larger assembly as a part and where the formula of B' has the formula as a special case. Since that is contrary to the assertion, we can conclude that embeddability is not maskable disposition of systems given the assertion (though it may be for mere assemblies). One way to think about why we get this result is that the stimulus for emebddability simply *is* interference, so long as it can be captured in the antecedent of a law, which all systems can be according to the assertion made in §3.3.3.

Putting the insights together

Consider again Ohm's law, but this time in its behaviour-modified form.

> **Ohm's law (behaviour-modified):** For all x, if x is a system S comprising a conductive component, a source of charge in the component and a voltage potential across it, then there is a system-type S' and behaviour-type B' such that,
>
> 1. $\forall x(S'x \to B'x)$ is a law,
> 2. all instances of S' have an instance of S as a part,
> 3. $\forall x(B'x$ if and only if the resistance of the conductive component R, the voltage potential V and flow of charge through the resistor I are related by the formula K' which

has as a special case the formula,

$$V = IR)$$

.

As we just saw, behaviour-modification effectively incorporates the embeddability of instances of system-types. So, following the steps of expressing dispositional analysis in terms of system-types to form a biconditional of the schematic form of SysCA, but now drawing on the *modified* version of the law, we get the following analysis of conductivity.

> **Analysis of conductivity (second pass):** For any x, x is conductive if and only if it has a natural property F which if had by the component of a system S (of which there exist instances) which also comprises a source of charge in the component and voltage potential across it, then there is a system-type S' and behaviour-type B' such that,
>
> 1. $\forall x(S'x \rightarrow B'x)$ is a law,
> 2. all instances of S' have an instance of S as a part,
> 3. $\forall x(B'x$ if and only if the voltage potential V and flow of charge through the component I are related to resistance R by a formula K' which has as a special case the formula,
>
> $$V = IR,$$
>
> where $0 < R < \infty$).

As far as I can see, there are no plausible masks which present a counterexample to this analysis (at least, not without breaking a law of nature). If a circuit is passing through a magnetic field, it may in fact be the case that there is no current flow through a conductive component despite there being a voltage potential across it (a mask of the disposition under typical construals). But the analysis is consistent with that, so long as a law $\forall x(S'x \rightarrow B'x)$ which does describe the interaction of magnetic field and electrical component also satisfies conditions 1–3. In a non-gravitational context, one might plausibly argue that the Lorenz force law is such a law, although I will not aim to prove that here.

In summary, SysCA remains a satisfactory schema for conditional analysis of component dispositions so long as we treat the behaviour-predicate B (and so also B^{V_D}) in each case as modified in the way suggested in §3.3.3. Consequently we have a way of analysing the dispositional features at component-level and law-level in RCJLs by drawing on the connection between the problem of masks for dispositional analysis and the problem of non-universality for laws.

6.1.4 Concerns about regress, circularity and triviality

I have argued that some kind of conditional analysis, viz. SysCA, can be defended against masks in the case of the dispositional properties which feature in laws' conditions of assembly. Notice that this is like none of the defences of SCA, or derivatively of StoCA, which I discussed in §6.1.2. It is not like Choi's because it accepts that dispositions are genuinely subject to masks, e.g. an electrical component remains conductive even when in a magnetic field which masks its conductivity. Like Bird's defence SysCA does rely on a distinction between dispositions which can be provided a strict conditional analysis and those which may not. But the distinction I advocate is not one required to carve along the divide between fundamental and non-fundamental properties or laws. Rather, it distinguishes systems which feature in strict (though potentially behaviour-modified) laws and mere assemblies which don't. It might in fact be that this exactly extends to all and only the fundamental dispositional properties which cannot be multiply realised, but there is no *in principle* reason for that (unless we simply defined 'fundamental' to apply to such properties).

SysCA is also unlike Manley and Wasserman's analysis in that it maintains a strict relationship between stimulus and manifestation rather than weakening the relationship to a probabilistic one under contextual parameters. This way SysCA does not suffer from the kind of relativism that contextual sensitivity brings. The analysis also treats masks as influencing the manifestation of dispositions on a continuum rather than either entirely masking or not masking at all as Manley and Wasserman's analysis does. This way SysCA can better represent the many quantifiable causal relationships between masks and dispositions science appears to reveal to us.

Of course, our ultimate goal in this chapter is to analyse *all* the causal features of RCJLs, and the combination of SysCA and the analysis of embeddability only gives us a way to analyse the *dispositional* features of RCJLs. However, I think there are foreseeable concerns as to whether they can even do this.

For example, one might be concerned about the fact that the modified B^{V_D} predicate on the right hand side of SysCA makes reference to other laws. I have already dispelled concern that this would result in some kind of unending regress. As I said in §3.3.3 and reiterated above, the hypothesis that there is always some law describing accurately some system in which others are embedded is requisite, but it needn't imply that we know such laws or that there are many instances of such laws. I believe we should simply commit to the idea that some laws are strictly universal without the requirement of behaviour modification.

Another reason to be concerned is that the further laws referenced in those which *are* modified won't be RCJLs. The point of accounting for the dispositions specifically in RCJLs is due to the expectation, motivated in §5.4.5, that other laws could be defined derivatively from them. So it would be a problem if the laws referenced in RCJLs with modified behaviour-predicates were not themselves RCJLs. However, I think it is reasonable to think that they will be. Take any law $\forall x(S'x \rightarrow B'x)$. Now assume that there is an RCJL $\forall x(Sx \to Bx)$ where S is a part of S' and the behavioural formula of B is a special case of the behavioural formula of B'. Then *by definition*, $\forall x(S'x \to B'x)$ will be an RCJL. This is because the dispositional components of S will also be dispositional components in S' and any instance of a system-type of a law which is universal without modification will be embeddable trivially (since the behaviour of all assemblies its instances are part of will be defined by the behavioural formula itself). Moreover, because of the embeddability of the instances of the law's system-type, the causal asymmetries between variables in the associated formula in B will be causal asymmetries in B'. Hence, $\forall x(S'x \to B'x)$ satisfies all the requirements of being an RCJL.

Of course, this suggests that wherever a behaviour-modified law can be pressed into service to analyse a component-disposition, there will be a law which is not in need of modification which can do the job equally well. Moreover, we might think the latter law is better than the former, since it will be more precise (specifying actual behaviour rather than special case behaviour). I am happy to accept such an outcome of the approach developed above. However, SysCA provides a way for us to use actual laws that we know to analyse component-dispositions. Practically, then, it serves enormous benefit.

A different set of worries one might have about the analysis which certainly needs addressing concerns the involvement of laws in the analysis of dispositions. Earlier I rejected any analysis of dispositions which employed counterfactuals (like SCA, MWCA) since the truth of counterfactuals appear to be grounded themselves in laws. According to standard semantics, the laws need to be established independently of dispositions in order for counterfactuals to analyse dispositions. Since I want to use dispositions to account for laws, the analysis would be circular. However, the extensional analyses for component-level dispositions which I have come down in favour of also draws on laws. SysCA was formulated with the insight that component-level dispositions could be defined with the conditions of assembly and behavioural formulae of the laws in which they feature as (respectively) stimulus and manifestation properties.

Despite the apparent concern, I do not think it would be right to deem the employment of SysCA given the current project as problematic as SCA and other counterfactual-based schema of analyses. SysCA does not require an *independent* analysis of laws, but only a *rearrangement of the information contained within* laws. The analysis of component-level dispositions and the laws which they are rearrangements of come together as a single informative set of generalisations. In using certain laws to form analyses of dispositions we do not, thereby, prohibit a non-circular account of laws which defines some of them—viz. RCJLs—to have dispositional conditions of assembly.

But even if the analysis does not open the project up circularity, there may appear in this claim cause for a related concern. For even if laws need not be established independently for the analysis of dispositions in order for the latter to work, there appears to be a resulting triviality in our understanding of both phenomena. For RCJLs have been defined partly by their system-types having dispositional conditions of componency and SysCA defines those dispositions in terms of the very RCJLs in which they appear. It is reasonable therefore to conclude that a dispositional requirement on a law somewhat trivialises the law itself. For example, having a component which is conductive is a condition of componency for the system-type in Ohm's law. But the disposition's stimulus conditions are defined by the other conditions of assembly for that same system-type and its manifestation conditions are defined by the law's corresponding behaviour-predicate. So, according to the analysis of conductivity (p.230), by saying that something is conductive we say (roughly) that it can play the role of a certain component in instances of Ohm's law. Consequently, as a condition of componency for Ohm's law, *having a conductive component* simply amounts, roughly speaking, to requiring that something play whatever role is required to fit with the rest of an assembly in order to bring about the relevant behaviour.

It therefore tells us very little about what to look for if we want to build or observe the relevant system-type.[9]

Perhaps this triviality is no surprise, especially given the often-noted 'dormitive virtue' concerns with dispositional explanations in general. Indeed, I think one should just embrace that some laws have conditions of assembly for their system-types which require, roughly speaking, that a component is *whatever will fit in with the rest of the assembly such that it satisfies the law's consequent behaviour-predicate*. The triviality of one condition of assembly is not obviously a problem for the analysis or the causal-junctions account of laws. After all, it is perfectly consistent with a disposition being partly defined by some behaviour that the component which has it is causally relevant to the behaviour. We need only recall Davidson's (1967) observation that the fact the cause of some event is its cause is completely trivial, but not false, to notice this. More relevantly, (Ellis 2002, Mumford 2004) have argued that dispositional explanations may have an element of vacuity regarding the causally relevant property of some behaviour but are still able to inform us of the causally relevant *entity* (it is trivial that it is the *dormitive virtue* of some pill rather than another of its properties that makes one sleepy, but not trivial that taking the *pill* rather than some other action makes one sleepy).

A more pressing concern regarding the triviality of dispositional conditions of componency is that if each condition of assembly for a law were dispositional then a law would end up conditioning behaviour on itself. I take it that *this* would be a disaster, rendering laws completely uninformative. So, in saying that certain laws are robust, in the sense that they have dispositional conditions of componency, we need to make sure that they are not so robust that they are true trivially. However, in general I don't think we should be

[9] Of course, given the extensional analysis, what the analysis actually requires is that the component shares some natural property with some object which actually *does* play that role in an assembly and that all things with that property which are in such an assembly also play that role. Perhaps this makes the analysis marginally more informative than I make out in the main text.

too concerned that this is going to be the case. As well as conditions of componency, all laws which have dispositional conditions of componency also have conditions of organisation too (recall that SVLs have neither). Whilst conditions of componency might conceivably consist entirely of dispositional conditions, conditions of organisation will not. This alone should be enough to save any such law from complete vacuity. And I think it's reasonable to suspect further that many laws' behaviour predicates go far beyond that which is necessary to define their system-types' component-level dispositions. For example, conductivity is a dispositional condition of componency for many laws (e.g. Kirchhoff's laws, Ampere's law, etc.), but its definition is exhausted by the information contained in Ohm's law. So, while Om's law may be trivialised *somewhat* by the dispositional condition, these other more complex laws will be even less so.

6.1.5 Summary

This concludes my analysis of the dispositional features of RCJLs. Both component-level and law-level dispositions have been analysed using an extensional conditional analysis along the lines proposed initially by Storer (1951). The idea has been to analyse the law-level disposition of embeddability by treating the stimulus condition to be a system's parthood of a larger system and defining the manifestation condition in terms of the behaviour-modification of idealising laws proposed in §3.3.3. Component-dispositions have been analysed in terms of specific laws whose system-types they are conditions for, treating the rest of the system-type's conditions of assembly as stimulus conditions and the behaviour as a manifestation condition. If we treat the behaviour-predicates of those laws as modified (since their system-types are embeddable), then I have argued neither analysis suffers from the problem of masks.

Due to its extensional form and quantification only over first-order particulars and categorical properties, if the analysis does its job, then the robustness of RCJLs can be given a Humean treatment. This takes us one step closer to a Humean account of laws in general which respects the causal-junctions conception of laws. The other necessary step is a Humean treatment of the other pair of causal features associated with RCJLs, viz. variable-level and law-level causal asymmetries. It's to that task I now turn.

6.2 Analysis of RCJLs' causal asymmetry

In §5.2.3 and §5.2.4 I drew attention to two kinds of causal asymmetry which exist in instances of RCJLs. 'Law-level asymmetry' is that which exists between an instance of a law's system-predicate and its behaviour-predicate, e.g. between an object's satisfaction of the conditions required to be an electrical conductor and a behavioural property defined by having variable-properties which satisfy the equation $V = IR$. This asymmetry was distinguished from 'variable-level asymmetry', which exists among the variables referenced in the behaviour-predicate, e.g. as a result of V and R being causes of I. In the case of RCJLs, the asymmetries at both levels combine to form a *causal junction*, in which there is a single effect or 'focus' of all the causal relations at that level.

In this section I propose a Humean treatment of the asymmetries in causal junctions. My basic idea is to develop a probabilistic analysis for what I will call 'causal theories for causal junctions'. Under certain plausible assumptions, causal junctions will have instances or 'models' which exhibit a probability distribution over their variables unique to a certain kind of causal structure. This is due to the fact that the junctions' foci in these models can be identified as approximating 'unshielded colliders' along a set of causal paths (Spirtes et al. 2000). By identifying the relevant conditions for these models with those in cases in which the causal junctions may be embedded, we can infer the existence of the same causal junction even if there is a qualitatively different probability distribution. Ultimately, what this means is that the proposal allows the asymmetries at the variable-level of RCJLs to be determined by the mosaic. A similar method of analysis will then be provided for the asymmetries at the law-level.

I begin in §6.2.1 with a defence of *actual frequentism*, an interpretation of probability which I take to be consistent with a Humean approach to the causal-junctions account of laws. In §6.2.2 I then introduce the formalism behind causal modelling and in §6.2.3 develop the concept of a *causal theory for a causal junction*. The probabilistic analysis of causal junctions at the variable-level ensues in §6.2.4 where I draw on the Causal Markov Condition and the condition of Minimality in order to define the assumptions required for the analysis to work. In §6.2.5 I then show how the existence of causal junctions at the law-level can be treated with the same probabilistic analysis thereby giving us an analysis of law-level causal asymmetry as well.

6.2.1 In defence of actual frequentism

It would be deceptive to offer a probabilistic analysis of causal phenomena under the auspices of a Humean methodology without acknowledging the various interpretive issues that surround probability. For some interpretations of probability are commonly thought to be unHumean, and others incompatible with the current project for more specific reasons. After mentioning some of these and defending Humean interpretations in general, I will defend the actual frequentist interpretation of probability over other Humean options. This interpretation deems the probability $P(A)$ of some outcome A in some reference class B to be the actual frequency of outcomes A in B.

Certainly, some interpretations of probability, e.g. a *propensity* interpretation, are often noted for their incompatibility (or inharmoniousness) with the Humean approach (Hájek 1997, Hájek 2012, Bradley 2015). One could hope to pursue a Humean analysis of propensities (perhaps along the lines of analysis for dispositions offered in §6.1), but I will assume such an approach is not optimal. Other interpretations also seem not to be in keeping with the current project. For instance, while I suspect a subjectivist interpretation of probability (e.g. Ramsey 1950 (2^{nd} edition, de Finetti 1937) could be made to work with the Humean approach, it might rob us of the desire to develop an account of laws which is properly objective. Pending any reason to think this cannot be achieved, I will assume probability is to be ridden of the any hint of subjectivity at play in the causal analysis.

Interpretations of probability more naturally compatible with Humeanism typically determine probabilities to be a function of what happens in the mosaic throughout the entire history of the world. As with laws, such interpretations are non-local, in that they determine the probability of some isolated event partly on the basis of what's going on throughout the entirely of time and space. However, since the very trials to which we assign probabilities are part of history, it is inevitable that the Humean interpretations end up partly grounding the probability of a certain trial's outcome in that very trial itself. This has lead to a number of complaints echoing those presented to Humean accounts of laws, many of which can be characterised under the basic intuition that probabilities should obey a certain sort of explanatory independence from the trials in which they are revealed (Hájek 1997). But as much as the Humean has learned to dismiss complaints over their view of laws, I believe many of these complaints can be dismissed on the grounds that they are either unmotivated or simply mischaracterise the Humean position (in what follows I draw for complaints exclusively on Hájek's 1997 compilation of arguments against finite frequentism).

For instance, it has been complained that Humean interpretations suggest curiously simple arguments against solipsism or for backwards causation due to their implication that probability is dependent on events beyond those under scrutiny (including events in the future). A Humean should simply respond that such criticisms stem from the undermotivated assumption that probability is an intrinsic characteristic of individual trials.

A similar complaint is that Humean interpretations suggest that probabilities about our personal futures implausibly imply that we should be concerned about the fate of others in the same reference class beyond mere empathy. This is an analogue of Kripke's 1980 Humphrey-objection to possible worlds analysis of modal claims and can be dismissed in the corresponding way by pointing out that under Humean views probabilities about our futures do depend on events external to our own causal futures but this doesn't prohibit them from *concerning* ourselves.

Another issue often levelled at Humean interpretations of probability concerns relativity to a reference class. Any view which determines probabilities at least partly by frequency of outcomes must say exactly what class or sequence counts as a relevant trial for the probability of an outcome. One problem with this is that it means the intuition that events simply have a probability *simpliciter* cannot be satisfied. As Hájek (1997) suggests, I think this is something the Humean should simply embrace, pointing out that 'knowledge of inequalities between conditional probabilities might be all that we need in order to control our environment in desireable ways' (p.74). Anyway, I doubt that the notion of a probability simpliciter is really any more intuitive than conditional probability.

A different problem is the difficulty associated with specifying exactly what the reference class is in any case. Under the Humean approach probabilities either seem ambiguous, vague or undefined. In response to this problem, Humeans will likely want to fall back on the resources of the mosaic itself. If we are of Lewis's persuasion, then the perfectly natural properties and individuals will determine the appropriate reference class (see Lewis 1994); if we are more pluralist about the contents of the mosaic, then something more nuanced must be said (see Cohen and Callender 2009, 26–30). I will assume some position along either of these lines is sufficient.

Admittedly, I think the Humean may have to accept some genuinely counter-intuitive results. One such intuition is that stability of a frequency of a certain outcome across a number of trials should be explained by the probability of that outcome. But under Humean approaches, probability is grounded in frequency (in some manner or other) and so arguably cannot explain it on pain of circularity. The problem here is much the same as that presented in §2.3 against laws' apparent lack of explanatory power under Humeanism. I argued there that even if we accept the result against Humeanism, it is far from clear that any other approach has done a better job of restoring the explanatory power of laws. I predict the same result for interpretations of probability alternative to a Humean one (see, for example, Lewis's 1994 arguments against the explanatory power of nomic necessitation with regard to probabilities).

Some complaints with Humean interpretations of probability don't stick under certain further assumptions. For instance, some Humeans have embraced NP as a modal characteristic of laws (see §2.2). NP entails that all laws remain true under any counterfactual supposition consistent with all of them (at least in scientific contexts). If that's the case then worlds in which an event goes without trial (e.g. the half-life of some atom of plutonium in a short-lived universe) may nonetheless retain the same probability as that in the actual world so long as the counterfactual supposition under which that world is considered is not inconsistent with the laws. Therefore, it's no good to complain that the Humeans are committed to an overly positivistic operationalism about probability (Hájek 1997, 79-80). Yes, for the Humean, the probabilities determined by our laws must be determined by what is trialed in this world; but, under the assumption of NP, it is also possible that a sequence can vary wildly from what is probable so long as this remains consistent with the laws.

Another problem is supposedly avoided by Humeans who accept a best system account of laws. This is the notorious 'problem of the single case'. If some outcome is trialed only once (its reference class has only one trial), or even not at all, then under a strict actual frequentism, the probabilities will be deterministic (the outcome has a probability 1 or 0 of happening) or else go undefined. According to Hájek, there are indeed such one-off events: 'a football game, a horse race, a presidential election, a war, a death, certain chancy events in the very early history of the universe' (p.81), subsequently actual frequentism can seem to be in trouble. One might wonder, of course, whether such events as these are really 'one-off'. Surely any attempt to assess the probability of an outcome in these cases will draw on other football games, wars, deaths, quantum mechanical outcomes. If these assessments can be understood as a recourse to other instances in some larger reference class then these trials are not one-off after all.

But the supporter of best-systems analysis need not push for such a response. The best systems approach aims to determine probabilities on the basis of initial or boundary conditions together with the laws which emerge from a best systematisation of the history of the world which best compromises strength, simplicity and fit (see §4.2). Fit is a function of frequency in the mosaic, hence if the actual frequencies diverge wildly from what is probable given a theory, this counts against the theory. But since the importance of fit in determining the correct theory for the world is compromised by strength and simplicity, the best-system interpretation suggests that probabilities may diverge from actual frequencies if it would make significant gains according to these latter parameters. As a consequence, it is plausible that the best system interpretation can provide non-trivial probabilities for single cases and even instanceless cases.

Now, the best system interpretation of probability is typically thought to rely on the success of the best system account of laws. One may recall that, in §4.2 I argued against the best systems account of laws due to its implicit commitment to the causation-mirroring conception of laws. But regardless, the interpretation would appear to be anyway unavailable since the best system account aims to determine probabilities with laws established via other means (i.e. through systematisation). In contrast, my project aims to provide an account of laws by first defining that subclass of laws, viz. RCJLs, partly *in terms of* the probabilistic relations which constitute the causal asymmetries associated with them. Perhaps the best-systems interpretation of probability could be employed in support of my account of laws if it were not assumed, as Lewis does, that the systematisation was a way to get both probabilities *and laws*. But I will not consider this option further. For I'm tempted to think it simply gets things wrong about probability anyway; a closer tie with fit is in fact preferable. Let me explain.

A commonly agreed on platitude of objective probability (or 'chance') is that it should constrain our rational credences. This idea is formally captured in the *Principal Principle*,

$$\textbf{(Principal Principle)}\ C(A|E.P(A) = x) = x$$

where $C(A|E.P(A){=}x)$ is the idealised rational credence we should give to an outcome A given the following conditions: any *admissible* evidence E (which may concern facts about the history of the world up to the trial which tests whether A and any general theory providing history-to-probability conditionals), and that the probability (or 'chance') $P(A)$ of A is x.[10]

The basic connection between probability and credence which the Principle aims to characterise is typically understood as something approaching self-evident, a conceptial truth (Hall 2004c), or analytic (Schwarz forthcoming). Moreover, formal arguments have even been presented in its favour drawing on the epistemic utility of credence functions which match objective probability (Pettigrew 2012, 2013). So there are good reasons to suppose it is correct (or thereabouts; see fn.10).

[10] Lewis (1994) argued that the Principle should be updated to accommodate the possibility of undermining futures. What I say here should be consistent with the update.

The important point for the sake of this discussion is that the Principal Princple suggests that whatever interpretation we give of probability, it should be one which makes sense of probability's role in constraining what is *rational to believe*. Now, the nature of rationality itself is a thorny issue. But I think we would do well to assume that something is rational to believe if it maximises expected utility; that is, if believing it brings about what is maximally subjectively valuable (other things held equal; see Briggs 2017). If that's the case, then surely interpretations of probability which are exclusively constrained by frequency are the most plausible. Consider a scenario in which God can tell you either the exact frequencies of each number for a series of die-rolls (which I assume does not violate any condition of inadmissible evidence), or tell you some predicted value for each roll determined by compromising information about frequency with considerations of strength and simplicity. If you had to make a bet on the outcome, you would surely rather the exact frequencies. It is frequencies and frequencies alone knowledge of which maximises utility. Hence, as Glynn admits, 'the advocate of the BSA requires some general explanation of why the best-fitting probabilities (the actual overall world frequencies) aren't automatically the best players of the chance role in guiding rational credence: i.e. why simplicity considerations constrain rational credence' (2017, 306).

Both hypothetical frequency interpretations (which interpret probability as frequency in infinite sample sizes) and propensity interpretations of probability also seem to compromise this utility-maximising feature of probability with a desire to satisfy intuitions about probability's independence from frequency. But whilst the actual frequentist interpretation of probability is often criticised for positing an overly tight connection with frequency, the above reasoning makes it seem reasonable to press precisely the opposite intuition: if probability is to play that role of maximising utility (as seems plausible given the Principal Principle), then actual frequentism comes out on top.[11]

[11] Arguably, Pettigrew's 2012, 2013 arguments support the same conclusion on the basis of consideration about maximising *epistemic* utility.

A similar point can be made about other concerns presented against actual frequentism. For instance, (Hájek 1997, 'problem 12') has complained that under the interpretation, any number of trials which is not a multiple of the possible outcomes will necessarily bias some outcomes (as would be the case if the total number of coin tosses were odd). He also complains (problem 13) that the interpretation implies many more correlations between apparently independent sequences simply due to the fact that the number of trials in the smaller of any two samples is unlikely to exactly divide that of the larger. And again along the same lines, Hájek complains that as the sample size increases the likelihood that bias will be introduced into the frequency increases (compare your expectation that a fair coin lands heads exactly 50% of the time in 2 and in 100 tosses). However, in all these cases, if we want to uphold a close connection between probability and what knowledge we should most desire in order to maximise utility (beyond knowing the exact order of outcomes!) it seems that actual frequency is exactly the right interpretation. After all, if a coin which is otherwise considered to be fair is known to come up heads 7 times in 11 tosses, it's better for the sake of that sequence to calibrate one's expectation of heads in any randomly chosen toss to 7/11 (assuming one doesn't also know the exact order of outcomes) rather than 1/2.

Lastly, I think the tight connection between frequency and probability should be willingly embraced in the case of existential and universal statements. Hájek complains that under actual frequentism, statements of the form 'the F exists' and 'all Fs are G' (including laws) can have no non-trivial probability of being true. *Pace* Hájek, that seems to me exactly right. But notice that this doesn't prohibit non-trivial probabilities for statements concerning the *process behind how the F came about*, or the *likelihood of things being G in general*.

Admittedly, actual frequentism has other strange results which might lead us to remain sceptical of the interpretation. For instance, it implies that there can be no irrational probabilities (assuming all probabilities' respective trials are finite). So it turns out that irrational number solutions to the modulus square of the wave-function can't represent exact probabilities of particle position, hence some tweaking of the 'Born rule' is required. But this shouldn't be considered a philosophical snub of science itself! The wave-equation may still be our best practically attainable method of determining probabilites for particle-position, save knowing the frequencies themselves. Moreover, the actual frequentist may simply wish to question the typically assumed tight connection between objective probabilities and what our best laws say.

I suppose that despite my foregoing efforts, one might anyway remain a sceptic about the actual frequentist interpretation of probability. It is worth stressing before moving on that it is not essential to the following analysis of causal asymmetries in laws that probability have this interpretation. There are at least two conceivable alternatives. Under one alternative, we provide a different interpretation of probability appropriate to the Humean methodology and the causal-junctions account of laws. On the second alternative, the analysis is pursued in terms of actual frequencies directly rather than probabilities. Having said that, I personally cannot see how the first alternative would proceed and I think it's better not to have to rely on the second, since the plausibility of the following analysis is greatly facilitated by the widely accepted connection between causation and probability. So I will continue under the assumption that probability *is* the actual frequency of outcomes within the appropriate reference class defined (somehow) by the contents of the mosaic.

6.2.2 Structural equations, causal models and causal theories

It is a central observation of the analysis of the causal junctions inferentially associated with RCJLs that they will have instances in which the junction's focal-variable approximates an 'unshielded collider' on a causal path leading from one of its cause variables to another. This section introduces some of the terminology required in order to develop this idea formally, specifically, *structural equations models, causal graphs* and *causal paths*.

Structural equations modelling is a technique employed (though not exclusively) for analysing causal relationships. According to the framework, a causal model for some situation in which the variable properties of objects are causally related is represented as an ordered pair $\langle \mathbf{V}, \mathbf{E} \rangle$, where $\mathbf{E}$ is a set of structural equations $\{E', E'', ...\}$ each of which relates the values of variables $V', V'', ...$ in the set $\mathbf{V}$.

Despite being equations, the members of $\mathbf{E}$ for any causal model do not represent purely symmetrical relationships but rather represent the type-level causal influence of the left hand variables on the right hand variables which are invariant across certain ranges of values. So, for example, an equation of the form $Z{:=}f_Z(X, ..., Y)$ should be understood as representing the fact that there is a range of values $x_1, ..., x_n, ..., y_1, ..., y_m$, such that X taking value x_i, ..., and Y taking value y_j (for $1 \leq i \leq n$, ..., and $1 \leq j \leq m$) are type-level causal influences of Z taking the value determined by the function $f_Z(x_i, ..., y_i)$.

I follow the convention of using the symbol ':=' to relate left and right hand sides of structural equations in order to keep structural equations visibly distinct from algebraic equations. Left-hand variables are denoted 'effect-variables' and right-hand variables 'direct cause-variables' relative to the equation (since an effect-variable in one equation may be a cause-variable in another and vice versa). Crucially, since the equations aren't symmetrical, it's not possible to rearrange them while preserving their content: what appears on the left hand (right hand) side stays on the left hand (right hand) side relative to each equation. However, this doesn't mean that rearrangements cannot on occasion preserve truth (perhaps implying cyclic conditions), nor that a re-arrangement is never useful in order to determine the numerical values of the variables.

Occasionally it is useful to represent a model graphically. One way this can be done is by placing the variables at the nodes of directed edges which represent the direction of causal relevance. Hence an equation $Z{:=}X+Y$ can be represented according to Figure 6.1 (p.248).

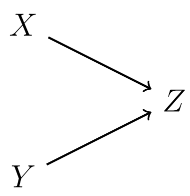

Figure 6.1

Using graphs we can define a number of useful notions (I follow here the terminology of Spirtes et al. 2000). A *direct cause* of some variable V is any variable at the node at the other end of a directed edge leading into V and a *causal path* is any sequence of variables in which neighbouring variables in the sequence are directly causally related. A causal path is *directed* if the direct causal relations go in the same direction in the sequence and is otherwise *undirected*. However, graphical representations are limited. It is not obvious, for example, how to distinguish graphically equations like $Z := XY^2$ or $Z := X-Y$. Typically the graph for both of these equations would be just the same as in the above diagram, even though the pattern of influence is radically different (see the discussion in Woodward 2003, 44-45). This limitation of graphs has drawbacks and benefits. While I will often prefer representation by structural equations for the sake of maintaining numerical precision, it will often be useful to refer to graphs when it comes to discussing embedded systems in which numerical relations are lost but the qualitative relations of causal influence are maintained.

One overall benefit of the structural equations approach (in comparison with, say, neuron diagrams; see Hitchcock 2007b) is that the values of the variables can be mapped to any set of numbers (natural, real, imaginary, complex, etc.), thereby enabling them to capture the precise kinds of functional dependencies which Mach (1900, 1905) and Russell (1912) decried vaguer understandings of causation for failing to accommodate (see §5.1). This is of particular benefit for the project of investigating the kinds of causal relationships which exist in *laws*, since the laws of science often describe complex numerical relationships between variables themselves.

Another similar benefit of structural equations modelling is that the contribution of many types of cause can be captured with familiar mathematical devices. Hence, negative effects can be captured with minus-signs or divisors, compound effects with multiples or powers, etc. Moreover, structural equations preserve numerical transitivity: causally correct equations remain causally correct if we substitute for any causally influential variable (on the right hand side of an equation in **E**) the function of its own causal influences in the model (see, e.g., Cartwright 2007, 154). This is helpful if we want to show the way in which variables can play a role in the effects of their effects.[12]

A final aspect often noted in discussion of structural modelling is that it relativises causal facts to a model or theory. But where I am happy to accept some relativisation for the purposes of what follows, I am less inclined to accept it all. Of course, a model which simply doesn't include certain variables is unable to display causal relations between them. As I will show in §6.2.3, the behaviour-predicate in RCJLs can be defined by a structural equation which is limited to the variables and causal relations which feature along a set of paths with a single 'collider variable' (to be defined shortly). This means that causal dependencies which may exist between the cause-variables or to and from other variables beyond the paths will not get mentioned in the model associated with the law. To this extent, there is relativistation of causal information associated with a law. However, I will

[12] Accepting numerical transitivity is not the same as accepting the transitivity of causation, which many in the structural equations literature want to reject for various reasons (see Kvart 1991, Hitchcock 2001b). See the discussion in (Cartwright 2007, fn.3, ch.5).

not assume that all variable-selection and causal dependency is relative to laws. After all, the account of laws to be provided is supposed to reflect something objective about the world. I will assume that there are some objective matters of fact about what variable-properties are actually present in the Humean mosaic. This will be paramount for closing off some apparent counterexamples to the Causal Markov Condition in §6.2.4.

Before moving on with the analysis, it is necessary to introduce some more terminology concerning the various levels of causation being analysed. The sort of causal information licensed by a causal model is typically understood to be at the type-level. This is to be contrasted with causal information at the token-level 'to the effect that some particular token event has caused another' (Woodward 2003, 40). This difference in level need not be interpreted as reflecting a contrast in different kinds of causation itself. Rather, a type-level claim 'depends on relationships that do or would obtain [...] at the level of particular individuals' (ibid). We can maintain that causation occurs at the level of token events, where individual events are classified as the instance of a value of some variable property. A type-level claim of the form "X is causally relevant to Y" then amounts to the claim that in general, when a value is instanced by variable property X it causes a value to be instanced by variable-property Y.[13]

Woodward has claimed that type-level causal relations are supposed to be 'cases in which the relationship between variables X and Y is general or reproducible in the sense that Y exhibits some sort of systematic response when the same changes in the value of X are repeated, at least in the right circumstances' (Woodward 2003, 41-42). The *repeatability* of relations between variables mentioned in type-level causal claims validates the suggestion that the relationship is at a higher level of abstraction than that picked out by token-level causal claims. But it is not always made clear at

[13]More correctly, there exist values of X which when instanced cause values of Y. Although this nuance will be irrelevant for the examples I'm considering.

what level of abstraction these types are supposed to be. Getting clearer on this will be important for what follows.

Woodward has admitted elsewhere with Hitchcock that 'to count as invariant and hence explanatory, a generalisation must describe a relationship that holds for certain *hypothetical* values of X and Y *possessed by the very same object* o' (Woodward and Hitchcock 2003, 20). Putting aside the authors' argument for invariance, it is clear from the context that they are talking of generalisations formulated as structural equations. Though operating at the level of variables-properties, a 'type-level' causal claim appears to be still restricted to a particular (*token*) system—*'the very same object'*. To avoid misunderstanding, I will make such restrictions explicit from hereon by adding subscripts to the variables in causal models which denote the system they apply to, e.g. causal model $\langle \mathbf{V_a}, \mathbf{E_a} \rangle$ is the particular token system a).

Making the restriction to particular systems explicit allows for the representation of a less restricted sort of type-level causal claim. This takes into account the abstraction from causal relevance between variables-property of particular systems to causal relevance between system-*types*. Drawing on the terminology of model-theoretic semantics, let's say that a 'causal theory' has causal models which interpret the causal structure of the theory as causal relationships exhibited by a particular real-world system. In such a case I will say the system 'models' the theory. A causal theory is denoted $\langle \mathbf{V}, \mathbf{E} \rangle$ (NB no subscripts!), where $\mathbf{E}$ is a set of equation-*types* $\{E', E'', ...\}$ and $\mathbf{V}$ is a set of variable-*types* $\{V', V'', ...\}$. The members of $\mathbf{E}$ and $\mathbf{V}$ thus correspond, respectively, to the equations (e.g. $\{E'_i, E''_i, ...\}$) and variables (e.g. $\{V'_i, V''_i, ...\}$) of each of the theory's interpreting models $\langle \mathbf{V_i}, \mathbf{E_i} \rangle$.[14]

[14] Technically, it is necessary to put some constraint on what counts as a model for a theory. A model is any instance of the causal theory exhibited by a causally related assembly of objects. But if models were restricted to occasions with the exact same enduring component parts it may turn out that certain variables cannot vary in value. For instance, it may not be straightforward within an assembly for a component part's mass to vary (in classical contexts at least). I suggest that models of causal theories be restricted to occasions under the same 'experimental context', where a context remains the same so long as component parts (with varying mass, say) are swapped in and out one at a time.

In what follows I will generally refrain from the added complexity of referring to the numerical identities between 'equation-types' of 'variables-types'. This would require denying, for instance, that $F = ma$ is strictly an equation relating variables—strictly speaking, it is an equation-*type* relating variable-types. In what follows I'll often use the word 'equation' and 'variable' to cover both theory and model levels of description and let the more formal representation of variables with or without subscripts do any needed disambiguation that isn't otherwise clear.

In the same way causal models license type-level causal claims which relate the variables of specific systems (instances of system-types), *causal theories* license causal claims relating the variable-types of system-types. Clearly both forms of causal knowledge can be represented in the form of structural equations and graphically in a similar manner. Once again, a distinction in levels of abstraction at which the claims are characterised does not commit us to a distinction between kinds of causation.

To make this distinction in levels of causal analysis clearer still, let's consider a highly simplified example. The dry-weight yield W_c of a particular crop in some field some year is causally influenced by the density ρ_c of the crop and growth characteristic C_c of the crop. Using the structural equations framework, we can represent this pattern of influence with the following equation (where f is some specific function).

$$W_c := f(\rho_s, C_c)$$

The causal model for this scenario is thus represented as the ordered pair $\langle\{W_c, \rho_s, C_c\}, \{W_c{:=}f(\rho_s, C_c)\}\rangle$. Here W_c, ρ_s, C_c represent the variables for the properties of some particular system c, viz. the field with planted crops. The model licenses causal claims such as "ρ_c is causally relevant to W_c". Also, in conjunction with information about the particular values taken by the variables (possibly with some further contextual facts) the model also licenses token-level causal claims, such as "the $10mm^2$ density of seeds sewn was a cause of the crop yield being 10 tons". However, we will also find it useful to be able to make causal claims of the form "ρ is causally relevant to W" where ρ and W are variable-types related by the equation-type $W := f(\rho, C)$. This latter causal claim draws on the existence of a causal *theory* about a *type of* system (hence no subscripts) and says that any system which models the theory (including c) will exhibit relations of causal relevance between the variables which are instances of the theory's variable-types.

As I will understand it, a causal theory and its various models are structurally isomorphic. That is, for all and only the causal relations between variable-types in the causal theory, there are causal relations between the corresponding variables in each model. That is not to say, however, that it is trivial what it takes for some model for a causal theory to exist. Alongside a causal theory (defined by a set of variables and structural equations) there will also be *model-conditions*. These provide the criteria according to which a particular system models the causal theory. In the above example, we may expect the model-conditions to require that systems comprise a particular crop which has been sewn in a particular field. Essentially, model-conditions are equivalent to conditions some assembly must satisfy in order to be an instance of the system-type for a law. Accordingly, we can expect the model-conditions to be divisible into conditions of componency and conditions of organisation (see §3.2.1). Those conditions in the considered example are, I take it, the same which define the system-type for Yoda's power law, in which the behaviour-predicate is defined by the non-structural equation,

$$W = C\rho^{-\frac{3}{2}}.$$

In general, it is this requirement for model-conditions for a causal theory which will ultimately help to tie the analysis of causal asymmetries at the variable-level developed in §6.2.4 to those at the law-level in §6.2.5.

On a final point about terminology, I realise that my chosen way of parsing things may clash with a related use of the term 'model' as something which is used to represent a target system (e.g. in the papers in Morgan and Morrison 1999, and see §3.2.2 above). Its not obvious to me, however, that the issue over terminology choice amounts to any philosophically significant disputes regarding the thesis put forward here.

6.2.3 Causal theories for causal junctions

We are now ready to start applying the terminology to the issue of determining the causal asymmetries in causal junctions. For the sake of simplicity, I focus in this section and the next on variable-level causal junctions and extend the analysis to law-level causal junctions only in §6.2.5.

As we have seen, structural equations are not the equations of laws. The former display causal asymmetries among their variables explicitly, the latter do not. Nonetheless, in §5.2.3 I suggested that there are indeed causal asymmetries among the variables in many laws' equations. We might think, therefore, that there should be an *interpretation* of such laws' equations in the form of a structural equation. Take, for example, Ohm's law. In §6.2 we saw that there is strong evidence that current I is treated as an an effect of resistance R and voltage potential V, its cause-variables. Textbooks on electronics will claim that current is 'resisted' to degree R by the component it passes through and that the voltage 'produces' the current. Rearranging the formula so that I is isolated on the left hand side and changing the identity sign from that of an 'unstructural' equation to one in a *structural* equation gives us,

$$I := \frac{V}{R}.$$

This tells us that voltage drop and resistance are causal influences of current. We can thus construct a causal theory of electrical resistance $<\{I, V, R\}, \{I := V/R\}>$.

Although nothing in the way I've introduced structural equations requires it, the structural interpretation of the equation in Ohm's law has a single variable on the left hand side and multiple variables on the right (some require this explicitly; see, e.g., Hitchcock 2001b). Notice that this is exactly the requirement for representation of a causal junction, since causal junctions have a single focus (in this case, I) which is causally influenced by the other variables (in this case, R and V).

To give some indication of how RCJLs' formulae may be given an interpretation as a structural equation, Table 6.1 (p.256) gives the structural intepretations of all the causal junctions tabulated in §5.2.3 (p.187). As before, the table is by no means exhaustive.

In each of these cases, the junctions' focal-variable is isolated on the left hand side. Notice how this is different from a plausible structural interpretation of the equation for a non-RCJL, such as that of the ideal gas law:

$$PV := nRT$$

Since P and V are both effects of temperature and quantity of substance, there are two variables on the left hand side and so there is no focal-variable.[15]

Each structural equation in Table 6.1 comprises the sole equation in what I will call a *causal theory for a causal junction*. Formally, causal theories for causal junctions have the form $\langle \mathbf{V}, \mathbf{E} \rangle$ in which the one and only structural equation $E \in \mathbf{E}$ relates a single endogenous variable to multiple exogenous

[15] Although in §5.4 I discussed various system-types (e.g. isobaric, isometric) in which one of these variables is singled out as the focal-variable. Presumably, the equations of some non-RCJLs will have no structural at all, e.g. the Planck-Einstein equation.

Table 6.1: Structural interpretations of laws' equations

Law	**Structurally interpreted behavioural equation**
Ohm's law	$I := \frac{V}{R}$
Total force law	$\mathbf{a} := \frac{\mathbf{F}}{m}$
Lorentz-force law	$\mathbf{F} := q(\mathbf{E} + \mathbf{v}\times\mathbf{B})$
Law of mutually gravitating bodies	$\mathbf{F} := Gm_1m_2/r^2$
Law of pendulum motion	$T := 2\pi\sqrt{\frac{l}{g}}$
Snell's law	$\sin\theta_2 := \sin\theta_1\frac{v_2}{v_1}$
Compton shift law	$\lambda' := \lambda + \frac{h}{mc}(1 - cos\theta)$
Archard's wear law	$Q := \frac{KWL}{H}$
First law of Thermodynamics	$\Delta U := Q - W$

cause-variables. Like other typical causal theories, it has certain non-trivial model-conditions which define when each theory has an instance. As I suggested in the case of Yoda's power law (§6.2.2), I will assume that the model-conditions for each causal theory are exactly the conditions of assembly for the system-type in each corresponding law.

This assumption is not trivial, for it implies that conditions of assembly are enough to maintain identical causal structure across instances. It comes under particular scrutiny when we consider cases in which systems are embedded with respect to theories for variable-level causal junctions in more complex scenarios. In such circumstances I will continue to say that the causal junction has a model in the context. However, the structural equations in that model may no longer be quantitatively accurate.

Taking what we have learned from the analysis of embeddability in §6.1.3 we can say that if some system is embeddable with respect to a causal theory then the single structural equation which describes it will always be a special case of some structural equation in a more accurate causal theory (or equally accurate in the trivial case where the system is embedded only in itself) for which the system contributes by satisfying part of the model-conditions. So, being embedded with respect to some causal theory entails not merely the mathematical fact that the characterising algebraic equation is a special case of the other but also that the causal structure of the model is maintained in embedded systems as well as.

To see how this works, return to the example of the causal junction associated with Yoda's power law. For a system c to be embeddable with respect to the causal theory for the causal junction associated with Yoda's power law must be the case not only that the symmetrical equation $W_c = C_c\rho_c^{-3/2}$ is a special case of whatever formula actually applies, but that structural equation $W_c := C_c\rho_c^{-3/2}$ is as well. So, if some other variable X_a is also causally relevant in the more complex system of which c is a part, it must be the case that

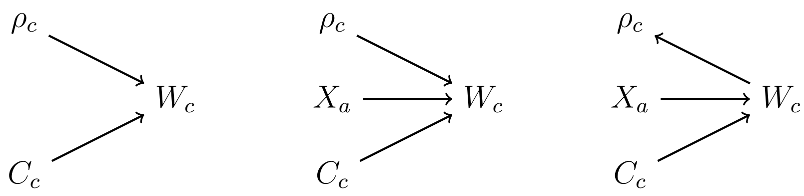

Figure 6.2

W_c is still preserved as the focal-variable of C_c and ρ_c. In terms of causal graphs we can say in general that the causal graph of the causal junction for associated with Yoda's power law is a 'subgraph' of whatever graph accurately represents the situation in which the model-conditions are satisfied, where for some graph G_1 to be a sub-graph of another G_2 the graphical structure of G_1 must be present in G_2.[16]

Figure 6.2 (p.258) comprises three graphs. The left hand graph represents a model of a causal junction associated with Yoda's power law by some assembly c; the middle graph represents a more complex model in which the model for the causal junction is a subgraph; the right hand graph represents another more complex model in which the model for the causal junction is not a subgraph. While the first and second graphs are consistent with the embeddability of instances of the variable-level causal theory for the causal junction associated with Yoda's power law, the third is not.

The structural interpretation of the equation of Ohm's law represents another example of a causal structure whos instances are typically embeddable. From the structural interpretation we can tell that any instance of the system-type in Ohm's law will be one in which the current is the effect of voltage potential and resistance. However, in many instances of the system-type of Ohm's law, the current will heat up the resisting component thereby affecting its resistance; and in some instances, a current is generated by a 'current pump' (e.g. in the sodium ion channels of the axons of the brain). These further causal

[16]The converse is not generally true, since graphs are not as precise as the causal theories or models they represent.

relations will be particular to the circumstances surrounding the model for the variable-level causal junction associated with Ohm's law. We might expect that those further causal relations will be predicted by the causal inferences licensed by some more accurate law whose system-type is also instantiated in those circumstances. Nonetheless, the structure of the causal junction entailed by a structural interpretation of Ohm's law is maintained in all these more complex cases.

In sum, we have now defined a causal theory for a causal junction and can refer to the equations which feature in laws in both structural and 'unstructural' interpretations. Of course, it is one thing to be able represent the causal relations in causal junctions when we know what the causal asymmetries are, but it is exactly finding an analysis for the latter which is in question. That is the aim of the following two sections continuing first with consideration of variable-level causal junctions.

6.2.4 A probabilistic analysis for variable-level causal junctions

The project of determining asymmetrical causal relations from the functional relations described in physical equations has been taken up on an number of occasions (see, e.g., Simon 1953, Spirtes et al. 2000, 46, Woodward 2003, 318).

Perhaps the most well known attempt at this project is that of Hans Reichenbach who makes explicit reference to the problem of formulae's ability to be rearranged or numerically 'converted'.

> It has been argued that, because of this convertibility of functions, physical laws do not define a direction or an order (cf. Russell 1912, quoted in §5.2.5). Functional relationship, it is said, is a symmetrical relation; if y is a function of x, then x is a function of y. Now there is no doubt that the latter statement is true, that

> functional relationship is, indeed, a symmetrical relation. But the question remains whether we have to conclude that causality is a symmetrical relation. The relation of causality may be of such a nature that it is not exhaustively characterized by the general concept of functional relationship. (1956, 28)

Reichenbach's proposal was a probabilistic analysis. It draws on the plausible principle capitalised on in many subsequent analyses of causation that *causes raise the probability of their effects* (see, e.g., Hempel 1965, Suppes 1970, Salmon 1977, Cartwright 1979, Eels 1991, Hausman 1998, Spirtes et al. 2000). As Reichenbach noticed, this is not alone sufficient to demarcate causal relationships. For it is a consequence of fairly undisputed probability axioms that if causes raise the probability of their effects then effects raise the probability of their causes; moreover, the effects of common causes raise each other's probability.

To get an asymmetry out of probabilistic analysis, Reichenbach noticed he could make use of the asymmetrical causal structure of 'future-directed causal forks' in which the probabilities of multiple effects of a common cause are 'screened off' by conditioning on that cause. But as important as this observation was (recall its relevance to the arguments from Frisch 2014 summarised in §5.1), it has been noticed that future-directed forks are in fact not the most appropriate phenomenon to analyse the asymmetry of causation Hitchcock (2016). This is especially the case if there is no time-order information to draw upon (as is the case with many variable-level and all law-level causal junctions).

It is now thought that asymmetries are most accurately captured by capitalising on the existence of *past*-directed forks, in which a single effect-variable has multiple causes. These past-directed forks are effectively causal junctions which, in situations satisfying certain specific conditions, exhibit a unique probabilistic character meaning that their asymmetrical structure can be unambiguously determined from statistical data.

Under a different terminology, this unique probabilistic character has been explored extensively by Pearl (1988) and Spirtes et al. (2000). Their project was one of developing algorithms for causal discovery. With some data for a set of variables, the algorithms they provided aimed to say what the causal structure is underlying those variables. Many of the algorithms were developed with the aim to aid discovery of causal structure in particularly complex scenarios where the data may may be incomplete or partially erroneous.

For the purposes of explaining how the objective probability distribution determines the causal structure between the variables in a law, we needn't investigate the success of these algorithms. After all, the project at hand is not one of causal discovery, but one of causal *analysis*. However, I believe that we can, to a certain extent, simply transfer many of the observations by Pearl, Spirtes *et al.* to form the basis of an analysis of causal junctions in terms of those probabilistic relations: according to such an analysis, the kinds of probabilistic dependencies and independencies which indicate a causal junction actually *constitute* what it is to be a causal junction.

The specific way which I will argue characterises the probabilistic analysis of a causal junction at the variable-level of a law can be given as follows.

The Probabilistic Analysis of Variable-level Causal Junctions

A causal theory $\langle \mathbf{V}, \mathbf{E} \rangle$ is a causal theory for a causal junction at the variable-level if and only if, it has at least three variables in $\mathbf{V}$ and only one structural equation E in $\mathbf{E}$ where,

(I) for each model $\langle \mathbf{V_x}, \mathbf{E_x} \rangle$ there is a system-type S' and behaviour-type B' such that,

  (i) $\forall y(S'y \rightarrow B'y)$ is a law,

  (ii) instances of S' satisfy (at least) the model-conditions for $\langle \mathbf{V}, \mathbf{E} \rangle$,

  (iii) $\forall z(B'z)$ if and only if z has variable-properties related by the formula K' which has a valid structural interpretation which has E as a special case.

AND,

(II) there is a variable W in $\mathbf{V}$ such that in at least one model $\langle \mathbf{V_i}, \mathbf{E_i} \rangle$,

  (i) W_i is probabilistically dependent on all other variables in $\mathbf{V_i}$, and,

  (ii) at least one variable in $\mathbf{V_i}$ other than W_i is approximately probabilistically independent of at least one other variable in $\mathbf{V_i}$ other than W_i, except when conditional on W_i.

The purpose of condition I is to make sure that all models of the theory exhibit the right *quantitative* relationships between the variables to count as an instance of a law. It incorporates a version of the behaviour-modification (see §3.3.3) tailored to causal theories thereby allowing us to include among the causal theories for causal junctions those associated with idealising RCJLs. Hence, the structural interpretation of E_x for any model of the theory need not accurately predict the values of the focal-variable given the values of its cause-variables. But by being a special case of a structural equation in a more accurate model, it is expected that whatever causal relations it describes will be at least qualitatively preserved in all instances. As before, if the modification is to be plausible for causal junctions associated with RCJLs, it requires that there do in fact exist laws which get things absolutely right (see §3.3.3). Moreover, these laws will also be RCJLs by virtue of having RCJLs as special cases (see §6.1.4).

Condition I also assures that the models for causal junctions associated with laws will always be either 'deterministic' or 'pseudo-indeterministic'. That is, the sole endogenous variable (the focal-variable) is either a function of other variables in the model or else is a function of other variables in some larger model of which the first is a special case. Note, this doesn't imply that the world must be deterministic for the analysis to work, only that the most accurate *laws* are deterministic in this sense.[17] This is crucial for minimising the assumptions required to support it (specifically, it blocks one reason to assume the condition of Faithfulness; see fn.18).

The purpose of condition II is to provide the conditions according to which a the theory will exhibit the right *causal* relationships between variables. The identification works by supposing that every causal junction has at least one model (i.e. one system which models the theory) in which the focal-variable (W) 'reveals itself' in the probability distribution over the model's variables thereby establishing the causal asymmetries which exist in every model of the theory.

This identification works in part due to the equivalence between causal junctions and what we may refer to in graphical terms as a set of undirected causal paths containing exactly three variables each of which has one and the same 'unshielded collider' (Spirtes et al. 2000). A *collider* is a variable along a causal path that has the effect of both adjacent variables, and a collider is

[17]This seems consistent with the genuinely indeterministic results of quantum mechanics. After all, both Schrödinger's equation (featuring in the behaviour-predicate of Schrödinger's law; see p.86) and the Born rule are deterministic within non-relativistic quantum theory. It is only *measurements* of their system-types which exhibit indeterminism.

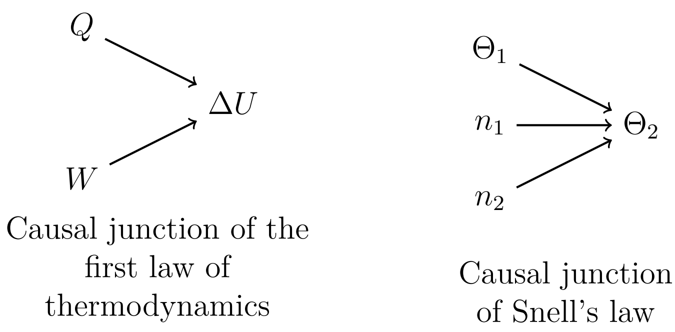

Figure 6.3

unshielded if there is no directed path in the model or theory from one of its causes to the other.

In the simplest case, a causal junction has one undirected path of exactly three variables, e.g. the causal junction associated with Ohm's law or the first law of thermodynamics (see Table 6.1). The collider on such a path is the middle variable in the sequence. But a causal junction can be made of many such causal paths (e.g. the causal junctions of Snell's law and the special force laws) so long as each path has the same unshielded collider as every other. This entails that the collider will be the effect of every other causal variable represented in the causal junction rendering the collider the focus of the junction. Figure 6.3 (p.264) comprises the graphs of the causal junctions associated with two example laws.

The probabilistic analysis relies on the idea that a particularly unique probabilistic profile associated with causal theories for causal junctions will occur in at least *some* model. But this is only plausible under certain non-trivial assumptions concerning the extent to which the *Causal Markov Condition* and *Minimality* hold in various circumstances. I will consider each in turn.[18]

[18] It is common in the probabilistic analysis of causal models to assume a further and more controversial assumption of *Faithfulness*, that all causal relations in the model are entailed by the Causal Markov Condition. This, in effect, precludes the causal influences of a variable on another cancelling along multiple causal routes Hitchcock (2016). Despite existing attempts to defend the principle (e.g. Spirtes et al. 2000, Steel 2006), it has also been harshly criticised (see Cartwright 2007, 66-68). While I'm inclined to think it can be defended in the case of lawful systems, it seems to me more impressive to show the extent

The first assumption required for the probabilistic analysis to work concerns the extent to which models satisfy the *Causal Markov Condition* (CMC). CMC has been utilised in a number of discussions of causal modelling (see, e.g., Pearl 1988, Spirtes et al. 2000, Pearl 2000) and can be defined in the following way.

> **Causal Markov Condition (CMC):** A model $\langle \mathbf{V_x}, \mathbf{E_x} \rangle$ (for any system x) satisfies the Causal Markov Condition if and only if, for every W_x in $\mathbf{V_x}$, W_x is probabilistically independent of all distinct variables which are not effects of W_x, given the direct causes of W_x in the model $\langle \mathbf{V_x}, \mathbf{E_x} \rangle$.

Two variables (continuous or discrete) are probabilistically independent if the joint probability density function of the two variables is equivalent to the product of the marginal probability density functions for each variable; otherwise they are dependent. So long as CMC holds approximately of a model then we know that any variables which are not directly causally related in the model will be probabilistically independent conditional on their direct causes.

Satisfying CMC is certainly not trivial and there are many known reasons why it may not hold for a particular model. However, in discussing potential failures, it will be important to get clear on what is actually required by the probabilistic analysis of causal junctions. Some mooted counterexamples will need to be dismissed out of hand, others may exist but must not be ubiquitous, the acceptance of others' existence depends on context.

One kind of failure of CMC was observed by Salmon (1984a) in cases where there exists an 'interactive fork' between the effects of a common cause. Consider the direction of a cue-ball and an 8-ball after a collision in a game

an analysis of causal asymmetry can be provided without it. The main reason we need not posit Faithfulness is that the causal models we are considering are pseudo-indeterministic. Under this condition, no failure of Faithfulness by a causal model can exhibit a variable which can be misidentified as the focus of a causal junction (specifically, such failures will not satisfy II(i) or II(ii)).

of pool. In some circumstances, the angles can be such that whether or not the 8-ball goes in one corner pocket depends statistically on whether or not the cue-ball will go in the neighbouring pocket. Both are distinct effects of a common cause, viz. the initial collision, but given the cause, the status of each ball's potting remains dependent on the other's.

In a case such as this, the interactive fork appears to be a consequence of the particular variables we choose to model it. After all, if we individuated the collision in terms of the precise momentum of the cue-ball, we would expect to render the effects independent once more. In this case, at least, the failure of CMC is dependent on the choice of variables in the model. In general I think it is plausible that the natural properties of the mosaic will be sufficient to provide the kind of variables which are not subject to interactive forks which fool us into attributing causation where it does not exist. Many apparent interactive forks are a consequence of poorly chosen variables which don't reflect the actual detail of a situation. But the mosaic itself does not have this problem. Famously, there still remains the problem case of quantum entanglement (see Spirtes et al. 2000, 63-64, and also Elby 1992). However, I will not presume that it is incumbent on the analysis to be provided here to analyse whatever causal relations may or may not exist in cases of quantum entanglement. For instance, I think we can agree with Spirtes *et al.* who argue that,

> the apparent failure of the Causal Markov Condition in some quantum mechanical experiments is insufficient reason to abandon it in other contexts. We do not, for comparison, abandon the use of classical physics when computing orbits simply because classical dynamics is literally false. The Causal Markov Condition is used all the time in laboratory, medical and engineering settings, where an unwanted or unexpected statistical depedency is prima facie something to be accounted for. If we give up the Condition everywhere, then a statistical dependency between treatment assignment and the value of an outcome

> variable will never require a causal explanation and the central idea of experimental design will vanish. (2000, 64)

We can, I think, remain confident that interactive forks are at most rare outside of the quantum domain and for the sake of much analysis can be assumed not to exist. Moreover, for an interactive fork to be involved in a causal model whose probability distribution allows for a false identification of a causal junction, it would have to be the case that the putative collider is dependent on multiple other variables as well as the one it is interacting with, while that interacting interacting variable must remain independent (in at least one model) of some of those other variables. Though logically possible, it seems to me that such a circumstance would be very strange, even for the quantum domain. I will assume from hereon that interactive forks can be ignored as a problem case for CMC.

A different kind of failure of CMC which must be considered can occur if a model isn't large enough. For example, if a model includes two effects of a common cause but not the cause itself, then those variables will be dependent and yet have no cause in the model to screen off their dependence. When it comes to models of the causal junctions associated with laws this may often be a possibility. For instance, the electrostatic force experienced by each of the plates in a charged capacitor is the focal-variable of a causal-junction associated with Coulomb's law were the charges of each plate are its cause-variables. However, if the two plates have been charged by the same voltaic cell, then their opposing charges will not be probabilistically independent, despite being uncausally related in the causal junction's model. (Recall that models and theories for causal junctions only include the variables in the junciton itself.)

Clearly, by including the power from that cell in the model, the probability distribution would obey CMC once more. But the assumption I wish to defend does not demand expanding the models of causal junctions to include the cause-variable's causes. It merely requires that there exist *some* instance which models the causal junction in which causally unrelated variables in the model are probabilistically independent. As with any causal junction liable to be referenced in a law, I think this is plausible in the electrostatic case. There are, for example, plenty of instances of electrostatic attraction between charged objects (e.g. charged sub-atomic particles and ions) which are independently charged. Such instances are all that are required for the first assumption to go through. So, unlike models with interactive forks, the existence of models which fail to satisfy CMC because the model isn't large enough do not need to be dismissed out of hand for the probabilistic analysis of causal junctions to go through.

The final problem case I will consider for CMC is one in which there do not appear to be *any* instances in which the direct-causes are independent. For instance, in any model of the causal junction associated with the total force law in which mass varies across token instances, we should not expect force and mass to be probabilistically independent. This is because mass has a positive effect on force in models of the causal junction associated with the Gravitational law, and wherever there is a model of the total force law's causal junction in which mass varies, there is also a model for the Gravitational law. Hence, for every model of the total force law (except the model comprising the entire universe), there is an expanded model in which the first is embedded which represents acceleration as a *shielded collider*, i.e. there is a causal influence between its cause variables in the model (viz. force and mass). As a consequence, no model of the causal junction associated with the total force law will exactly satisfy CMC, since these restricted models do not acknowledge a causal relation between mass and force despite the two failing to be probabilistically independent conditional on their causes in the model.

It is examples like this (perhaps the only example!) which leads me to include the weaker requirement of *approximate independence* between cause-variables in condition II(ii) of the probabilistic analysis for causal junctions. For despite the constant influence of mass on force, there are models of the causal junction associated with the total force law which are clearly *very close* to satisfying CMC. A nonzero charge in two bodies, for example, will typically significantly outweigh the effects of gravitational attraction in small bodies. This can be seen by comparing of the extreme difference in the constants $k_e = 8.99 \times 10^9 Nm^2C^{-2}$ in Coloumb's electrostatic law and $G = 6.67 \times 10^{-11} Nm^2kg^{-2}$ in Newton's gravitational law. It will be clear from the probability distributions over variables in models where mass and charge vary that mass and force are very nearly independent in comparison with the significant dependencies between, on the one hand, mass and acceleration and, on the other hand, force and acceleration.

The weakened requirement of approximation to CMC may seem problematic. However, it's worth pointing out that this the notion of approximation comes with a quantifiable measure. We could, if necessary, put an exact value on the threshold for approximation to CMC in order to dispel vagueness, although I doubt this would be helpful. Having said that, if we keep the boundary vague this need not entail that there are any boundary cases—the world may be kind enough to make each case clear enough. But even if there are some cases of indeterminacy in the causal structure for some systems, I take this to be entirely acceptable under the Humean world-view.

Alternatively, if the move to approximation turns out to be too problematic there are other options. One is to revise the set of assumptions which constitute the analysis such that the focal-variable is determined in 'expanded models' in which the indirect causes of the focal-variable are also taken into account in the probability distribution. If, for example, neighbouring masses, charges and distances were included in the model, we would reclaim CMC and the focal-variable would be determined once more. Another response would be to argue that the force law should not be represented as having a single force-variable, but rather a variable for each component force. Although this would only solve the problem for this specific case, there seems to me some independent justification for it given the mysteriousness of 'total force'.

But despite various alternative options, for the sake of simplicity I will stick with an analysis in terms of approximation. The weakening to approximate independence is something any hard-nosed actual frequentist should be open to anyway, since according to that interpretation failures of independence can occur simply if the proportion of correlated variables do not cleanly factor into the reference class (see Hájek 1997, 223-224 and also the discussion in §6.2.1).

Whereas the assumption that CMC holds approximately in relevant models provides the assurance that probabilistic dependencies are indicative of causal relations, the second assumption, that *Minimality* holds in all such models, provides the assurance that probabilistic independencies are indicative of *non-causal* relations. Minimality (MIN) is also widely assumed in causal modelling (see, e.g., Pearl 1988, Spirtes et al. 2000) and can be defined in the following way.

> **Minimality (MIN):** A model $\langle \mathbf{V_x}, \mathbf{E_x} \rangle$ (for any system x) satisfies the Minimality Condition if and only if no model obtained from $\langle \mathbf{V_x}, \mathbf{E_x} \rangle$ by removing any direct causal relations satisfies CMC.

In this context 'removing a direct causal relation' implies fixing one of the structural equations in the model to be such that the coefficient of one right hand variable is set to zero. Graphically, this is equivalent to removing the arrow (or 'edge') which joins that right hand variable to the left hand variable in that equation. For the deterministic and pseudo-indeterministic settings we are considering, so long as a model satisfies MIN, any two variables which are directly causally related will either be probabilistically dependent on each other, given their other direct causes, or else their relationship will be undefined (in failures of Faithfulness; see fn.18). Crucially, in neither case can the identifying features of a causal junction be

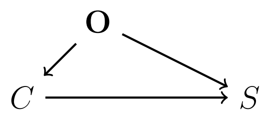

Figure 6.4

exhibited.

Like CMC, MIN is not trivial, and there are cases where it appears to restrict too much. An example exists if all the direct causes in a deterministic model of some cause-variable C for a collider S are the other cause-variables $\mathbf{O}$ for the same collider (Zhang and Spirtes 2011; see Figure 6.4, p.271). In such a case the model will not satisfy MIN because S will be independent of $\mathbf{O}$ conditional on C and independent of C conditional on $\mathbf{O}$, and the removal of the causal relation either between $\mathbf{O}$ and S and or between C and S will still satisfy CMC. However, although plausible, the failure of MIN is no threat to the probabilistic analysis of causal junctions since in none of these models (either the 'true' model or any models obtained by removing causal relations) does any variable exhibit the same probabilistic character of a causal junction which satisfies CMC. This is a result of the collider in the example being failing to being *shielded*.

A different way a model can fail to satisfy MIN is if one of the variables does not vary at all. In that case, that variable will be probabilistically independent of all other variables even if it should be counted as a cause-variable or focal-variable. Presumably there are many such models for many causal junctions, particularly when the variable is a conserved quantity (e.g. mass or charge). Nonetheless, the probabilistic characterisation of causal junctions only requires that there be *some* models which satisfy MIN, which there plausibly will always be.[19]

For deterministic and pseudo-deterministic models which satisfy CMC and

[19] So long as we can identify models by an experimental context in which components (with different mass, say) can be swapped in and out, rather than by a single permanent system of enduring objects (see fn.14).

MIN the pattern of probabilistic relationships exhibited by unshielded colliders turns out to be unique. CMC implies that the direct causes of an unshielded collider will be probabilistically independent conditional on their causes (although not when conditional on their effects!), and MIN implies that the unshielded collider will not be probabilistically independent of any of its direct causes. Hence among models which (approximatly) satisfy CMC and MIN, a model for a causal junction comprising a set of causal paths which share the same unshielded collider will be unambiguously determined by the probability distribution over it. Supposing there is one such model for every causal theory for a causal junction, the causal asymmetries in the theory itself can be determined.

Let's consider an experimental example. Imagine an experiment e in which one goes into a series of rooms one after the other and measures the time it takes for the reverberation RT_D of a repeatable complex sound wave of constant amplitude to drop a certain number of decibels D, the volume V of the room, its surface area S, and a constant of absorption a for sound which depends on the material of the surfaces in the room. It turns out that each of the variables RT_D, V, S, a can be expressed as a function of the others and, moreover, that while RT_D is dependent on every other variable, there exist approximate independences among some of the remaining variables. Specifically, S and V are independent from a_m. According to the probabilistic analysis of causal junctions, this information is sufficient to formulate a causal theory for a causal junction $\langle\{RT_D, V, S, a\}, \{R_T := f(V, S, a)\}\rangle$ in which RT_D is the focal-variable.

Like any other causal theory, this one has model-conditions. Formulating these will take into account the circumstances required for a measurement to be valid, e.g. that the sound in each case is the same, that the sound is made *inside* the room, that the room is *closed*, etc. Whatever the specifics, it is an implication of the probabilistic analysis of causal junctions that those circumstances which satisfy the model-conditions will be models of the same causal theory of which experiment e is a model. In any circumstance in which the same sound is made inside a series of closed rooms, that circumstance will be a model of the causal theory for the causal junction just described. Notice that this will be the case even if the model itself does not satisfy CMC or MIN. Only one model of the theory for a causal junction needs to have the special probabilistic features defined in the probabilistic analysis and it was a stipulation that the experiment e provided such a model.

As I have said, one model is all that is required to determine the type-level causal relations within a causal junction and all its models. Presumably this limiting situation is very often not the case and there are numerous models with the characteristic probability distribution captured in condition II of the probabilistic analysis of causal junctions. However, the fact there *need only* be one might seem problematic. To placate the concern somewhat, notice that at least one prima facie undesireable consequence does not in fact follow. For the fact that only one model need exist with the right probability distribution in order to establish the truth of a causal theory for a causal junction doesn't commit us to saying that if the model hadn't existed, the causal relations predicted by the theory wouldn't exist in any of its instances. Recall that if we accept NP (see §2.2) we will accept that laws remain laws under any counterfactual supposition consistent with them. It seems continuous with NP that if we accept a causal junctions account of laws, we should hold that the causal theories associated with RCJLs will hold under such counterfactual suppositions consistent both with them and the associated causal theories. Since it is consistent with RCJLs and the causal theories for causal junctions associated with them that they lack models which satisfy CMC and MIN, those models need not exist in order for the laws and causal junctions to still accurately capture the causal relations in other models. The only constraint I wish to assert, therefore, is that there must be at least one model in the *actual* world which satisfies these conditions in order for the probabilistic analysis of causal junctions to be possible. (For analogous remarks concerning actual frequentism, see §6.2.1, and concerning extensional analysis of dispositions, see §6.1.1.)

6.2.5 A probabilistic analysis for law-level causal junctions

So far in §6.2 I have put aside law-level asymmetry to concentrate on analysing the variable-level causal asymmetry associated with RCJLs. I just provided such an analysis by using observations drawn from Spirtes et al. (2000) about the unique probabilistic character of certain models of causal theories for those junctions. Crucially, for the discussion on law-level asymmetry, I also suggested that the conditions of assembly for system-types could be identified with model-conditions for those causal theories. One might presume this to be sufficient in order to supply an analysis also of law-level asymmetry, i.e. deeming the asymmetry between system-type and behaviour in a law to be captured by the intrinsic difference between model-conditions and the causal theory itself (recall §4.1.2). However, I think there is something more informative we can say.

First notice that the satisfaction of various conditions of assembly can be denoted as value-instances of variables. Consider a causal model for some particular electrical conductor r, an instance of the system-type of Ohm's law. We can introduce new binary variables: $C_r = 1$ if r has a conductive component, 0 otherwise; $P_r = 1$ if r has a voltage potential, 0 otherwise; $Ch_r = 1$ if r has a source of charge, 0 otherwise; $O_r^{across} = 1$ if r is organised such that the voltage potential is across the conductive component, 0 otherwise; and $O_r^{in} = 1$ if r is organised such that the source of charge is in the conductive component, 0 otherwise. The behavioural property of Ohm's law is having properties voltage, V, current I and resistance R satisfying the equation $V = IR$. We can accordingly introduce the variable 'Ohm_r', where $Ohm_r = 1$ if r instantiates the behavioural property, 0 otherwise.

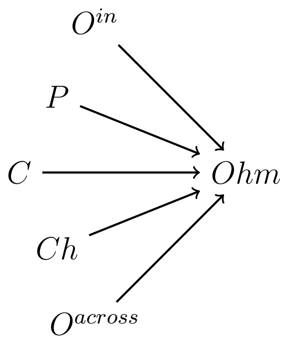

Figure 6.5

Given the characterisation of variables, the law-level causal asymmetry in r can be captured in the following structural equation.

$$Ohm_r := C_r \times P_r \times Ch_r \times O_r^{across} \times O_r^{in}$$

What we have described is the model of a causal theory $\langle\{Ohm, C, P, Ch, O^{in}, O^{across}\}, \{Ohm := C \times P \times Ch \times O^{in} \times O^{across}\}\rangle$. The causal information of the theory is represented graphically in Figure 6.5 (p.275).

It is noticeable that the causal theory is that of a causal junction, although this time at the *law-level*, with *Ohm* as the focus. We might expect, therefore, that a probabilistic analysis akin to that provided at the variable-level will apply here also. I say 'akin', since the theory for the law-level causal junction would appear not satisfy Condition I in the probabilistic analysis.[20] But it certainly appears to satisfy Condition II.

To see that this is the case observe first that in the model just considered, it is reasonable to assume that Ohm_r will be statistically dependent on the other variables in the model. After all, if any one of the conditions of assembly fails

[20] A question I have not pursued in the main text is whether there are model-conditions for causal theories for causal junctions at the law-level. If so, we might be tempted to posit further laws in which the system-type is defined by those conditions. However, I suspect, that the model-conditions will be trivial: being satisfied by any assembly so long as it is meaningful to assess whether or not it satisfies the system-type in Ohm's law or not.

to hold, the behaviour will not result. Furthermore, many of the conditions of assembly in the model will be statistically independent of each other. For example, whether a component is conductive or not is independent of whether a source of charge exists within it. Admittedly, it is plausible that not every pair among a system-type's conditions of assembly will be independent. But as long as some condition of componency is independent from some other condition, then the probability distribution on the model will sufficient to reveal Ohm_r as the focus.

In general, the probabilistic features which help determine the asymmetries in a causal junction defined in Condition II of the probabilistic analysis for causal junctions at the variable-level works just as well for causal junctions at the law-level. Indeed, it seems to be in some sense even *more* stable, since there is not the same issue of embeddability (hence why Condition I is irrelevant in the law-level case).

Drawing on the suggestion in §5.3, I take it that we will often be able to collapse causal theories for both levels into one. For example, we might think that in any instance of the system-type of Ohm's law, the existence of a source of charge in the conductive component is a direct causal influence of both the value of resistance and the current flow. Similarly, we might think that the existence of a voltage potential and its organisation across the conductive component is a causal influence of the voltage drop across the conductive component. Taking these causal relations into account and adding them to the graph representing the variable-level causal junction associated with Ohm's law gives us the graph represented in Figure 6.6 (p.277).

Although each case in which causal theories at the law-level and variable-level are collapsed into one causal theory must be considered independently, I suggested in §5.3 that it is invariably the case that causes of either junction will causally influence (directly or indirectly) the focal-variable (i.e. the focus of the variable-level causal junction). This is clear in the above graph where

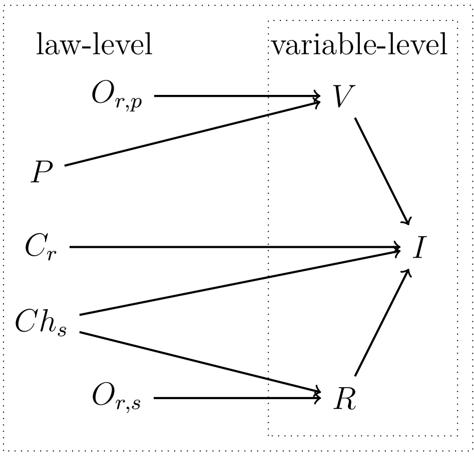

Figure 6.6

all causal paths end up at I and the general idea will be of central importance on for the analysis of token causation developed in §7. Although it is beyond the scope of this thesis to offer any general probabilistic analysis for causal theories of the sort represented graphically in Figure 6.6.

But we do not need to see the compatibility of the two levels of theory to notice what's important for the purposes of this chapter: that causal junctions at the law-level as well as the variable-level can be provided a probabilistic analysis consistent with the Humean methodology.

6.3 A *Humean* causal-junctions account of laws

Over the course of this and the previous chapter a new account of laws has emerged consistent with the Humean methodology. The account comprises two components. The first component was the conclusion of §5: a causal-junctions account of laws. This accounts for laws as any generalisation which has the form $\forall x(Sx \rightarrow Bx)$ and is either a robust causal junction law (RCJL) or else associated inferentially with an RCJL in some derivative or supportive way (see §5.4.5). Although the list of appropriate ways provided is open ended (hence 'account' rather than 'analysis'), I take it that it provides an informative enough generalisation to play a role in accounting for laws. Nonetheless, the account in the form presented in §5 took various causal features as primitive, hence the need for a second additional component of the account specifically designed for those who want to avoid such unHumean posits. In effect this means providing an analysis of RCJLs in Humean terms. This second component has been the task of this chapter.

In §6.1 a proposal for analysis of the dispositional features contributing to RCJLs robustness was provided in terms of two extensional conditional analyses StoCA and SysCA. In §6.2 a proposal for analysis of the causal asymmetrical features contributing to RCJLs association with causal junctions was provided in terms of a probabilistic analysis of certain models of those causal junctions' causal theories which satisfy conditions of CMC and MIN.

This completes the discussion on accounting for laws under Humean methodology. In the final chapter I show how it can be put to use in developing an informative analysis of token causation.

7 A nomological analysis of token causation

The causal-junctions account of laws proposed at the end of §5 builds heavily on the inferential associations of laws with type-level causal features. However, the analyses provided for those causal features in §7 do not directly supply us with an analysis of type-level causation *in general*. For instance, the extensional conditional analyses of dispositional properties provided in §6.1 gives us an analysis of dispositions which feature in laws' conditions of componency or at the law-level, in the case of embeddability. But the analysis does not extend to dispositions which do not feature in laws of which there are plausibly many. The probabilistic analysis of causal junctions provided in §6.2 also does not extend to an analysis of type-level causal asymmetries in general, since it only aims to provide an account of those asymmetries which exist *in a single causal junction*. Therefore, it presumably cannot be used directly to account for the causal relationship between smoking and contracting cancer, or between the production of greenhouse gases and climate change, which may in fact be linked by long chains of causal connection.

An account of type-level causation in general is beyond the scope of this thesis. I do, however, suspect that a plausible way one might approach such an account is via a generalisation over the token-level causal relations. If that's to work we need first an analysis of *token-level causation*, i.e. instances of type-level causal relations relating a temporally and spatially locatable cause and effect.

It is my understanding that the notion of a robust causal junction law can be put to work relatively straightforwardly in accounting for token causation with only one essential further commitment concerning event-identity. This chapter aims to show how this can be done and moreover, how such an account may improve on other analyses of token causation which do not draw on laws. The result is a defence of a nomological account of token causation.

While it is not as popular as other approaches to understand laws of nature to play the central role in the analyses of token causation, it is noticeable that laws have rarely been completely absent from popular theories. The early process-theories of causation (e.g. Aronson 1971, Fair 1979) were heavily motivated by observations about the structure of laws, and later variants (specifically Dowe 2000, 172) make use of laws to assure the sufficiency of their causal criteria. Among views of causation which analyse causation in terms of counterfactual or probabilistic dependency laws have typically played a role in making sense of the dependency relations themselves (e.g. Jackson 1977, Lewis 1986b, Maudlin 2007 for counterfactuals, Hempel 1965, Lewis 1994, Hall 2004c for probabilities). Moreover, laws have also crept in into the analysans more explicitly in certain cases (see, e.g., Paul 2000, Schaffer 2001, Hall 2007).

Still others believe that a workable set of (possibly necessary and sufficient) criteria for causation might be provided with very few analysanda *other than* laws (e.g. Armstrong and Heathcote 1991, Armstrong 2004a, Maudlin 2004). One of the most promising of the 'purely' law-based accounts is surely Maudlin's, and yet he admits that 'I do not think that there is any uniform way that laws enter into the truth conditions for causal claims' (2004, 430). I disagree with his scepticism and my purpose will be to describe something approaching a 'uniform way' drawing heavily on the thesis so far.

In §7.1 I provide a recap of the key points in the foregoing chapters which will contribute to the subsequent analysis of token causation. In §7.2 I develop the analysis by first proposing an analysis of 'intra-system causation', the direct and indirect token causation which takes place within single instances of laws' system-types, and then proposing an analysis of 'inter-system causation', the token causation which takes place across multiple instances of system-types. The discussion will lead me to conjecture on event-identity and draw attention to the contrasting features of the analysis with typical entailment-centred nomological accounts. §7.3 presents an extended example of how the analysis applies to causal explanation using the Space Shuttle Challenger disaster as a case-study.

I then move on to motivate the analysis provided by comparing its successes in different ways with other approaches to token causation common in contemporary literature. In §7.4 I show how the analysis addresses classic problem cases of preemption and negative causation which have formed the foundation of many complaints with dependency and physical connection views of token causation, respectively. In §7.5 I show how the analysis addresses an 'accommodation criterion' concerning how token causes and their effects are related in time which dependency views and powers-based views seem to struggle with. §7.6 concludes.

7.1 Recap

In order to develop the analysis of token causation I will be drawing heavily on a number of claims from the foregoing chapters which I will repeat briefly here, making use of the example of Sabine's equation. This is also a good pace to remind oneself of the key conclusion of the thesis so far.

In §3.1 I argued that laws should not be considered identical to the equations typically employed to stand for them. Instead, laws are conditionals of the form

$$\forall x(Sx \rightarrow Bx)$$

where 'S' is replaced by a system-predicate which denotes a system-type and

'B' is replaced by a behaviour-predicate which denotes a type of behaviour (typically described by a formula). For example, consider Sabine's equation,

$$RT_{60} = \frac{24V \ln 10}{Sac_{20}}$$

which relates the reverberation time RT_{60} of a sound in a room to decay 60 decibels with the speed of sound c_{20} in that room to the volume V of the room, the surface area S of the room and the average absorption coefficient a of the room's surfaces. According to the argument I provided in §3.1, the equation itself is not a law. But we might nonetheless posit a law—'Sabine's law'—in which the satisfaction of the formula by some assembly of objects is conditioned on the assembly satisfying the right system-type—what we might call a 'Sabine Room'.

In §3.2.1, I suggested that we understand system-types in general to be individuated by 'conditions of assembly', further divisible into conditions of componency and organisation. Arguably, the conditions of assembly for a Sabine Room are satisfied by all and only assemblies comprising a closed room in which a constant sound-wave is made (see the example of experiment e in §6.2.4).

I then discussed, in §3.3, equations which do not accurately describe the behaviour of any system-type. I suggested that the corresponding law can be rendered true by modifying the behaviour-predicate to require that the contained equation be a special case of a more accurate equation in a law whose system-type has the modified laws' system-type as a part. This behaviour-modification aims to save the universality of problematic idealising laws by capitalising on the idea that laws describe the behaviour of embedded systems. Notably, the solution worked without having to restrict the class of systems to which the idealising laws describe thereby making sense of the full breadth of cases they are applied to in practice.

With our example, we know that Sabine's equation doesn't accurately apply to any system perfectly accurately. The presence of objects in the room and moving surfaces can complicate the relationship between its variables. Nonetheless, we might suspect there to be a more accurate equation which takes into account these further conditions and which will have Sabine's equation as a special case. The law with which this more accurate equation is associated will have a system-type whose conditions of assembly have those for a Sabine Room as a part. Consequently, the truth of Sabine's law can be saved if it is rendered as follows.

> **Sabine's law:** For all x, if x is a Sabine Room, then there is a system-type S' and behaviour-type B' such that,
>
> 1. $\forall x(S'x \rightarrow B'x)$ is a law, and,
> 2. instances of S' have an instance of a Sabine Room as a part, and,
> 3. $\forall x(B'x$ if and only if x has variables RT_{60}, c_{20}, V, S, a related by the formula K' which has as a special case the formula,
>
> $$RT_{60} = \frac{24V \ln 10}{Sac_{20}})$$

In the analysis that follows I will make explicit use of all of these general observations from §3. The first step is to provide an analysis for 'intra-system causation' (in §7.2.1 and §7.2.2), the token causation among the behavioural variables (e.g. RT_{60}, c_{20}) and conditions of assembly of laws' system-types (e.g. being a closed room with a repeatable soundwave). Here the behaviour-modification will serve to show how laws can still be used to analyse token causal relations even in scenarios in which they are not accurate. Hence, we should expect to be able to use Sabine's law for causal analysis despite its inaccuracy.

In §4.1 I discussed the plausibility of a typical way in which laws and causation have been thought to be inferentially connected, viz. the *causation-mirroring* *conception of laws.* This conception interprets laws as having a logical form whose instances can, like causation, concern distinct individuals and times. I defended the logical form proposed in §3 against this conception thereby showing that the way in which laws analyse causation (if they do at all) cannot be simply one of 'covering' causal relationships. For example, Sabine's equation cannot be applied in such a way as to *entail* anything about some system in the future based on the values of some potentially distinct system at a time. Rather, the equation is known to apply to systems *at the very time that they themselves are Sabine Rooms.*

In what follows I will uphold this prohibition of 'covering'-type analyses of causal relations. The relationship will be instead much more similar to one of physical connection in which the existence of instantiated system-types provides the means of connection. However, I will also show, in §7.4, why the account improves on physical connection accounts by being able to accommodate a certain degree of negative causation. In §4.2 I used the rejection of the causation-mirroring conception to reject the best system account of laws, which I argued was implicitly committed to the conception. A more promising account, I argued, would draw on a more accurate causal conception.

To this end, in §5.2 I drew on four causal features which I argued are inferentially connected with many laws: component-level and law-level dispositionality, and variable-level and law-level causal asymmetry. I then argued, in §5.3, that there is a class of laws—the *robust causal junction laws* (RCJLs)—which exhibit all four features in a particularly notable and prevalent way. RCJLs are considered to be robust since the conditions of assembly for their system-types contain dispositional requirements on their components and instances of their system-types' are embeddable with respect to the laws' behavioural formulae in the instances of system-types of more accurate laws. Perhaps more relevant to the analysis to follow, the causal asymmetries in these laws at both variable and law-levels combine to form *causal junctions*, in which a single effect or 'focus' has multiple causes. Moreover I argued that the focus of the variable-level causal-junction—the 'focal-variable'—can be considered the effect both of the other variables in the variable-level junction and the conditions of assembly in the law-level junction.

Returning to our current example, Sabine's law appears to be just such an RCJL. The law is rendered robust partly due to the fact that a condition of componency for being a Sabine Room is that the room's surface is closed with a material which reflects a non-zero quantity of sound energy. It is also rendered robust by the fact that instances of its system-type are embeddable with respect to the Sabine equation in more complex scenarios, such as those in which there are objects present and moving surfaces. Brief consideration of the kinds of explanatory relationships and interventions available should also make it clear that RT_{60} is a focal-variable, causally explained by reference to the other variables and the conditions of assembly which individuate the class of Sabine Rooms.

In what follows I will make essential use of the idea that token causal relationships can exist whenever the system-type of an RCJL is instantiated. These causal relationships exist between the token instances of any cause-variable from either the law-level or variable-level causal junction associated with the RCJL and the focal-variable. The central intuition of the account is that any token effect can be identified with some focal-variable of the variable-level causal junction associated with some RCJL taking a specific value. For ease of reference I will refer to the focal-variables of the causal junction associated with a law L of the form $\forall x(Sx \rightarrow Bx)$ as the 'focal-variable of B', 'the focal-variable of L', 'the focal-variable of a law', etc. Hence, the focal-variable of Sabine's law is RT_60.

In §6 I argued that we can provide an analysis of the type-level causal features of RCJLs appropriate to the Humean methodology advocated in §2.

Obviously, if the analysis of token causation in terms of RCJLs is to be genuinely non-causal it is vital that these type-level causal features admit of *some* non-causal analysis themselves. I will assume that the Humean analysis provided in §6 lends substantial support to this being the case. More specifically, if the analysis of RCJLs' type-level causal features is genuinely Humean then the analysis of token-level causation which makes use of such laws will be Humean also. Consequently, in combination with §6, the result of this chapter will be a Humean account of token causation.

Nonetheless, in what follows I will not make any essential reference to the Humean analysis of RCJLs' dispositionality or causal asymmetry. All that is needed to develop the analysis of token causation is the notion that an RCJL has a focal-variable which is the effect of all other variables in the behavioural formula and all the conditions of assembly for the laws' system-type. I will often write as though it is clear which is the focal-variable of a law assuming that some argument could be provided for this in the manner of the foregoing chapters (e.g. either by drawing on knowledge of what can be manipulated via certain interventions, or thinking more specifically about the probabilistic dependencies). While differences of opinion on which is the focal-variable may jeopardise the success of my examples, it would only jeopardise the overall approach if the very concept of a focal-variable were in question. I assume enough has been done in §5 and §6 to block that concern.

7.2 The Proposal

My proposal for causal analysis of token causation is to make use of the system-relativity of focal-variables. A particular instance of the value of a variable may be the instance of the focal-variable for one system, but the instance of a *cause*-variable in another. It is this feature of causal events which allows for token causation to be analysed via chaining together instances of laws' system-types. These chained causal relationships I will call 'inter-system causation', and my proposal is that all token causation is inter-system causation. In order to show why this might be plausible, I need to start first with the links in the chains of inter-system cauation, viz. 'intra-system causation'.

First we need to establish some terminology and notation. I denote causal events (whether causes or effects) with logical constants $a, b, c, \ldots$, nominalisations (e.g. *the Beatles' last gig, the sinking of the Andrea Doria*) and also as 'value-instances' of the form '$\Theta^t_y \Leftarrow \theta$', to be read 'the instantiation of the value θ at t of the variable-property Θ by object y'. For the sake of simplicity, I will for the most part treat the times of such events as *instants* of zero duration. However, I work under the assumption that any event a which has a temporal part which causes a temporal part of another event b is a cause of b. Hence, by providing an analysis for causation between instants, we get (via the assumption) an analysis for enduring events also (see §7.3 for an example of this). When I want to denote an event comprising a summation of value-instances between two specific times, e.g. $t1$ and $t9$, I will use the notation $\Sigma^9_{t=1}\Theta^t_y \Leftarrow \Theta(t)$, to be read 'the summation of value-instances from $t1$ to $t9$ of the variable-property Θ by object y'. Events will constitute the sole relata for all token-level causation analysed in what follows.[1] Their method of denotation should not, however, be taken to imply a commitment to any criterion of identity for events; this issue is discussed below in section §7.2.5.

I will also make use of the notion of a 'system-event'. This can be denoted in a similar way to other events, but here, the variable which takes the place of Θ will indicate whether or not a particular system-type is instantiated or not, y by the name of an assembly of objects and θ by a 1 to indicate that the system-type is instantiated (0 would indicate that the assembly does not instantiate the system-type). Also, the time-variable will be prefixed by a 'Δ' to signify that system-events endure for some period of time denoted by a single variable (e.g. 't') or by a range (e.g. '$t1-t9$'). So, a system-event comprising

[1] The account will, therefore, be inconsistent with a fact ontology (e.g. Mellor 1995, 2004) or a states-of-affairs ontology (e.g. Armstrong 1997, 2004a) of causal relata.

the instantiation of a system-type S which features in the antecedent of at least one law by an assembly of objects a for some duration Δt will be denoted '$S_a^{\Delta t}\Leftarrow 1$'. Assuming the respective law(s) are true, then a system-event allows us to predict the satisfaction of the laws' behaviour type by that same assembly. I will also make use of the idea that system-events may have their own focal-variables when the instantiated system-type is that of an RCJL.

Finally, I will say that two events p and q, where $p = \Theta_y^{t1}\Leftarrow\theta$ and $q = \Phi_x^{t2}\Leftarrow\phi$, are *system-paired in* $S_a^{\Delta t}\Leftarrow 1$ if and only if, x and y are both part of a, $t1$ and $t2$ fall entirely within Δt, and for some law $\forall x(Sx \to Bx)$, for either event, either,

1. it is the satisfaction of a condition of assembly for S,[2] or,
2. its variable (Θ or Φ) features in the behavioural formula in B.

So, for example, a system instantiating the system-type for Ohm's law can be expected to incorporate system-paired events between any of the following events (given appropriate times of occurrence): the event of the conductive component being conductive, the event of the source of charge being a source of charge, the event of the voltage potential being applied across the component's terminals, the event of the source of charge being in the electrical component, the event of the electrical component's resisting the charge to the a value of R, the event of the voltage potential being of a value of V and the event of the current-flow being a value of I.

We are now ready to start building up an account of intra-system causation. The basic idea behind intra-system causation is simple: all instances of the type-level causation associated with RCJLs are token causal relations. But when it comes to considering token-level causation there are a number of delicate issues concerning the temporal occurrence of the relata, which need to be dealt with carefully. In order to do this I divide intra-system causation

[2]More specifically, the disjunct should require that the event is a condition of assembly for S which actually constitutes *part of the reason* a is S.

into two varieties, direct and inertial, treating each in turn before characterising intra-system causation *simpliciter*.

7.2.1 Direct intra-system causation

In §5.3 and §6.2.5 it was suggested that in general we can assert that type-level causes of the focal-variable in a system can be either the cause-variables in the variable-level junction or the causes in the law-level junction, also describable as the value-instance of a (binary) variable (see §6.2.5). With this in mind, the most straightforward token-level cases are what I call 'direct intra-system causation', defined as follows.

Direct Intra-System Causation

c is a **direct intra-system cause** of e, where $c = \Theta_y^{t1}\Leftarrow\theta$ and $e = \Phi_x^{t2}\Leftarrow\phi$ and $c \neq e$, if and only if there is some robust causal junction law $\forall x(Sx \rightarrow Bx)$, such that,

1. Φ is the focal-variable in B,
2. c and e are system-paired in $S_a^{\Delta t}\Leftarrow 1$ (for some assembly a), and,
3. $t1 = t2$

To get an idea of how direct intra-system causation has been defined, consider Figure 7.1 (p.290) of an electrical circuit with a single-cell power source and resistor.

Say we want to find the direct intra-system causes of the event $e = I_r^t\Leftarrow i$, the circuit r having a particular current i through it at t. First, we note that there is a system-event in which the system-type which features in the antecedent of Ohm's law is instantiated by assembly r. In its unmodified form, and making the conditions of assembly explicit, Ohm's law can be expressed as follows:

> **Ohm's law:** For all x, if x is a system S comprising a conductive component, a source of charge in the component and a voltage

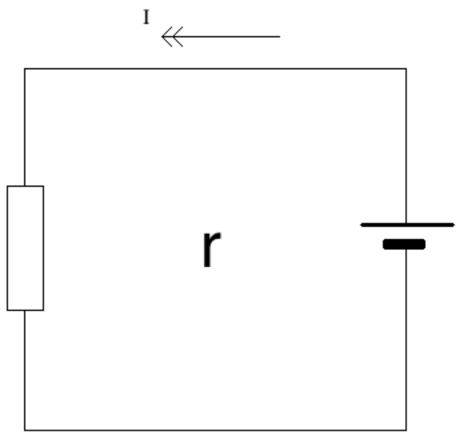

Figure 7.1: Cell, Resistor & Switch in Series

> potential across it, then the resistance of the conductive component R, the voltage potential V and flow of charge through the resistor I are related by the formula,
>
> $$V = IR.$$

According to the above definition, the direct intra-system causes of e will include the simultaneous value-instances of variables in the same behaviour-predicate and the simultaneous satisfaction of conditions of assembly by r appears. Since I is the focal-variable of Ohm's law, these will include the simultaneous value-instances of voltage potential V across r (the terminals of the cell), the resistance R of r, the conductivity of r, the existence and location of a voltage potential across r, and the existence and location of a source of charge in r (i.e. all the instantaneous instances of relations represented in Figure 6.6, p.277). I take these results to be intuitively correct.

Notice that nothing in the definition of direct intra-system causation requires that the equation in the law's behaviour-predicate accurately *predicts* the effect on the basis of its causes. Even at this stage, the nomological analysis does not require that the law be used to entail the effect given some conditions (thouh this may in fact be possible in the circumstances). Hence, if we wanted, we could employ the behaviour-modified variant (see §3.3.3) of Ohm's law in the case above and still get the same result.

Nonetheless, the current definition can appear deficient in a couple of ways. First, from a contemporary perspective, we will expect the most precise laws of physics to be relativistic. This suggests that the notion of absolute simultaneity in causal influence should be abandoned in favour of simultaneity relative to a frame of reference. Moreover, the causal influence between any two objects in an assembly may also take time due to other factors (e.g. light/signal-propagation, damping) and so not be simultaneous from within *any* reference frame. Indeed, we might take contemporary physics to imply that the only truly direct intra-system causal influence will be that which occurs between co-located objects, e.g. a charge in a field, or two variables of the same object, e.g. force and acceleration (see the discussion in §5.2.3).

A more careful analysis would, perhaps, have to assess the permitted time for causal influence between cause-variables, conditions of assembly and the focal-variable by taking into account the distances between components and the speed of lower-level processes which constitute the causal behaviour of the system. In what follows I will largely ignore this deficiency of the current analysis in this respect and assume that variables in laws are simultaneously correlated except where a law says explicitly otherwise.

But while we might *permit* causal influence mediated instantaneously, it is a further issue that direct intra-system causation is *restricted* to characterising only instantaneous causal relations. I don't think that intra-system causation should be restricted in this way. At least *some* intra-system causes should be allowed to *precede* the time at which causal influence is mediated to the focal-variable. Otherwise we risk not being able to make sense of any causation taking time at all! §7.2.2 discusses how we might go about opening up the definition of intra-system causation in general to accommodate this in a plausible way.

7.2.2 Inertial intra-system causation

An easy way to incorporate non-instantaneous causal influence within the current definition of intra-system causation might seem simply to allow that the times $t1$ and $t2$ of the respective events c and e be such that $t1 \leq t2$ (rather than $t1 = t2$). But here we have to be careful. For it seems to me that only limited variables are plausible earlier intra-system causes of the values of a system-event's focal-variable.

Some value-instances of laws' focal-variables seem obviously to be caused by past value-instances of their intra-system cause-variables. For instance, within a single system-event of the system-type in Malthus's growth law, past size of population is surely a cause of current size of population. But other value-instances of laws' focal-variables do not seem so clearly to be caused by past value-instances of their system-events' cause-variables. For instance, within a single system-event of the system-type in the law of pendulum motion, past lengths of a pendulum and past values of gravitational field do not straightforwardly seem to bear causal relevance to the value of the pendulum's period at a time. And within a single system-event of the system-type in Snell's law, past angles of incidence and velocity do not straightforwardly seem to be relevant to the refracted velocity at a time.

One could dispute the intuition in these latter examples. After all, if one intervened on a pendulum's length or the gravitational field, one would expect the period to be forever affected unless some further careful intervention was made to put things back to how they were before. Similarly, if one intervened on the incidence angle of a ray of light, one would expect the refracted velocity to be affected from then on pending any similarly careful rectifying intervention. But when we consider what has been argued for over the foregoing chapters, I think we will be less inclined to think that the causal relations these dependencies correspond with are licensed by the the pendulum law and Snell's law themselves.

If one accepts the causal-junctions conception of laws one will accept that many laws, including the classical pendulum law and Snell's law, are each associated inferentially with a different causal junction. In §6.2 I tried both to lend support to this intuition and analyse the causal asymmetries in these junctions in terms of probabilistic dependencies and independencies among the variables. Something that went implicit in the analysis was that the dependency that focal-variables have on their cause-variables in the junction was relative to the particular time at which the values of variables were taken. Putting aside relativistic effects, low periods of oscillation of pendulums at a time is probabilistically dependent on high gravitational field strength and (independently) low length *at that very time*. High velocities of refracted rays at a time are dependent on high velocities of the corresponding incidence rays *at that very time*. Insofar as the probabilistic analysis supports the causal-junctions conception, it supports only instantaneous causal relationships between variables (again, ignoring relativistic effects).

Moreover, the arguments for the logical form of laws presented in §3 and against the causation-mirroring conception of laws in §4.1 lent credence to the idea that laws condition behavioural properties on the simultaneous satisfaction of certain conditions of assembly. Hence, the logical structure of laws themselves also provides no support for non-instantaneous intra-system causation.

If we are to take seriously these points, it suggests that we should make sense of non-instantaneous causation within instances of systems in a rather more careful way. Although I take it to be an empirical matter, it seems to me that only a certain class of laws are legitimate for licensing inferences about non-simultaneous intra-system causation. Hempel agreed with this restriction suggesting that 'causal explanation by reference to antecedent events clearly presupposes laws of succession,' where 'these laws concern temporal changes in a system' (1965, 352).[3] As I understand it, a law of succession is any law

[3] Indeed, Hempel thought that *all* causation required explanation in terms of laws of succession, citing explanations of the contrasting *laws of coexistence* such as the pendulum law, Ohm's law and the gas laws as supporting non-causal explanation. Hempel thought,

which expresses a variable as a function of some other variables with respect to time.[4] Continuity laws are an example. In integral form these laws relate the flux $\mathbf{j}_\zeta$ of some variable ζ across the boundary of a closed surface S and the change over time of the quantity of ζ within the volume enclosed by the surface to the net rate Σ_ζ at which ζ is being generated or destroyed within the volume. Mathematically, this relationship can be presented as follows.

$$\frac{d\zeta}{dt} + \oiint_S \mathbf{j}_\zeta \cdot d\mathbf{S} = \Sigma_\zeta$$

Arguably, continuity equations have a structural interpretation in which the rate of change of their respective variable ζ over time (i.e. $d\zeta/dt$) is the focal-variable of a causal junction in which the flux $\mathbf{j}_\zeta$ and net rate of generation Σ_ζ are cause-variables. For instance, a continuity equation can represent how the rate of change of population ζ within a region ($d\zeta/dt$) is influenced by those entering and leaving the region ($\oiint_S \mathbf{j}_\zeta \cdot d\mathbf{S}$) as well as the birth-rates (a source in Σ_ζ) and death-rates (a sink in Σ_ζ) in that region. This would make the value-instance of a rate of change of population the direct intra-system effect of simultaneous value-instances of all the variable-properties in the equation. But what makes continuity equations seem relevant to *non*-instantaneous causation in systems is that they constrain the amount the quantity ζ can change over time. That is, the equation shows how variables like ζ have a certain degree of *inertia*. It is *this*, I take it, which makes it plausible that earlier instances of the variables in the relevant equation are causes of their later variables (so long as the system is maintained).

for example, that 'one surely would not want to say that the pendulum's having a period of two seconds was *caused* by the fact that it had a length of 100 centimeters.' I hope I have done enough in the foregoing chapters to undermine *that* general view. But in the case of causal explanation by reference to 'antecedent events' (i.e. earlier in time) Hempel seems to have been right.

[4] Therefore these laws are not the same as generalisations concerning temporally successive property instances mentioned in §4.2.1. As I understand them, Hempel's laws of succession are genuine dynamical laws whereas those generalisations considered earlier concern an unfolding sequence of events and are not reasonably considered to be laws.

In general, laws of succession help causally connect value-instances of certain variables at a time with value-instances of the same variable at a later time. They do this by describing the numerical relationships of a causal junction in which the particular variable's time-derivative is a focal-variable. The time-derivative of population is a focal-variable in the population-specific interpretation of the above continuity equation. Hence, in the system described, I suggest population itself is caused by earlier values of population.

If we are to accommodate this non-instantaneous causation within single system-events (so that it is an instance of intra-system causation) I suggest we introduce the idea of 'inertial intra-system causation', which links causally variables whose time-derivatives are focal-variables in laws of succession. For this I notate for the time-derivative of any variable β the same variable with a dot above it, i.e. $\dot{\beta}$; correspondingly, for some event f, where $f = \Upsilon^t_z \Leftarrow \upsilon$, the event $\dot{f}$ is the value-instance of the time-derivative of Υ at t by z, i.e. $\dot{f} = \dot{\Upsilon}^t_z \Leftarrow \nu$ for some value ν. The precise definition can then be given as follows.

Inertial Intra-system Causation

c is an **inertial intra-system cause** of e, where $c = \Theta^{t1}_y \Leftarrow \theta$ and $e = \Phi^{t2}_x \Leftarrow \phi$ and $c \neq e$, if there is some robust causal junction law $\forall x(Sx \rightarrow Bx)$, such that,

1. $\dot{\Phi}$ is the focal-variable of B,
2. $\dot{c}$ and $\dot{e}$ are system-paired in $S^{\Delta t}_a \Leftarrow 1$ (for some assembly a),
3. $\Phi = \Theta$, and,
4. $t1 < t2$.

This allows us to say, for example, that past population is a cause of current population in a system since the time-derivative of population is the focal-variable in a law of succession whose system-type is instantiated by the system from the time of that past population to the present. It also allows us to say the same for any variable whose time-derivative has direct intra-system causes, e.g. the focal-variable of any continuity equation, including mass, energy and momentum.

7.2.3 Intra-system causation *simpliciter*

Figure 7.2 (p.296) represents diagramatically two possible arrangements of direct and inertial intra-system relationships between the focal-variable F, some cause-variable N in the variable-level causal junction with F and some condition of assembly C for the instantiated system-type (other cause-variables are omitted). Events in both diagrams are represented by black dots and intra-system causation is revealed by the grey arrows.

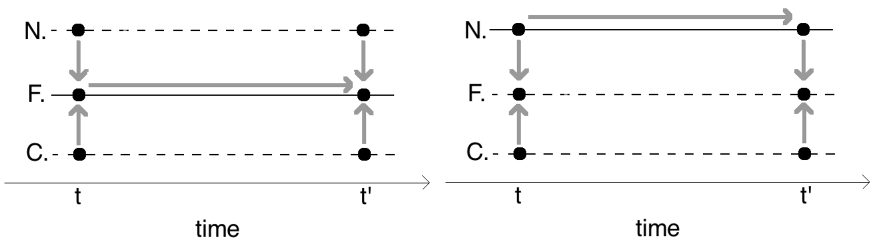

Figure 7.2: intra-causal route

In the left diagram, the time-derivative of F is also a focal-variable in a law of succession. F might, for example, be potential energy, N height in a gravitational field and C the component's having mass. Since F is the focal variable, its value-instances are directly intra-system caused by the instantaneous value instances of N and C. But since energy is also a conserved quantity, its time-derivative is the focal-variable of a law of succession, viz. the law of conservation of energy. This means it's value-instances will be inertially intra-system caused also by its earlier value-instances.

In right diagram, the time-derivative of N is a focal-variable in a law of succession. N might, for example, be population, F the quantity of social networks of some kind and C the habitability of the environment. Since F is the focal variable, its value-instances are directly intra-system caused by the instantaneous value instances of N and C. But since the time-derivative of population is a focal-variable in a law of succession, viz. growth laws, past value-instances of population will be an intra-system cause of later value-instances of population.

The diagrams might seem to imply that past events and contemporary events can *overdetermine* value-instances of focal-variables. Is this right and if so is it a problem? Here it's worth recalling that part of the campaign of §4.1 was to undermine the relevance to causation of the idea that effects are determined in any *logical* sense by their earlier causes by law. After all, there is simply no way to *determine* what the energy value of some macroscopic object is given its energy at a previous moment. Indeed, as far as I can see, no amount of additional premises about the way the world is restricted to that past moment will do so either. So whatever kind of overdetermination relation one might worry about here, it is not one of *logical* determination. But then in what sense is there a causal overdetermination worry remaining? Perhaps it is some kind of 'nomological overdetermination' which is causing the worry. But while effects can be connected via law-abiding systems both to past and simultaneous causes, the removal of logical relationships makes it harder to say why we should think direct and inertial intra-system causes are independently sufficient to bring about the effect—which I take to be the crux of the worries about systematic overdetermination.

Insofar as we have any kind of overdetermination in cases of intra-system causation, it is rather different from those usually worried about. Direct intra-system causation is *always* instantaneous (or at least as fast as causation can be mediated between variables) and exists between *distinct causal variables*. In contrast, inertial causation exists between property-instances which are causes of their own later property-instances and is *never* instantaneous. The two kinds of causation connect the cause with the effect effectively in different dimensions (on the one hand *laws' behavioural formulae*, on the other hand, *time*). From hereon, I assume that the two forms of intra-system capture the full breadth of token causal relations within a system-event without being subject to any problems as a result of overdetermination. (I will consider worries about the disjunctive nature of intra-system causation in §7.2.6.)

A final theoretically optional by exegetically useful step is to provide a definition for intra-system causation *simpliciter*. Plausibly, causal relations exist between any pair along a path comprising either direct or inertial intra-system causal relations. e.g. between any value-instances along any continuous path comprising grey arrows in the above diagrams. This suggests the following complete definition of intra-system causation.

Intra-system causation *simpliciter*

c is an **intra-system cause** of e, where $c = \Theta_y^{t1}{\Leftarrow}\theta$ and $e = \Phi_x^{t2}{\Leftarrow}\phi$ and $c \neq e$, if and only if there is some robust causal junction law $\forall x(Sx \rightarrow Bx)$, such that,

1. c and e are system-paired in $S_a^{\Delta t}{\Leftarrow}1$ (for some assembly a), and,
2. c and e are linked by a chain of direct or inertial intra-system causal links.

For example, return to the cases of the pendulum law and Snell's law considered in §7.2.2. The reason we are inclined to think that an earlier change in length of pendulum would be a cause of current period is not, according to the analysis so far provided, simply due to the fact that length and period may be system-paired. Rather, it is because they are system-paired *and* length is associated with certain further continuity laws (e.g. of mass or momentum) so that past value-instances of length is an inertial intra-system cause of later value-instances of length which in turn is a direct intra-system cause of period (an example of the right hand diagram in Figure 7.2, where N=length and F=period). Similarly, the reason we are inclined to think that an earlier change in angle of incidence of a refracted ray would be a cause of later velocity of the refracted ray is because a change in angle of incidence is a direct intra-system cause of the refracted ray's velocity (according to Snell's law) which is an inertial intra-system cause of later values of velocity due to the continuity of the photon's moment (an example of the left hand diagram in Figure 7.2, where N=angle of incidence and F=velocity). So we can reclaim our initial intuitions by witnessing how intra-system causation *simpliciter* zig-zags its way between value-instances and across time via specific direct and inertial intra-system causal relations.

In general, I take the definition of intra-system causation to capture all of the causal relationships which go on within a single system-event. The next step is to analyse token causation between events which aren't even system-paired in any system-event.

7.2.4 Inter-system causation

Not all causation is mediated by a single system. The extension to causation in other systems is fairly straightforward and comes again in both direct and inertial varieties. Regarding the former, systems exhibit a degree of nesting and overlapping, and if we are happy to put to one side concerns with transitivity, we may assume that the direct intra-system causes of a variable within some system which is both the focal-variable in that system *and* a cause-variable in a second system will be direct causes of the focal-variable in that second system.

So, for example, the resistor in our above example may count as an instance of Pouillet's law (expressed in its unmodified form).

> **Pouillet's law:** For all x, if x is a system S comprising a solid object with continuous cross-sectional area A, then its resistance R, the length l and resistivity ρ are related by the formula,

$$R = \rho \frac{l}{A}.$$

I take it that the focal-variable of Pouillet's law is R. Hence value-instances of R will be directly intra-system caused by simultaneous value-instances of l, A, ρ and by the simultaneous satisfaction of conditions of assembly *being a solid object* and *having a continuous cross-sectional area*. But we have seen that R is also a cause-variable in the variable-level causal junction associated with Ohm's law. Therefore, value-instances of R can play the role of a 'system-link' between two system-events (in this case, where one is nested in the other), viz. the instantiation of the system-type in Pouillet's law and the instantiation of the system-type in Ohm's law. Consequently, when considering the effect in our earlier example $e = I_r^t \Leftarrow i$, we might plausibly infer that simultaneous value-instances of l, A, ρ and the fact that r is a solid object with continuous cross-sectional area to also be among its causes—more specifically, its 'inter-system causes'.

Regarding the inertial case, I think we should look again at quantities whose time-derivatives are focal-variables in laws of succession, such as conserved quantities. Such quantities allow for causation to be mediated across time within systems but also between them. For instance, a projectile's momentum can be influenced by multiple boosting and retarding variables if it successively moves through a number of systems. This happens when the momentum of a snooker ball is affected by the momentum of the cue in a collision system, then by the friction of the baize in frictional system, then by the momenta of other balls, cushions and pockets in further collision systems, etc. The system-links in these cases are value-instances of the snooker ball's momentum which appear in multiple overlapping system-events. Again, putting worries about transitivity aside, I take this to suggest that the intra-system causes of a variable whose time-derivative is the focal-variable of a law of succession are also the causes of that conserved quantity at later times (even when the system-types in which those intra-system causes existed have gone out of existence).

Incorporating these ideas we get the following definition for 'inter-system causation'.

<u>**Inter-system causation**</u>

Inter-system causation is the transitive closure of intra-system causation.

This makes inter-system causation effectively equivalent to intra-system causation without the requirement that cause and effect be system-paired (this is why I said earlier that intra-system causation *simpliciter* is 'theoretically optional'). Since intra-system causation is a limiting case of inter-system causation, we can say that all intra-system causation is inter-system causation.

Given the above example, it might not be immediately obvious why a distinction between inter-system and intra-system causation needs to be maintained. For we might imagine that in the example we can take the perspective of there being one single system which incorporates all the variables in Pouillet's law and Ohm's law, thereby reducing all the instances of inter-system to cases of intra-system.

In this example, I am open to such a move. Though the laws are kept apart in practice, positing a law which combines the two does not appear to conflict with scientific understanding. However, I think it would be wrong to assume this is always possible. In non-simultaneous inter-system causation, the system-links can join systems which do not endure for the same times, so that an earlier inter-system cause might never be construed as an intra-system cause (recall that for intra-system causation, the duration of the system-event must include both cause and effect). Moreover, there appear to be cases of system-linking which are not captured by system-events formed from any mere combination of laws. This will become clearer with the discussion of event-identity.

7.2.5 Type and token event-identification

The procedure for chaining systems together to form inter-system causal relations is fairly straightforward in the case when the system-link is the value-instance of a variable which features in the behaviour-predicate of both systems (as the specific value for R is in the previous example). But this need not always be the case. Sometimes a single event must be identified as the value-instances of two *distinct* variables. Here I discuss a way this can happen in which such identifications work for all instances of a type of value-instance and another way in which such identifications only work given the contingent token features of the situation.

Some non-robust causal junction laws (non-RCJLs) entail the value of one variable on the basis of another. A very clear example of this is the Planck-Einstein law (see §5.4.3), which allows us to infer from the energy of a photon what its frequency will be and vice versa. Arguably, there is no explanatory asymmetry between these two variables. Indeed, it seems that rather than describing a relationship between two distinct event-types, the Planck-Einstein law gives us a way to *identify* the same event in two different ways: a photon's having a particular energy at a time *just is* its having a certain frequency at that time. Hence, it is plausible that a system-link can be established by using the system-event of an instantiation of the Planck-Einstein law itself to mediate the connection. In such cases, there will be inter-system causation between two further distinct system-events in which the focal-variable of one is the energy (frequency) of a photon and the cause-variable of another is frequency (energy) of a photon.

As with the example of inter-system causation in §7.2.4, system-links identified under different descriptions related by law are *type*-identifications of events. That is because the law permits an identification of the event as an instance of either variable (e.g. frequency or energy) at a time in all cases where the system-type has an instances. Non-RCJLs can, therefore, be used to support type-identifications of events for system-links in the chains of inter-system causation.

Perhaps there are many such system-links. But my understanding is that they are not the only way in which inter-system causation is thought about (if implicitly) in practice. For there seem to be cases of system-linking in *token* cases where no law can be found which helps translate from characterisation in terms of one variable to that of another. Consider, for instance, the following example, in which a switch is placed in the circuit considered earier. The resulting circuit is represented in Figure 7.3 (p.303).

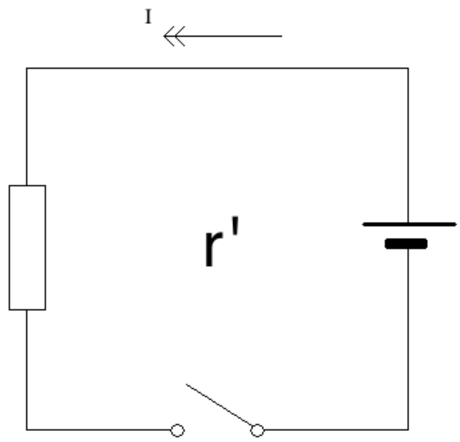

Figure 7.3: Cell, Resistor & Switch in Series

The current flow through the circuit will occur if the switch is closed. So, we might plausibly attribute any mechanical force applied to the switch which results in its closure as a cause of the resultant current flow. Here, the switch's closure seems to be playing the role of a system link between the application of a force and the current flow through the circuit. But it may not be immediately obvious how to make sense of this within the framework provided. After all, the variable concerning the position of the switch is not a variable which features in known electrical laws, and there are no obvious laws analogous to the Planck-Einstein law which license a translation of the event described as a spatial position of the switch to a description of the same event in terms of properties relevant to electrical laws.

However, it is a condition of the system-type in Ohm's law that its assemblies must comprise a conductive component. And we can see that the closure of the switch *in this token case* will allow the assembly represented in the diagram to satisfy the conditions of assembly for that system-type where it currently does not.[5] Therefore, it does appear that *in the specific scenario represented in the above diagram*, that the event of the switch's closure can be *identified* with the assembly's satisfaction of a condition of assembly for the system-type in Ohm's law. This is the case, regardless of any mediating law to license the identification.

Once we grant such a token event identification, it can form the system-link between the mechanical switch system and the electronic system. The total force F applied perpendicular to the switch-lever at a distance d from the pivot creates a torque, which we can causally link to the angle of the switch-lever ϕ by recognising that the switch is an instance of the antecedent of the following law.[6]

> **Torque and angle law:** All inertial rigid bodies rotated from rest about a frictionless pivot have distance d, moment of inertia I, applied force F, initial angle ϕ_0 and ϕ after time t related by the formula,
>
> $$\frac{d}{I}\iint_0^t F dt.dt + \phi_0 = \phi$$

I take it that in this law, ϕ is the focal-variable. Hence, we can infer that the

[5] Although I do not say that the assembly is *dependent* on closure for satisfaction of the conditions. For this is may not in general be the case (see the preemption case in §7.4).

[6] This law can be derived from the angular analogue of the total force law (where $\mathbf{L}$ is angular momentum):

> **Angular force law:** All rotating inertial rigid bodies have properties $F, d, \mathbf{L}$ related by the formula,
>
> $$Fd = \frac{d\mathbf{L}}{dt}$$

If we integrate twice the formula and substitute for momentum according to the relation $\mathbf{L} = I\frac{d\phi}{dt}$ we get the behavioural formula for the torque and angle law.

switch exhibits an intra-system causal relationship between force and angle. If we now incorporate as the system-link the event identified both as the switch taking a particular angle ϕ and as the circuit coming to satisfy a condition of assembly for the system-type in Ohm's law then we have an inter-system causal path from force on the switch to current in the circuit.

To adopt a view of events as identifiable under different descriptions is not, in and of itself, controversial. However, my treatment may appear incompatible with any one description being that which captures the event's 'true essence'. Otherwise the analysis would be at risk of deeming some intra-system causal relations merely superficial in comparison with others. This takes us away from Kim's (1976) and Lewis's (1986d) individuation criteria and towards a more Davidsonian view according to which token events are particulars (see especially Davidson 1976, but also Davidson 1967, 1985, Quine 1985).

Despite Davidson struggling to convincingly settle on a criterion of identity for events, I think his characterisation of events to be anyway the most promising of the three just mentioned. After all, redescription of token events in the way required for inter-system causation is ubiquitous. It is present when we say that an act of apology was someone's saying 'I apologise', or when we identify someone's intentionally burning some scrap paper with their unintentional burning of a valuable document (Davidson 1976, 296). As Davidson pointed out, we also make use of the redescription of events when we explain them.

> Last week there was a catastrophe in the village. In the course of explaining why it happened, we need to redescribe it, perhaps as an avalanche. There are rough statistical laws [generalisations] about avalanches: avalanches tend to occur when heavy snow falls after a period of melting and freezing, so that the new snow does not bind to the old. But we could go further in explaining this avalanche—why it came just when it did, why it covered the area it did, and so forth—if we described it in still a different and more

precise vocabulary. (Ibid.)

While Lewis and Kim's individuation of events has the facility for some redescription, they still require events to have something like an essence or fundamental description. For Kim, an event's essence comprises the property, object and time underlying an event; for Lewis, it is the (contextually defined) class-relation which identifies instances of the event across possible worlds. The benefit of Davidson's treatment is that events have no fundamental characterisation. In the terms of the nomological account, an event is not defined essentially by its role in any single system. In the token case, the closure of the switch just is the forming of a closed loop of conductive material and just is a lever moving through a certain angle ϕ. None of these descriptions is *the essential description*.

Much of the past concern with the Davidsonian view has been over which criterion of identity should be employed. Davidson famously vacillated between a causal-individuation and a space-time individuation. Both suffer problems. The causal-individuation can seem to result in circularity since causes and effects must be individuated first (although see Horsten 2010, De Clercq et al. 2014), and since they are events, we seem to need the criterion of identity for events to pick them out. But the space-time individuation can seem too coarse-grained, as when a sphere both rotates through an angle and heats up simultaneously (Davidson discusses this example at the end of his 1976 and further in Davidson 1985 and Quine 1985). Here is not the place to enter into any detail on the matter. It should suffice for the sake of the present account to note that the space-time individuation appears at least *necessary* condition on event-individuation, and I don't think we should take our current lack of a full analysis to indicate any sense of incorrectness with the view. After all, it's not like Kim's and Lewis's own views don't also have problems of individuation. Indeed, it is a merit of Davidson's approach that the issues of individuation can be laid bare so clearly.

I will from hereon assume that token-identification of events of the sort made plausible by Davidson are another way in which system-links can be understood when observing sequences of intra-system causation chained together to form inter-system causal relations.

7.2.6 Is all causation inter-system causation?

The proposal for a nomological analysis of token causation is completed by the following claim.

Token Causation

c is a token cause of e if and only if c is an inter-system cause of e.

But is this plausible? A first and obvious complaint is that there will be plenty of causation which doesn't fall into law-based causation. I cannot respond comprehensively to this kind of complaint here. After all, without counterexamples to the analysis or any substantial explanation as to why the analysis is limited to back up the complaint, it is hard to combat properly. But I can point out that the preceding points concerning event-identity mean that all that is required for all causation to be subsumed under the analysis is for it to consist in events which *under some description* can be linked by chains of instantiations of laws' system-types. Moreover, these descriptions don't even have to be remain at the same 'level' of description throughout a particular instance of inter-system causation. Having said that, if one thought that physics was on the way to being both causally closed, entirely law-governed and at a constituent level of every causal relationship then one might expect that a redescription of the events in the terms of physics would support the proposal (see, e.g. Davidson 1970, 1995).

A second conceivable complaint which again queries the completeness of the analysis is that the stipulation that causation follows the temporal order is overly restrictive. A response which should placate this kind of objection comes from noticing three things. First, while the proposal prohibits retro-causation out of hand (i.e. causation in which the effects precede their causes) it does not prohibit simultaneous causation. I will argue in §7.5 that this in fact a significant boon for the account. Second, almost *no* non-causal account of causation in the literature today is completely compatible with retro-causation. As I will point out later on, dependency accounts invariably have to posit temporal order at some point of the analysis in order to have a hope of avoiding counterexample. And I will suggest later why accounts which focus on power-manifestations for their analysis of causation are arguably committed to causation which is *only* simultaneous, i.e. neither past-directed nor forward-directed. Third, the constraint on temporal order in the account comes from explicit stipulation. If we decided, for instance that there were determinable cases of causation which are retroactive, then the current account could be amended to incorporate them still making use of the idea that a nomological analysis will draw on connections with causal junctions. After all the defining feature of the proposed analysis concerns the causal asymmetries in causal junctions which have been defined purely probabilistically i.e. without the assumption of a temporal order. What needs to be the case for the account to work is that the causal order maps to *some* non-causal order such that the latter can define the former and be incorporated within the general nomological framework given above. If this order is not a temporal order, then maybe some adjustments will need to be made, but this is a long way off a complete overhaul of the proposed analysis.

A third complaint one might make against the proposed analysis is that the assumption of transitivity leads the analysis to make some unsound judgements of causation where there is none. Failures of transistivity have often been discussed in the causation literature (e.g. in McDermott 1995, Lewis 2000, Hitchcock 2001b, Hall 2004a,b) so it is incumbent on me to say something about such possibilities.

One way of response to apparent cases of transitivity-failure is simply to deny the cases are genuine: causation just is transitive and appearances to the contrary are to be explained away. This is the response of a number of admirable philosophers (Lewis 2000, Hall 2004a) and in some cases seems highly defensible. However, its worth pointing out that the systems-based account is not *in principle* prohibitive of a further qualification which denies transitivity in certain cases. It seems plausible to me that whatever it is that does lead us to suspect that transitivity has failed in certain cases is a feature of causal analysis which extends beyond that of basic causal structure. For instance, we might attribute transitivity-failure to cases where the net result of a causal influence does not change the effect we're interested in from what it would have been otherwise (Hitchcock 2001a), or from some value considered a default (along the lines suggested in Hitchcock and Knobe 2009). But in each of these explanations, we understand that there exists an underlying structure of transitive relations independent of net effects or contextual parameters. It is always possible, then, to take recourse to the view that the above analysis is for this basic kind of causal structure and a full analysis of causation proper is pending some further contextual parameters.

The final complaint I will consider is that the proposal is overly disjunctive. One disjunction exists in the analysis between the binary causal variables for system-types' conditions of assembly and the multi-valued variables in behavioural formulae. Another disjunction exists between intra-system causation and inter-system causation. Finally, a disjunction exists between the instantaneous direct intra-system causation and the non-instantaneous inertial intra-system causation. Is it plausible that token causation should admit of so many subdivisions?

Let me respond to the concern over each of these sub-divisions in turn. On the first disjunction, the distinction between binary variables for conditions of assembly and the multi-valued variables in behavioural formulae is a distinction which comes about from the way in which laws are characterised rather than from causation itself. While the binary variables represent properties at the law-level, multi-valued variables represent properties at the variable-level (although recall from the functional relationships considered in §3.1.3 that this need not be the case). But it is quite possible to conceive of the relations between all these variables as part of one single causal theory. In §6.2.5 I showed how this might be done specifically with the example of the causal asymmetries associated with Ohm's law. If this can in general be achieved, then the types of causation at the law-level and at the variable-level need not be understood as qualitatively distinct from a modelling perspective.

On the second disjunction, whether a causal relationship comes out as one of intra-system causation or inter-system causation may be relative to a particular description of the events and laws employed to pair them. Earlier, I argued that a resistor's length was an inter-system cause of the current flow in that resistor, since I drew on two laws and two corresponding system-events to show how they were related. But I also admitted that there does not seem to me any deep reason why Ohm's law and Pouillet's law could not be combined to form a new law in which the correspondent system-events related instances of length and current as those of intra-system causation. The title 'law' is often reserved for particular relations for reasons of mere historical contingency, and no doubt if a generalisation inferred from combining accepted laws is too overloaded with variables or too detailed in its conditions of assembly then we are unlikely to call it a law. But for the sake of causal *analysis*, combinations of laws which stick to the right logical form (or better still, are RCJLs) can be treated as laws themselves, as is suggested by the account provided in §5.5.[7]

On the third and final disjunction, while the previous subdivisions are fairly superficial, the distinction between simultaneous and non-simultaneous

[7]Although I don't take this wide scope to imply the much stronger and, to my mind, dubious thesis that any consequence of a law is a law; see the discussion in §2.2.

causation is, I think, actually quite fundamental. Shortly, I will show why too much focus on one or the other has restricted the scope of many accounts of causation. But that doesn't mean the two aren't importantly different. Arguably, a crucial failure of those who adhered to the 'causation-mirroring conception' of laws (see §4) can be attributed to the assumption that causation is the same in both cases. Indeed, insofar as token causation is something analysable by the kinds of laws which we know exist in this world—roughly, laws of succession and laws of coexistence, see fn.3—I think we should expect a clear division in the analysis between simultaneous and non-simultaneous causation.

In sum, I take the disjunctive nature of the analysis provided for token causation to be in some respects tolerable, in others admirable. The analysis presents a way of using laws to analyse causation which seems to capture well the kind of explanatory measures we commonly take. In §7.3 I will draw on an extended example to show just how adept and continuous with explanatory practice the analysis is. But despite this continuity the analysis is fairly novel from the perspective of philosophical literature on causation. The reason for this, I believe, stems from the long shadow cast by the misguided causation-mirroring conception (see §4) which has lead to a fairly systematic misinterpretation of the logic, and therefore also the capabilities, of laws. One of the main symptoms of the misinterpretation seems to me the idea that laws explain by supporting entailments of the effects from the causes. It is, therefore, a particular novelty of the analysis of causation proposed here that it uses laws as the key analysans without making this claim.

7.2.7 Nomological analysis without entailment

I have proposed a nomological analysis of token causation summarised by the claim that all token causation is inter-system causation. The analysis is nomological because according to it, it is instantiations of laws' system-types (i.e. 'system-events') and behavioural properties which provide the conditions according to which token causal relations exist.

Under the right specification, it is often possible to use the law to *deduce* the value instanced by the effect-variable (or 'focal-variable') on the basis of its intra-system causes. In such cases intra-system causation would admit something akin to lawful entailment of the effect. But despite the potential for entailment from causes to effects, under the proposed analysis of token causation, this is neither sufficient nor necessary for intra-system causation.

Entailment isn't sufficient (at least at the variable-level) because laws typically license symmetrical inferences. It is only with some further non-logical analysis of the type-level causal junctions associated with laws (e.g. in terms of probabilistic dependencies, see §6.2.4) that we can actually tell which entailments correspond to a causal asymmetry.

Lawful entailment also isn't necessary for intra-system causation since in cases where a system-event is the instantiation of the system-type of an idealising law, the behavioural formula do not provide precise entailments. This is clearest when we consider the behaviour to be modified in the way suggested in §3.3.3. This modification says that the formulae in idealising laws is a *special case* of whatever behaviour is actually exhibited by the system and so cannot be expected to entail accurate values for the instantiated variables.

For longer causal chains, i.e. inter-system causal relations, the notion of lawful entailment becomes even less relevant. Value-instances of the focal-variable in one system-event are identified with value-instances of causal-variables in other system-events which can occur in the distant past. This step can be entirely contingent on the particular set up of the scenario and on the kind of token identifications of events available under different descriptions. Hence, any kind of entailment is typically out of the question.

It's worth comparing the view developed here once again with the commonplace assumption that insofar as laws help analyse token causation, they do so by entailing the effect from the conclusion (see, e.g., Hempel 1965, Armstrong and Heathcote 1991, Armstrong 1997, 2004a, Paul 2000, Schaffer 2001). Ever since Hempel and Oppenheim (1948) suggested their covering law account of causal explanation, it has been commonplace to criticise such accounts. But in criticising them, theorists seem to have missed the opportunity to make use of laws in analysis in the rather different way I have presented above. Clearly according to the current analysis, entailment is beside the point. We may use the mathematically precise features of laws to make predictions, but this is not what constitutes the ability for laws to *analyse causation*.

7.3 Extended example: The Challenger Disaster

In this section, I will present a brief representation of how the nomological analysis as I have described it can be used to explain extended causal relationships in a real-world example. I will refrain from going into specific details about numerical values; after all, often we don't have access to these. Indeed, given the foregoing section it should be considered a merit of the analysis that it doesn't require specific values to work. Moreover, I think the example draws attention to the plausibility and informativeness of the analysis, particularly in its ability to help us provide the kinds of explanation for causal relations we actually might look for.

The example I will use is the explosion of the Space-Shuttle Challenger on 28th January 1986 in which all seven crew lost their lives. The enormity of such a public, expensive and heavily engineered tragedy instigated a very long and rigorous inquiry by a US Presidential Commission. Their conclusion was given as follows.

> In view of the findings, the Commission concluded that the cause of the Challenger accident was the failure of the pressure seal in the aft field joint of the right Solid Rocket Motor (*Report of the Presidential Commission on the Space Shuttle Challenger Accident*).

Let's call this conclusion by the Commission the *'primary causal claim'*. The justification for the primary causal claim was a result of the connection they were able to draw between the failure of the pressure seal to the explosion of the external tank through a number of intermediate events part of which is summarised in the following passage from the Commission's report.

> The destruction of the [solid rocket motor] seals caused hot gases to leak through the joint during the propellant burn of the rocket motor [40]. The flame from the hole impinged on the External Tank and caused a failure at the aft connection at the External Tank [76-77]. Since the internal pressure of the liquid hydrogen tank is at approximately 33 pounds per square inch, a sudden venting at the aft section will produce a large initial thrust that tails off as the pressure drops [67] (*Report of the Presidential Commission on the Space Shuttle Challenger Accident, Chapter IV*).

Subsequent events were captured in two moments in a timeline of the launch prepared shortly it after it happened. Two relevant entries are the following:

> T+73.124: The resulting forward acceleration begins pushing the [hydrogen] tank up into the liquid oxygen section in the tip of the external fuel tank.
>
> T+73.213: An explosion occurs near the forward part of the tank where the solid rocket boosters attach. (*Timeline on the Space-shuttle Challenger explosion assembled by United Press*

International's Cape Canaveral Bureau chief William Harwood and radio chief Rob Navias.)

There are a number of mediating events we can infer from the report and timeline. The following appear particularly salient: *seal-destruction*, the destruction of the seals; *gas-leak*, the gas leak from the right solid rocket motor; *tank-impingement*, the impingement of the gas leak on the external tank; *tank-failure*, the failure of the liquid hydrogen tank; *tank-thrust*, the thrust of the hydrogen tank; *tank-collision*, the fusion of the liquid hydrogen and oxygen tanks; *fuel-conflagration*, the spontaneous conflagration of the fuel in the external tank.

Assuming the defining moment of the accident can be identified with the last event in this sequence, the primary causal claim can therefore be supported by the following connected token causal claims:

1. *seal-destruction* caused *gas-leak.*
2. *gas-leak* caused *tank-impingement.*
3. *tank-impingement* caused *tank-failure.*
4. *tank-failure* caused *tank-thrust.*
5. *tank-thrust* caused *tank-collision.*
6. *tank-collision* caused *fuel-conflagration.*

In analysing these relations, a pre-emptive task will be to notice the different descriptions we can identify the relata under. Table 7.1 (p.317) suggests some plausible token relata which will be helpful in what follows. In each case, I suspect some of the identifications can be challenged. For instance, I have suggested that we might identify *tank-failure* both with the flux-density across the hydrogen tank reaching a critical point and with that area of the tank failing. Perhaps a more precise explanation of the Challenger disaster would distinguish these latter two value-instances and instead find a way to link them by inter-system causation. I suspect whether we do this in practice or not depends on the precision we require. For instance, in some circumstances we might identify a catastrophe with an avalanche, including along with it the destruction of houses, etc. In other circumstances, we might prefer to identify the catastrophe with the *consequences* of the avalanche, i.e. the effects of the pressure of the compacted snow on the walls of the buildings. I see no reason why admitting this would compromise the general fact that we can in general successfully identify events under different descriptions (compare the contextually relative validity of identifying the referent of 'Holland' with 'The Netherlands'.)

I will now describe how we can justify this causal sequence in terms of inter-system causation (for simplicity I omit the behaviour-modification of each law used in the analysis).

seal-destruction* caused *gas-leak. *Seal-destruction* was an event which took place in the right solid rocket motor as it warmed up during launch. Since the rubber o-rings which constituted part of the seal were slow to expand during the rocket booster's initial thrust, the seals between the booster's tang and clevis joint exposed a gap through which the internal ignited gases could be forced out under the booster's internal pressure.

We may identify the system-event $S\text{-}orifice\text{-}flow_a^{\Delta t1-t5}$ in which, from soon after the launch at $t1$ to the explosion at $t5$, an assembly a comprising the destroyed seal and escaping gases instantiate the system-type of the following fluid dynamical *robust causal junction law* (RCJL).

> **Orifice-flow law:** For all x, if x is a gas flowing from an orifice then it has variables pressure P, mass flow-rate $\dot{M}$ and area of the orifice A which satisfy the formula,
>
> $$P = \frac{\dot{M}}{2A}.$$

Since *seal-destruction* can be identified with the presence of an orifice at the seals and *gas-leak* with the mass-flow rate of gas at the orifice (see Table 7.1),

Table 7.1: Event-sequence of the Challenger accident

key times	descriptions		
t1 -t5	*seal-destruction*	$\Sigma_{t=1}^{5} A_{seals}^{t} \Leftarrow a(t)$ (the presence of an orifice at the seals)	
	gas-leak	$\Sigma_{t=1}^{5} \dot{m}_{gas}^{t} \Leftarrow \dot{m}(t)$ (the mass flow-rate of the gas at the orifice)	$\Sigma_{t=1}^{5} m_{gas}^{t1} \Leftarrow m(t)$ (the presence of some quantity of gas at the orifice)
t2-t4	*tank-impingement*	$\Sigma_{t=2}^{4} m_{gas}^{t} \Leftarrow m(t)$ (the presence of some quantity of the gas at the hydrogen tank)	$\Sigma_{t=2}^{4} T_{gas}^{t} \Leftarrow T(t)$ (presence of a temperature gradient across the hydrogen tank's walls)
t3	*tank-failure*	$\mathbf{q}_{tank}^{t3} \Leftarrow \mathbf{q}(t3)$ (the flux density across the hydrogen tank reaching critical)	$A_{tank}^{t3} \Leftarrow A(t3)$ (an area of the hydrogen tank's wall failing)
t3-t5	*tank-thrust*	$\Sigma_{t=3}^{5} F_{tank}^{t} \Leftarrow F(t)$ (the thrust of the hydrogen tank)	
t4-t5	*tank-collision*	$\Sigma_{t=3}^{5} \sigma_{tanks}^{t} \Leftarrow \sigma(t)$ (the stress on both tanks at the point of collision)	$\Sigma_{t=1}^{5} mix_{tanks}^{t} \Leftarrow mix(t)$ (the mixing of the fuels)
	fuel-conflagration	$\Sigma_{t=4}^{5} E_{fuel}^{t} \Leftarrow e$ (the combustion of the fuel)	

the two events are system-paired in *S-orifice-flow*$_a^{\Delta t1}$. Moreover, because they both occur between $t1$ and $t5$, and $\dot{M}$ is the focal-variable of the Orifice-flow law we can infer that every value-instance comprising *seal-destruction* is a direct intra-system cause of a value-instance comprising *gas-leak* for all times $t1 - t5$.

gas-leak* caused *tank-impingement. *Gas-leak* began at the destroyed seals in the rocket booster and flowed out towards the external tank. We may identify the system-event *S-gas-flow*$_b^{\Delta t1-t5}$ in which an assembly comprising the gas itself between $t1$ and $t5$ instantiates the system-type of the following conservation RCJL.

> **Gas-flow law:** For all x, if x is a continuous flow of fluid across some distance then it has variables mass M, density ρ and volume V which satisfy the formula,
>
> $$\dot{M} = \rho\dot{V}.$$

Since *gas-leak* can be identified with the presence of gas at the orifice and *tank-impingement* can be identified with the presence of gas at the tank, their time-derivatives are system-paired in *S-gas-flow*$_b^{\Delta t1-t3}$. Moreover, since *gas-leak* commences earlier than *tank-impingement* and $\dot{M}$ is the focal-variable of the Gas-flow law, we can infer that there exist value-instances comprising *gas-leak* which are inertial intra-system cause of value-instances comprising *tank-impingement.*

tank-impingement* caused *tank-failure. As the gas impinged on the external tank it heated the aft section up so much that the tank failed. We may identify the system-event *S-thermal*$_c^{\Delta t2-t3}$ in which an assembly comprising the gas and the walls of the external tank between $t2$ and $t3$ instantiates the system-type of the following thermal conduction RCJL.

> **Fourier's heat law:** For all x, if x is thermal gradient across a

> thermally conductive material then it has variables flux-density $\mathbf{q}$, heat-conductivity k and temperature T which satisfy the formula,
>
> $$\mathbf{q} = -k\nabla T.$$

Since *tank-impingement* can be identified with the presence of a temperature gradient across the tank hydrogen tank's walls (a condition of assembly for Fourier's heat law) and *tank-failure* can be identified with the flux-density across the hydrogen tank reaching a critical point at $t3$, the two events are system-paired in $S\text{-}thermal_c^{\Delta t2-t3}$. Moreover, since they both occur between $t2$ and $t3$, and $\mathbf{q}$ is the focal-variable of Fourier's heat law, we can infer that every value-instance comprising *tank-impingement* is a direct intra-system cause of a value-instance comprising *tank-failure*.

***tank-failure* caused *tank-thrust*.** As described by the presidential commission, 'a sudden venting at the aft section [of the external tank] will produce a large initial thrust'. We may identify the system-event $S\text{-}thrust_d^{\Delta t3-t5}$ in which an assembly comprising the pressurised liquid hydrogen in the hydrogen tank (with its failed aft-section) between $t3$ and $t5$ instantiates the system-type of the thrust RCJL,

> **Thrust law:** For all x, if x is liquid flowing from an orifice in a pressurised container then it has variables pressure of the liquid hydrogen P_l, area A and Thrust force $\mathbf{F}$ which satisfy the formula,
>
> $$F = P_l A.$$

Since we can identify *tank-failure* with the presence of an area of the tank which has failed and *tank-thrust* can be identified with the hydrogen tank's thrust force, the two events are system-paired in $S\text{-}thrust_d^{\Delta t3-t5}$. Moreover, since both occur between $t3$ and $t5$ and F is the focal-variable of the Thrust law, we can infer that every value-instance comprising *tank-failure* is a direct intra-system cause of a value-instance comprising *tank-thrust*.

tank-thrust* caused *tank-collision. The propulsion of the hydrogen tank into the oxygen tank caused a the two to undergo stress at the point of collision. We may identify the system-event $S\text{-}stress_e^{\Delta t4-t5}$ in which an assembly comprising the hydrogen tank and oxygen tank between $t4$ and $t5$ instantiates the system-type of the stress RCJL,

> **Stress law:** For all x, if x is a collision between two objects then it has variables stress σ, area A and force $\mathbf{F}$ which satisfy the formula,
>
> $$\sigma = \frac{\mathbf{F}}{A}.$$

Since *tank-thrust* can be identified with the hydrogen tank's thrust force and *tank-collision* can be identified with the stress on both tanks at the point of collision, the two events are system-paired in $S\text{-}stress_e^{\Delta t4-t5}$. Moreover, since the events both occur between $t4$ and $t5$, and σ is the focal-variable of the Stress law, we can infer that every value-instance of *tank-thrust* is a direct intra-system cause of a value instance comprising *tank-collision.*

***tank-collision* caused *fuel-conflagration*:** At the point of collision between hydrogen and oxygen tanks the fuels started to mix. We can identify the system-event $S\text{-}combustion_f^{\Delta t4-t5}$ in which an assembly comprising the hydrogen and oxygen between $t4$ and $t5$ instantiates the system-type of the chemical reaction RCJL,

> **Hydrogen combustion law:** For all x, if x is the combustion of hydrogen then it has variables energy E, lower heating value of hydrogen LHV_H and mass of hydrogen m_H which satisfy the formula,
>
> $$E = LHV_H \times m_H.$$

Since *tank-collision* can be identified with the mixing of the hydrogen and oxygen (a condition of assembly for the combustion law) and *fuel-conflagration* can be identified with the release of energy E the two are system-paired in S-$combustion_f^{\Delta t4-t5}$. Moreover, since they both occur between $t4$ and $t5$, and E is the focal-variable of the Hydrogen combustion law, we can infer that every value-instance of *tank-collision* is a direct intra-system cause of a value-instance comprising *fuel-conflagration*.

According to the nomological analysis I provided in §7.2, c is a token cause of e if and only if c is an inter-system cause of e, i.e. connected by a chain of direct and inertial intra-system causal relations. Since there are temporal parts of all the events in the causal sequence which are connected to some parts of subsequent events in the sequence either via direct or inertial intra-system causation, the above reasoning allows us to confirm the primary causal claim made by the Presidential Commission.

Working through extended examples like this exposes the immense power of the proposed nomological analysis of token causation. For it is natural in explanation of token causal relationships to draw on actual known laws where possible and yet, so far as I know, no account of token causation since Hempel's has proposed any general method by which this should be done.[8]

Hempel's (and Oppenheim's) approach was reprimanded for requiring that the effect be *deduced* or *statistically inferred* from the conjunction of cause with laws, and also for the fact that laws do not seem to exhibit the required type-level asymmetry. But the analysis I have proposed rectifies both these issues. It rectifies the first by requiring only that the cause and effect be linked via a chain of suitable system-pairs, where there is no expectation of deduction or statistical inference (see §7.2.7). The analysis also rectifies the second issue concerning the symmetry of laws' deductive/statistical entailments by showing how the asymmetrical structure of a causal junction is associated with the central class of robust causal junction laws which are

[8]There are accounts which bear similarities, e.g. Strevens (2008), Baumgartner (2008, 2013a), but it seems to me that none of these accounts provide any method for drawing explicitly actual laws which are actually drawn on in practice.

inferentially associated with type-level causal asymmetries (perhaps themselves analysable in probabilistic terms; see §6.2).

So much for the success of my nomological analysis in consideration of other nomological approaches to causation. I now turn to consider its success in consideration of other approaches entirely.

7.4 Classic problem-cases

In this section I show how the proposed nomological analysis of token causation tackles two kinds of causal scenario known to be problematic for other popular approaches to analysis of causation. While I don't take this to provide any clear conclusive indication of the superiority of the approach, it should certainly indicate its competitiveness.

Dependency theories understand causal relations to be defined in some manner by effects' depenency on their causes. Examples include counterfactual theories of causation (1973a, 1986g, 2000, 2001, 2007), and probability-raiser theories (e.g. Menzies 1996, Kvart 2004, Fenton-Glynn 2011). In their simplest form, dependency theories have often suffered from not being able to judge the causation in cases of preemption successfully. There have been many attempts to respond to these concerns which I will not delve into here. But where there has been success, it is typically due to conscious and baroque adaptation of the basic premise that causation is rooted in dependency. One could be forgiven, therefore, for thinking that the amendments to the theory are unreasonable post-hoc fixes to an essentially flawed approach. Hence, a different approach which is already able to deal with cases of preemption in its nascent state should bear some appeal.

Consider Figure 7.4 (p.323; developed from an example in Pearl 2000). Switch a is a single-pole double-throw (SPDT) switch and has two positions: up or down. Switch b is a single-pole single-throw (SPST) switch which has two positions: open or closed. If a moves down and b closes, the bulb will

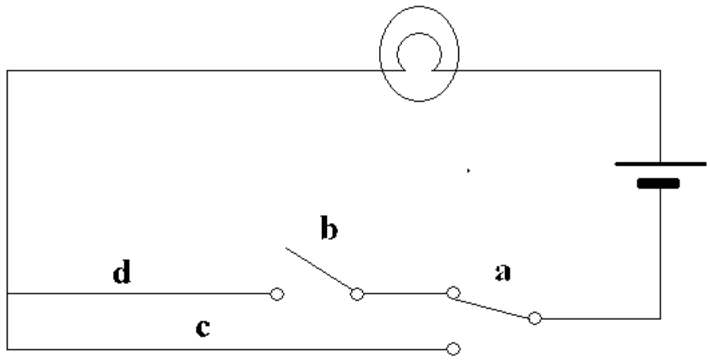

Figure 7.4: SPDT and SPST in series

illuminate. But a's movement will also prevent b from having any causal relevance to the lighting of the bulb, since by moving down, a prevents b from being in a circuit with the bulb. In such a scenario it is natural to treat a's moving down as the cause and b's closure as the pre-empted non-cause, or *'fizzler'* (Schaffer 2001).

As with all cases of preemption, the scenario involving Figure 7.4 poses a problem for basic dependency theories because the effect (the bulb's illumination) is not dependent on the cause (a's moving down). The example represents a case of causation without counterfactual dependence, since if a had not moved down, then the bulb would still have lit. If we assume the system is deterministic, then the example represents a case of causation without probabilistic dependence, since given that b closed, the probability of the bulb lighting was not increased by a moving down. Moreover, if we assume the system is indeterministic, then the example also represents a case of probabilistic dependence without causation, since given that a moved down, the probability of the bulb's lighting would still be increased somewhat by b closing

However, if we follow the nomological analysis provided in §7.2 and treat token causation as inter-system causation, we have a simple solution. When the bulb is lit it is in an instance of the system-type of Kirchhoff's voltage law.

> **Kirchhoff's voltage law:** For all closed electronic loops, the sum of resistances ΣR, current through each conductor I and the total electromotive force Σemf available in that loop are related by the formula,
> $$\Sigma emf = \Sigma R \times I$$

Since the bulb's illumination can be identified with (or at least has as an uncontroversial intermediary event) the current through the bulb, it will be a value-instance in the system-event consisting in the assembly instantiating the system-type in Kirchhoff's law. Moreover, since current is the focal-variable of Kirchhoff's voltage law, the bulb's illumination is directly intra-system caused by any simultaneous value-instance of another variable (e.g. Σemf or ΣR) or by any simultaneous satisfaction of a condition of assembly for the law's system-type. Clearly, the movement of switch a into the down position will be one of these latter intra-system causes. For it is the event of the switch becoming part of the closed electronic loop which the bulb is in. In comparison, the closure of switch b, is no such event, since switch b does not come to be in the closed loop with the bulb when moved down. Switch a is a direct intra-system cause, hence an inter-system cause, hence a token cause of the bulb's illumination according to the proposed nomological analysis. Switch b does not appear from the setup to be any kind of token cause under the analysis.

Despite the terminology I use to make the causal relevance of switch a clear in the example, I take it that the explanation charts a very basic sort of reasoning which we actually make in the scenario (an was in fact already made when introducing the example): that switch a is part of the relevant circuit/system and b is not. The analysis can capture formally this reasoning partly due to the fact that it does not rely on any notion of dependency to analyse token causation. Rather, it is concerned with the actual events which take place.

In this regard the approach is similar to physical connection views of causation. These views characterise causation in terms of the physical connection which causes have to their effects, e.g., via some flow of energy or momentum (e.g. Aronson 1971, 1985, Fair 1979, Ehring 1997, Salmon 1998b, Dowe 2000, Rueger 2006). The competitive edge of physical connection views over dependency views is most obviously their adeptness when dealing with preemption cases. However, they have typically struggled to handle a different kind of causal relation we called 'negative causation'.

Negative causation exists when at least one relatum consists in an object's failure to instantiate a particular property at some time, such as in cases of causation by omission or prevention. Despite some philosophical argument to the contrary (see Dowe 2000, Beebee 2004) it seems natural on occasion to treat negative causation as genuine causation (see Lewis 2004, Schaffer 2000, 2003). But since such relations include events consisting in the absence of energy or momentum, negative causation is not easily accommodated by the physical connection views.

The nomological analysis of causation as inter-system causation does not have such restrictions and is, subsequently, able to easily accommodate many cases of negative causation. Consider the circuit represented in Figure 7.5 (p.325).

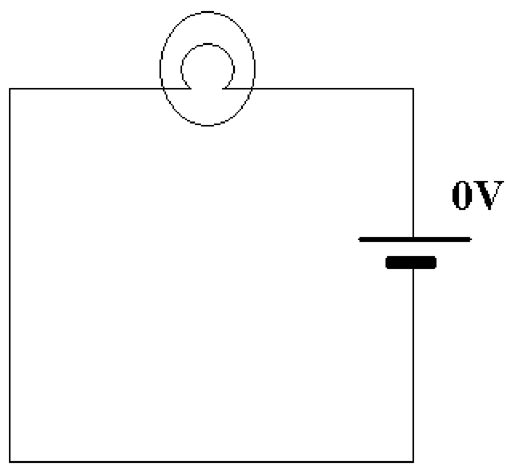

Figure 7.5: Drained Cell & Bulb in Series

In this circuit, there is no current supplied to the bulb because the cell is supplying no electromotive force (emf). This appears a case of negative causation since the cause of the bulb's failure to light is itself a negative event consisting in an object's *failure* to have nonzero emf. Subsequently, physical connection views will not (traditionally) treat such a case as one of causation; after all, there is certainly no flow of energy or momentum from the cell to the bulb.

Yet, notice how easy the nomological analysis gets an affirmation of causation in this case. For again, in this scenario we have an instance of the system-type in Kirchhoff's voltage law. That is because, despite there being zero emf produced by the cell, the whole assembly still forms a closed electronic loop. Subsequently, the event of the cell having zero emf is still an event *system-paired* with the value of current supplied to the bulb. Moreover, since current is the focal-variable of Kirchhoff's law, we can infer that the value of emf produced by the cell (viz. zero) is a direct intra-system cause, hence inter-system cause, hence token cause of the simultaneous zero current through the bulb. Notice also, that the solution doesn't entail that *any* drained cell is similarly a cause of the bulb's lack of illumination nor does it invoke contextually dependent default or deviant properties in order to avoid this result (this has been seen as a necessary recourse for some dependency views; see the discussion in Beebee (2004), Hall (2007)).

Many dependency views and physical connection views have found ways to deal carefully with instances of their respective problem-cases which are as simple as those have presented. But I take it to be a significant merit of the nomological analysis that it does not suffer from either kind of case in its very initial presentation. After all, the analysis was not developed specifically with such cases in mind (unlike many of the baroque dependency and physical connection views; see, e.g., Lewis 2000, Halpern and Pearl 2001, Hall 2007, Fenton-Glynn 2011 for advanced dependency views, Fair 1979, Salmon 1994, Ehring 1997 for advanced physical connection views).

7.5 The Accommodation Criterion

A final merit of the proposed nomological analysis I want to draw attention to is its ability to accommodate both simultaneous and sequential causation. While I remain unsure about the prospects of physical connection views on this point, it seems to present difficulties both for dependency views and another popular approach to analysing causation which we might call 'powers-based views'.

In the presentation of my proposal, I have assumed that we should prefer an account of causation which can accommodate both simultaneous and sequential causation. Some arguments for the existence of simultaneous causation have already been provided in the foregoing chapters (see, especially, §5.2.3). The one I consider to be the most persuasive came from Huemer and Kovitz (2003) who pointed out that the total force law and Lorentz-force law suggest (under a causal conception) that acceleration, force producing that acceleration, and the presence of electrical and magnetic fields inducing that force are all simultaneous. Notice that such simultaneous relationships bear many of the trademark symptoms of causation. They are handles for the control and manipulation of variables, they can be used to assign responsibility, and so on. Granting such examples shows that it is not so easy to dismiss simultaneous causation on the basis of special relativity. Whereas relativity is only a problem for spatially separated and simultaneous relata, in both of Huemer and Kovitz's examples, the causes are *co-located* with their effects.

Granting that some causation is simultaneous, should we thereby expect all of it to be? Strangely, the weaker view that causes are sometimes simultaneous with their effects has been most prominently defended recently via the stronger view that *all* causation is (for instance, Huemer and Kovitz 2003 seem to take this view, as do Mumford and Anjum 2011, ch.5). But this is highly problematic. To start with, although we may agree that some effects of an applied force are instantaneous with that force (e.g. acceleration) other phenomena are certainly not instantaneous with the application of a force despite being paradigm effects of it. For instance, the movement of an object initially at rest will not be simultaneous with the application of a force. If the force is applied at t_0, then the object will not be moving at t_0, but only (a vanishingly brief but non-zero amount of time) afterwards. This is the case regardless of how much force is applied. Problems like these can be generated for any relationship governed by a differential relation (see the discussion in Easwaran 2014).

Moreover, one might think, as I do, that the distinctness in times of causes and their effects is not limited to cases in which the times *overlap*—causes can happen in the extended past of their effects too! But these kinds of causal influence have been objected to in the past. A common line of argument, started by Russell (1912), is that if an influence entirely precedes its effect, then it cannot be contiguous with it, and hence cannot necessitate the effect. Putting aside whether or not necessitation is an essential feature of causality, the very premise needs questioning. As has been pointed out by Chakravartty (2007), among others, causes may be understood to occur in the open interval up to but not including the time of their effects. And yet Chakravartty has gone on to complain that if causes are continuous with their effects this leaves no way to properly individuate them. So the argument goes, the only part of the cause which can truly have any power to bring about the effect is the very last temporal slice, and assuming time is contiguous, this renders the cause interpreted as an enduring event undefinable.

I don't find this line of reasoning particularly persuasive. Chakravartty's point seems to amount to the point that causes will have vanishingly brief final temporal parts which 'screen off' their influence of the effect from the earlier temporal parts. This would be the case if time is continuous and, for example,

> '... we have [...] a causal chain [... $c \to d \to e$], in which positive

> time is indicated by arrows. [... O]nce we know that d has occurred, it is no longer necessary to know that it was preceded by c; event c is no longer relevant to the prediction of e. The contribution of c to e has been absorbed in d, so to speak; and d may be said to *screen off* x from e' (Reichenbach 1956, 189).

Such cases of screening off seem entirely plausible, perhaps even pervasive. But I see no reason why this fact should preclude some event extended in time and entirely preceding another counting as the cause of that other event. This would seem especially natural if one thinks that causation has plenty of transitive instances. Indeed, it can seem in such cases that screened off causes of an effect are indeed causes *because* they are also causes of the intermediate screening events. After all, we don't claim that only the top layer of bricks in a house (let alone an indefinably small part of the top layer) holds up the roof because they are the only layer in contact with the roof!

But perhaps such indirect causation could be accepted whilst still maintaining that where there is screening off of earlier temporal parts by later temporal parts, those later temporal parts will inevitably be simultaneous with the effect. If true, this might appear to reclaim the fundamentality of simultaneous causation. However, we know that this interpretation cannot in general be true. To show this, let's look at an example.

Consider the situation represented in Figure 7.6 (p.330). A ball is in a stationary but unstable equilibrium between the jaws of a grip up to time t_0 (see diagram in Figure 7.6, p.330). At t_0 a small force F_p is applied which releases the ball downwards out of the grip. The force is applied until t_1 and is removed entirely by t_2, at which point the ball is in freefall. We can ask: is the force F_p a direct influence of the freefall? If it isn't then we must find some relatum entirely distinct from the push which screens it off from the freefall (NB. this is the case even if the push is claimed to be no cause *at all*).

But this will be difficult under any interpretation of screening off. For there is no obvious event distinct from the freefall itself which absorbs the influence of the push on the freefall. Nor is the energy supplied by the push mediated by another object or power. These facts can be stipulated of our hypothetical example to hold so long as there is a smooth and continuous removal of F_p—a perfectly physically plausible scenario.

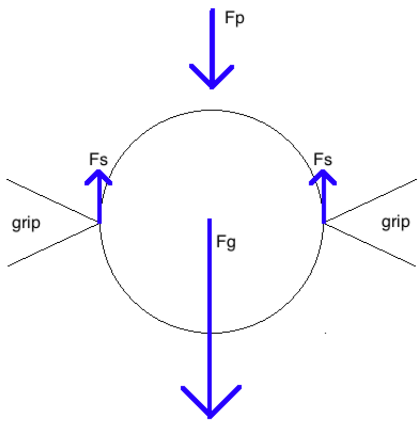

Figure 7.6: ball in grip

But now notice also that the force F_p and the freefall do not overlap in time. This is because the ball cannot be in freefall if the pushing force is still being applied. Hence, the push and the freefall must be *purely sequential*: there are no temporal parts of the cause which both screen off the earlier parts and are simultaneous with the effect. Consequently, the view that putative causes are always screened off by legitimate causes simultaneous with their effects cannot be true.

It has been argued that the sort of case just described is not one of genuine causation. This is a view made explicit by Mumford and Anjum (similar sentiments are present in Huemer and Kovitz 2003, Heil 2012) who say that 'a key point to note [...] is that one causal process is not seen as the cause of another but only as an enabler; thus, we do not need to invoke causation between temporally distinct events' (2011, 127).

But how convincing can such a view be? First, we know it cannot be exactly right simply from consideration of the non-zero time it takes for a force to bring a resting mass to a nonzero velocity. And even if such cases can somehow be incorporated within this view, there are further difficult questions to be addressed. First, from a neutral standpoint, the view seems wildly unmotivated. Sure, if we have some particular metaphysics we're already committed to (as Mumford and Anjum do), then maybe we must deny sequential causation, but from a pre-theoretical position, the view seems unquestionably counter-intuitive. For it denies, ultimately, that the Big-Bang did not cause a rapid expansion of space-time, that over-lending for sub-prime mortgages did not cause the 2008 credit crisis, that my staying up all night did not cause me to be tired this morning. Compare such claims, for example, with those of other philosophers who have tried, and in general failed, to convince us that causation has a wildly different extension than we typically think it does, such as Dowe's (2000) distinction between negative causation and genuine causation, or Russell's (1912) eliminativism about causation in altogether. The literature which responds to these texts reflects the fact that we are more inclined to see their radical conclusions as a *reductio* on their positions rather than anything else (see, e.g., Schaffer 2000, 2003, Bontly 2006).

But why do we find it so implausible when philosophers try to persuade us that causation is not where we thought it once was, specifically between relata separated in time? One reason is that we want to allow that experiments which rely on our ability to affect changes across time are making use of genuine causation. The administration of a drug, for example, can surely be a causal influence on someone's health. If it wasn't then how come it's so effective—or are we not allowed to say even that it is effective? What are we supposedly missing out on by confidently bringing about things in the future that we get when bringing them about simultaneously? To be persuasive, Mumford and Anjum need to show us how it can be that despite our intuitions, our use of the concepts of causation, influence and efficacy can be so consistently wrong. This requires, at least in part, an explanation of how such concepts can mean one thing, when most of us use it to mean another. Not even the most hardened causal reference theorist should doubt that use has some impact on meaning (see, e.g., the discussion on names in Evans 1982, Dummett 1978, McDowell 1998)—and even if they did, presumably their view about reference could not be maintained under a denial of sequential causation anyway.

Another point in support of sequential causation is that it is incumbent on theorists of the ontology of causation to provide an account about what distinguishes real processes from pseudo processes. A beam of light is a real process, the movement of the spot of light across a stage is a pseudo process (see Salmon 1984a, Hitchcock 1995, Dowe 1995, Salmon 1998b, Choi 2002). This distinction is one which appears to demand a causal explanation. But theorists like Mumford and Anjum might seem committed to denying that any temporal stage of a process could be the cause of later temporal stages of the same process. And their denial would therefore seriously complicate the issue of how single processes gets to be extended through time at all. On the one hand, if past stages of a process can't cause latter stages, then how do those latter stages come to be? Mumford and Anjum's use of the word 'enabler' would seem to be appropriate at this stage, but if enabling is capable of doing the important work of bringing about sequences of events or processes, then why is enabling not just plain old causation after all? On the other hand, if past stages of a process can cause its own latter stages, then why can't we say the same for distinct processes too?

This leads to a final point about the need for sequential causation. There is an old problem first raised by Hume, which questions how simultaneous causation can ever provide the grounds for a universe which has duration at all. On the one hand, if enabling were a new metaphysical category, maybe it could be used to make sense of this. But then the view promoted by Mumford and Anjum becomes far less ontologically parsimonious at no obvious benefit to clarity or understanding. On the other hand, if enabling is not some distinct metaphysical category, but a construction from non-causal and real causal relations, then if all causation is simultaneous (or mediated by a vanishingly brief amount of time), then we might expect enablers must be as well. This would make it utterly mysterious how earlier moments of history are explanatory of later moments *at all*.

All in all, there is some proportion of relations which we typically think of as causal and which are extended throughout time. If we are to be persuaded that they are not causal, this won't erase the requirement to understand those relations as metaphysically and explanatorily significant. If enabling is that relation, then it's something philosophers like Mumford and Anjum should be worrying about just as deeply as they do about causation.

The preceding comments are enough for me to believe it is worth steering away both from views which are not open to purely simultaneous causation and views which are not open to purely sequential causation. My conclusion in general is that an analysis of token causation should *accommodate causes which are both simultaneous and sequential with their effects*. That is, it should—*pace* Hume, Mellor, Glynn—allow that certain causes can be simultaneous with their effects, but also that—*pace* Huemer, Kovitz, Mumford and Anjum—allow that causes can completely precede the time in which their effects occur. I call this the *'accommodation criterion'*.

The accommodation criterion sets a constraint on any analysis of token causation if it is to be plausible. There are countless analyses around, but in what follows I will focus on two broad approaches which appear to me particularly prominent in the current metaphysics of science literature, viz. powers-based views and dependency views before considering the nomological analysis proposed above. It turns out that while the former two each struggle to accommodate the accommodation criterion in different ways, the nomological analysis does not.

7.5.1 Powers-based views and the accommodation criterion

Could a powers-based view of causation support the accommodation criterion? My inclinations are to say ‘no’. Often, the dialectic of powers-based views is to argue that an ontology of powers is more suitable for capturing genuinely simultaneous causation than one of related events, such as we find in dependency views. This may bear some truth, as we will shortly find out. But regardless, it is important here to focus on what an ontology of powers cannot do, and it seems the suitability of handling simultaneous causation has, in this case, come at the expense of any ability to handle sequential causation: powers-based views seem *committed* to simultaneity.

To see this, consider first the phenomenon of manifestation. It is sometimes said by powers-theorists that causation is not a relation (see Lowe 2012, Heil 2012). The argument for this idea is that causation should be understood as the manifestation of a power by an agent in some patient. There is, therefore, only actually one relatum to causal occurrences, namely, a manifestation. For example, the dissolving of the sugar-cube is done by the water, but it was just one thing—a dissolving—which happened to the sugar-cube. Causation occurs according to these powers theorists when two things do some one thing together. In this way, causation is not understood to be a relation. But if causation is not a relation, then simultaneity is entailed trivially, since one monadic occurrence has no other relata to be temporally distinct from (although this needn’t preclude the manifestation being a temporally extended process).

True, some powers theorists do not like the Aristotelian idiom of agents and patients, opting instead for a focus on *reciprocation* in causation between ‘power partners’ (see Heil 2003, Martin 2008). In this case, there may not be an agent or patient as such in a causal occurrence, but rather two entities which manifest their powers mutually in each other. If this is the story, then there may indeed be more than one relatum in a causal occurrence. For instance, the sugar-cube is dissolved by the water, but the water also gets sugary because of the sugar-cube. However, simultaneity is still demanded by the fact that the manifestations are *reciprocal*.

A form of analysis of powers which might seem to countenance cross-temporal causation is that given by Bird (2007a). For in discussion of finks he suggests that 'the process whereby a disposition manifests itself will typically take time' and that 'finkish dispositions arise because the time delay between stimulus and manifestation provides an opportunity for the disposition to go out of existence and so halt the process that would bring about the manifestation' (p.25). A closer reading of Bird reveals, however, that his view of powers (or in his terms 'potencies') do not take time to manifest. Indeed, his central argument against the possibility of finks at the fundamental level appears to rely on this fact (assuming time is continuous; see pp.60-62).

I do not presume that powers theorists will find it impossible to accommodate sequential causation, but given this brief survey it seems that they are likely to have difficulty in doing so; or at least, their ontology does not naturally dispose them to such accommodation.

7.5.2 Dependency views and the accommodation criterion

I now turn to consider the compatibility of dependency views and the accommodation criterion.

In their most basic form, counterfactual-dependency analyses tell us that causes of some effect are those distinct events which are such that if they hadn't happened, the effect wouldn't have happened either (see Lewis 1973a, for an early version of the idea). But to make sure the analysis captures the asymmetry of causation, it must be that those depended-on events would still have occurred even if the phenomenon hadn't. If this were not a consequence, then the theory would make causes dependent on their effects just like effects are on their causes and so lack the required asymmetry.

Obviously, the source of this asymmetry will be reflected in whatever semantics is appropriate to give for counterfactuals. So it is a useful place to look if we want to understand the theory better. Many of the philosophical treatments for counterfactuals' assess whether the truth of their consequents are true in worlds where the antecedent is true but otherwise the history is exactly like the actual course of history up to the time of the antecedent (see Jackson 1977, Maudlin 2007; Lewis's semantics is discussed below). If the consequent turns out to be true in those worlds, the counterfactual is true. This approach will capture the asymmetry of sequential, forward-tracking influence, since in the relevant worlds where some effect doesn't happen, the cause has still occurred by virtue of the stipulation that the worlds be like ours up to the time of the antecedent. But this semantics seems to fail when it comes to simultaneous influence for that very reason: all worlds in which some phenomenon doesn't occur, will be ones in which its simultaneous causes don't occur either.

A similar problem befalls many of the probabilistic-dependency theories of causation. Typically, probabilistic-dependency theories deem a relationship between two relata as causal if the probability of the putative effect occurring given that the putative cause occurs in conjunction with some background conditions is higher, or at least different from, the probability of the first occurring without the cause. Surveying the various accounts, background conditions have been characterised in a couple of ways. One way is Cartwright's (1979), who claimed that they should be understood as the independent causes of the putative effect (see also Eels 1991). But since we're looking for a non-causal analysis of causation, this suggestion must be passed over. The other way, advocated by Reichenbach (1956), Suppes (1970), Kvart (2004), Fenton-Glynn (2011) is that the background conditions are facts about the situation's history before the putative influence occurs. As with the counterfactual semantics just considered, the problem with this is that in the case of purely simultaneous causal influence, such a choice of background conditions is not sufficient to preclude the analysis returning the result that an event is a cause of its simultaneous causes.

Returning to counterfactual dependency, these views might appear untouched by the above concerns if, instead of holding fixed the past by stipulation, they treated counterfactuals (or counterfactual probabilities) as assessed by some other means. Lewis (1986b), for example, suggested that one might avoid a stipulation on historical background by assessing the truth of counterfactuals according to an independent asymmetry of *causal overdetermination*.[9] To motivate this idea, Lewis argued that it is harder to eradicate all the effects of a cause than it is to prevent a cause in the first place. One would have to suppose that *many* of a cause's effects did not occur before being able to infer that a single cause did not happen either. But supposing a single cause didn't happen is typically enough to know that a number of its consequences didn't either.

Such an asymmetry makes no obvious mention of temporal direction and so might be thought, unlike those theories which explicitly condition on historical backgrounds, to accommodate simultaneous causation. Indeed, Lewis spoke openly about desiring to accommodate causal occurrences mediated by tachyons, advanced potentials (in Lewis 1986b), or (more fantastically) an unobservable battle of goblins in causal relations with the co-spatiotemporal presentation he was giving on the idea (see Lewis 1986d).

Nevertheless, I do not think Lewis's idea can bear the weight required of it. Elga (2000) has shown that an asymmetry of overdetermination is not something born out in fundamental physics. Any minimal set of jointly sufficient conditions for a phenomenon would have to include information about the Cauchy surface (in non-relativistic terms, a hypersurface of space

[9]Although Lewis introduces this idea in deterministic terms, he extends the idea to indeterministic worlds (see Lewis 1986c, postscripts). Hence 'overdetermination' should not be thought of as a logical entailment.

at a time) cutting across the entire future or past lightcone of a phenomenon, otherwise there could be some event in the past or future which might ultimately collude to determine that the phenomenon does not occur (or occurs differently). But by including the entire past or future lightcone, we entirely preclude the possibility of an asymmetry of overdetermination in time. In Lewis's terminology, the Cauchy surface which cuts the past lightcone of a phenomenon is no more and no less a 'minimal set of jointly sufficient conditions' for an event than is the Cauchy surface which cuts across the future lightcone (see Frisch 2005 for a more detailed version of this criticism). If counterfactual asymmetries cannot be grounded in temporal overdetermination relatiokns, this implies that the only asymmetries of causal overdetermination supplied by physics that there *could* be must be simultaneous ones. Strangely, Lewis's criterion therefore can seem to be *limited* in the way I suggested powers-based views are, only to an analysis of simultaneous causation.

7.5.3 The nomological analysis and the accommodation criterion

Unlike the previously considered approaches, the nomological analysis of causation as inter-system causation is clearly able to accommodate both purely simultaneous and purely sequential causation. Inter-system causation has been defined as a chain of direct or inertial causal relations and whereas direct Intra-system causation is defined as existing only between simultaneous relata (the cause is entirely simultaneous with the effect), inertial intra-system causation is defined as existing only between non-simultaneous relata (where the cause entirely precedes the effect). Admittedly, I suggested we would have to adjust the criteria for direct intra-system causation to accommodate relativistic effects, but this would not lead to a rejection of simultaneous causation altogether. Co-located value-instances, like those referenced by Huemer and Kovitz (2003) would still be simultaneous under any such adjustment.

The success of the nomological analysis might suggest to powers theorists and dependency theorists alike that incorporating some element of nomological analysis would be beneficial. Indeed, some dispositionalists have openly admitted a place for laws when it comes to analyses of causation (see Cartwright and Pemberton 2013, Heil 2015), although they typically maintain that the metaphysical priority lies with powers suggesting, for instance, that laws derive from facts about powers essences (Bird 2007a) or describe regularities among them (e.g. Tahko 2015, Demarest 2017). However, the preceding points should lead us to question this priority. Not only can laws handle a class of causal relations which dispositional analysis seems to struggle with (viz. sequential causation), but it is consideration of laws themselves which actually lent *credence* to the existence of such causation. Moreover, laws seem to do a lot more than connect dispositions in the way suggested by powers theorists. For instance, some laws (e.g. Newton's first and the law associated with the free-particle solution to the Schrödinger equation) describe what something does when in an inertial state, i.e. when it is exercising no power at all and no dispositions are acting on it. Other laws, e.g. the second law of thermodynamics, seem to concern what happens as a matter of the contingent orientation of its components, i.e. nothing due to the intrinsic dispositional character of the system or its components. Finally, in §6.1, I offered an analysis of dispositions which suggests that when we look to describe the stimulus and manifestation conditions of certain powers or dispositions, we do well to look to the laws. One could, therefore, be forgiven for thinking that it is after all *laws* which are the more basic analytical category, with powers and dispositions being drawn from laws in specific cases.[10]

In the case of dependency views of causation, it is more common to invoke

[10] Although, strictly speaking the analysis provided in §6.1 does not imply a direction of metaphysical priority between powers and laws.

laws explicitly. For instance, Hall (2007) endorses Maudlin's counterfactual semantics which explicitly makes use of laws for his structural equations account of token causation, and both Schaffer (2001) and Paul (2000) introduce the idea of lawful entailment to support their respective probability-raising and counterfactual dependency accounts, etc. Nonetheless, given that we now have a perfectly workable analysis of causation purely in terms of a laws (viz. my above proposal) I see no reason for sticking with a general line of analysis in terms of dependency which tends to accept the importance of laws anyway.

7.6 Summary

In this chapter I have developed nomological analysis of token causation drawing on the causal-junctions account of laws developed throughout in §5 and §6 and the notion of a system that it brings with it. With the added discussion on event-identity, I suggested we can link together such causal relations by identifying the same event in different intra-system causal relations. This gave us the notion of 'inter-system causation', which I suggested generalises over all token causal relations. If the Humean treatment I gave in §6 of robust causal junction laws is successful then the nomological analysis of token causation presented in this chapter also satisfies the Humean methodology.

But we can, I think, go further than this. For in §7.4 and §7.5 I have tried to show how the nomological analysis has a natural ability to succeed where other approaches have failed. The nomological analysis avoids certain formulations of problems of preemption and negative causation which naive dependency and process-based views respectively struggle with, and it accommodates both simultaneous and sequential causation, which dependency and powers-based views respectively struggle with. Admittedly, there are many complex variants of the considered approaches which may fair better than the simplistic ones I have drawn attention to. There are also, presumably, other kinds of approach entirely which I haven't considered. But although these points will not amount to the final word on the matter, I nonetheless believe I've presented enough evidence to support a new and promising analysis of token causation in terms of laws.

8 Conclusion

For a Humean, causal relations and laws of nature are something to be accounted for (assuming they are not to be eliminated) in terms of the mosaic. But beyond that, there is a lot of scope as to how the accounts should proceed. After defending the Humean methodology in §2, the subsequent five chapters of this thesis have argued for a particularly novel view of the relationship between laws and causation.

The first step in §3 was to establish a clear logical form for laws. I argued that all laws are 'system-laws' which have the form of a conditional with an antecedent which specifies a system-type and a consequent which specifies a behavioural formula. This was also the first place where the idea of a law describing an 'embedded phenomenon' was mooted in order to characterise laws' application in idealising contexts. I argued that a modification of idealising laws' behaviour-predicate could serve to reestablish their universality by specifying that the idealising formula is a special case of whichever laws' formula actually applies.

In §4 I defended the view that laws are system-laws against a common trope in philosophical explorations of laws: that laws mirror the structure of causal relations. Laws, I argued, condition behaviour on properties of the very same system at the very time at which the behaviour is predicted. Moreover, those properties on which behaviour is conditioned are conditions of assembly, rather than behavioural properties themselves. Hence, the laws cannot be chained together in the way the causation-mirroring conception suggests. This critique allowed me to also criticise the best systems account of laws, which I argued is implicitly committed to such a conception.

In §5 I set about offering an alternative 'causal-junctions conception' of laws, pointing towards four causal features associated with many laws two of which concern dispositionality, two of which concern causal asymmetry. I pointed to a specific class of laws—the *robust causal junction laws* (RCJLs)—which exhibit all these features in a way which is particularly informative for when it comes to knowing how to intervene and control our environment. I then propose a 'causal-junctions account' of laws which understands RCJLs as that central class of laws from which all others get their lawhood derivatively.

In §6 I provided an analysis of the causal features in RCJLs consistent with the Humean methodology. This involved an extensional analysis of the law-level disposition of embeddability and a schema for extensional analysis (i.e. SysCA) of component-level dispositions in terms of the very laws in which they feature. I then proposed that we analys the asymmetries in causal junctions at both the variable-level and law-level by means of probabilistic analysis in which the effect-variable in each junction is revealed as an unshielded collider on the set of causal paths which comprise the junction. The key consequence of these analyses was that it enabled the 'causal-junctions account' of laws to be proven consistent with the Humean methodology, thereby providing the means for a fully Humean causal-junctions account.

Finally, in §7 I presented a nomological analysis of token causation. The analysis drew on the fact that RCJLs are themselves grounded in type-level causal relations. By showing where the system-types of these laws have instances, we can simultaneously show where their type-level causal relations have instances. Chaining together these instances via system-linking events (which may be identified under different descriptions) gives us the concept of inter-system causation which, I argued, provides a competitive analysis of causation.

Returning once more to the diagram presented in §1, it should be clear now how the connections of analysis have been provided.

One might notice that I've added a further node to the diagram. For although it was not addressed in the foregoing thesis, I think we have reason to suspect the utility of this account of laws and token causation in providing an analysis of type-level causation in general. For example, if it is a type-level causal fact that smoking causes cancer then we might hope to analyse such a fact by recourse to the token instances in which individuals' acts of smoking causes them to contract cancer. These relations in turn will be analysed by an understanding of the intra-system causal relations and event-identities in such circumstances. Intra-system causal relations are ultimately to be explained by RCJLs and their grounding in specific kinds of type-level causation which occur in lawful systems (perhaps at a microscopic cellular or even bio-chemical level). Hence very specific forms of type-level causation which take place within laws' system-types (which I have suggested are analysable in terms of probabilistic dependencies) may ultimately serve to analyse the rest of type-level causation.

Extending the framework to account for all type-level causation is beyond the scope of this thesis, but the possibility of such an account confirms, I think, the comprehensiveness and plausibility of what has already been provided: a Humean account of laws and (token) causation.

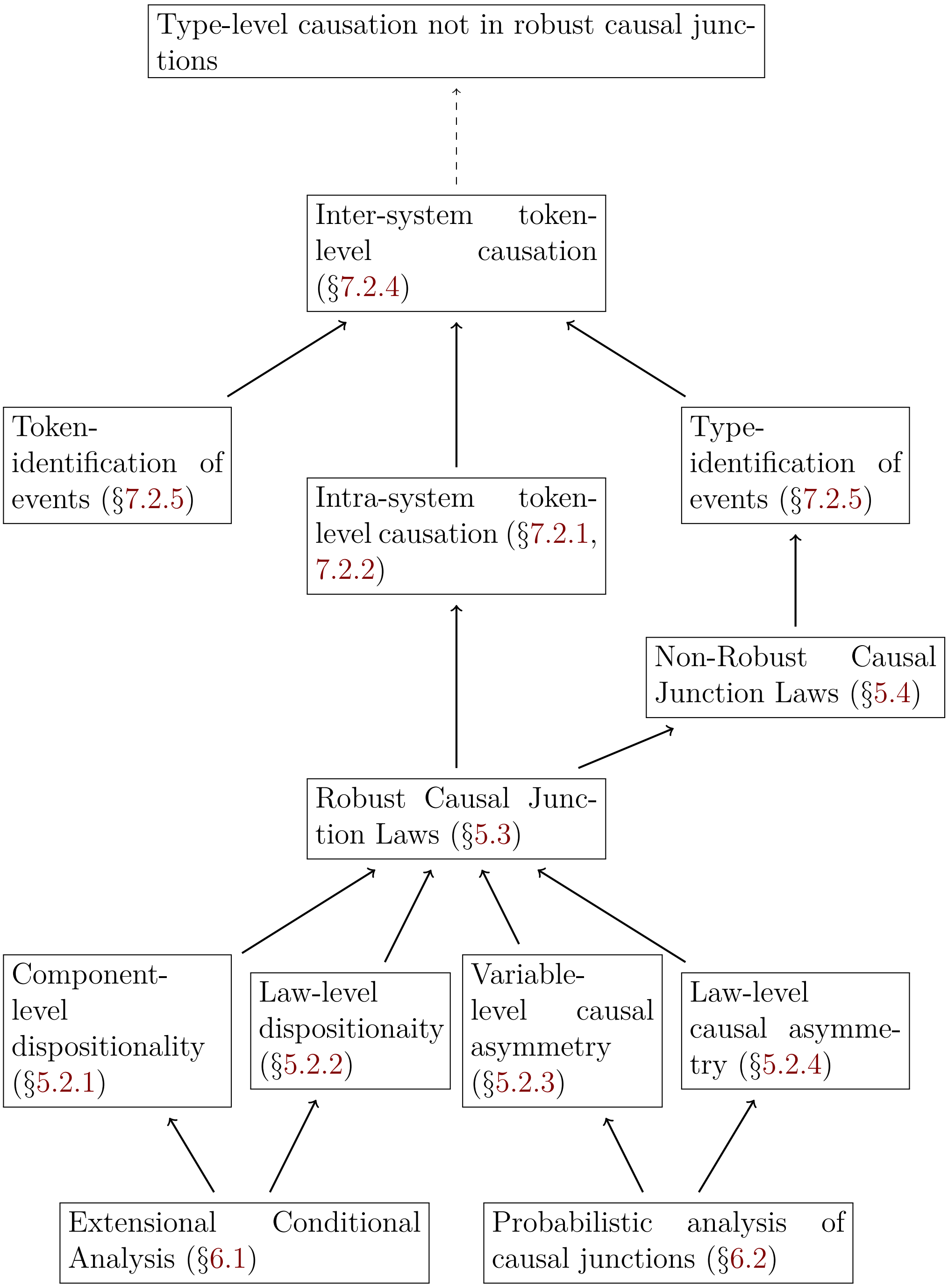

Figure 8.1: Thesis structure from §5 onwards

Bibliography

Akaike, H.: 1973, Information Theory and an Extension of the Maiximum Likelihood Principle, *in* B. N. Petrov and F. Csaki (eds), *2nd International Symposium on Information Theory*, Akademiai Kiado, pp. 267–81.

Albert, D.: 2000, *Time and Chance*, Harvard University Press.

Anscombe, G. E. M.: 1971, *Causality and Determinism*, Cambridge University Press.

Armstrong, D.: 1968, *A Materialist Theory of Mind*, Routledge.

Armstrong, D.: 1983, *What is a Law of Nature?*, Cambridge University Press.

Armstrong, D.: 1993, The Identification Problem and the Inference Problem, *Philosophy and Phenomenological Research* **53**(2), 421–422.

Armstrong, D.: 1997, *A World of States of Affairs*, Cambridge University Press, Cambridge.

Armstrong, D.: 2004a, Going through the Open Door Again, *in* J. Collins, N. Hall and L. Paul (eds), *Causation and Counterfactuals*, MIT Press, chapter 19.

Armstrong, D.: 2004b, How do particulars stand to universals?, *in* D. Zimmerman (ed.), *Oxford Studies in Metaphysics*, Oxford Clarendon Press.

Armstrong, D.: 2005, Reply to Bird, *Analysis* **65**(3), 264–265.

Armstrong, D. and Heathcote, A.: 1991, Causes and Laws, *Noûs* **25**, 63–73.

Arntzenius, F.: 2010, Reichenbach's common cause principle, *in* E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, fall 2010 edn.

Aronson, J.: 1971, The Legacy of Hume's Analysis of Causation, *Studies in History and Philosophy of Science* **2**, 135–156.

Aronson, J.: 1985, Conditions Versus Transference: A Reply to Ehring, *Synthese* **63**, 249–257.

Ayer, A. J.: 1999 [1956], What is a Law of Nature?, *in* M. Tooley (ed.), *Laws of Nature, Causation, and Supervenience.*, Garland Publishing, pp. 1–52.

Bach, K.: 1994, Conversational Impliciture, *Mind and Language* **9**, 124–62.

Barker, S.: 2013, The Emperor's New Metaphysics of Powers, *Mind* **122**(487), 605–653.

Baumgartner, M.: 2008, Regulartity Theories Reassessed, *Philosophia* **36**, 327–354.

Baumgartner, M.: 2013a, A Regularity Theoretic Approach to Actual Causation, *Erkenntnis* **78**, 85–109.

Baumgartner, M.: 2013b, Rendering Interventionism and Non-Reductive

Physicalism Compatible, *Dialectica* **67**(1), 1–27.
Beatty, J.: 1980, What's Wrong with the Received View of Evolutionary Theory?, *Philosophy of Science* **2**, 387–426.
Beatty, J.: 1995, The Evolutionary Contingency Thesis, *in* G. Wolters and G. Lennoz (eds), *Concepts, Theories, and Rationality in the Biological Sciences*, Pittsburgh University Press, pp. 45–81.
Bechtel, W. and Abrahamsen, A.: 2005, Explanation: a mechanist alternative, *Studies in History and Philosophy of Biological and Biomedical Sciences* **36**, 421–441.
Beebee, H.: 2000, The Non-Governing Conception of Laws of Nature, *Philosophy and Phenomenological Research* **61**, 571–594.
Beebee, H.: 2002, Contingent laws rule, *Analysis* **62**, 252–55.
Beebee, H.: 2004, Causation and Nothingness, *in* J. Collins, N. Hall and L. Paul (eds), *Causation and Counterfactuals*, MIT Press, chapter 11.
Beebee, H.: 2006a, Does Anything Hold the Universe Together?, *Synthese* **149**, 509–533.
Beebee, H.: 2006b, *Hume on Causation*, Routledge.
Beebee, H.: 2011, Necessary Connections and the Problem of Induction, *Noûs* **45**(3), 504–527.
Bird, A.: 1998, Dispositions and Antidotes, *The Philosophical Quaterly* **48**, 227–234.
Bird, A.: 2001, Necessarily, salt dissolves in water, *Analysis* **61**(267-74).
Bird, A.: 2002, On whether some laws are necessary, *Analysis* **62**(3), 257–70.
Bird, A.: 2007a, *Nature's Metaphysics: Laws and Properties*, Oxford Clarendon Press.
Bird, A.: 2007b, The regress of pure powers?, *Philosophical Quarterly* **57**(229), 513–534.
Bird, A.: 2010, Causation as the Manifestation of Powers, *in* A. Marmodoro (ed.), *The Metaphysics of Powers: their Grounding and their Manifestations*, Reidel.
Black, M.: 1954, Metaphor, *Proceedings of the Aristotelian Society* **55**, 273–294.
Black, R.: 2000, Against quidditism, *Australian Journal of Philosophy* **78**(1), 87–104.
Bontly, T.: 2006, What is an empirical analysis of causation?, *Synthese* **151**(2), 177–200.
Braddon-Mitchell, D.: 2001, Lossy Laws, *Noûs* **35**(2), 260–277.
Bradley, D.: 2015, *A Critical Introduction to Formal Epistemology*, Bloomsbury Academic.
Briggs, R.: 2017, Normative theories of rational choice: Expected utility, *in* E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, spring 2017 edn, Metaphysics Research Lab, Stanford University.
Bromberger, S.: 1966, Why Questions, *in* R. G. Colodney (ed.), *Mind and Cosmos*, University of Pittsburgh Press.
Carey, S.: 2009, *The Origin of Concepts*, Oxford University Press.
Carnap, R.: 1928, *The Logical Construction of the World and Pseudoproblems*

in Philosophy, Open Court Publishing.

Carnap, R.: 1936, Testability and Meaning, *Philosophy of Science* **3;4**, 419–471;1–40.

Carroll, J.: 1994, *Laws of Nature*, Cambridge University Press.

Carroll, J.: 2012, Laws of Nature, *in* E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*.
URL: *http://plato.stanford.edu/archives/spr2012/entries/laws-of-nature/*

Carroll, J. W.: 2016, Laws of nature, *in* E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, fall 2016 edn, Metaphysics Research Lab, Stanford University.

Cartwright, N.: 1979, Causal Laws and Effective Strategies, *Noûs* **13**, 419–437.

Cartwright, N.: 1983, *How the Laws of Physics Lie*, Oxford University Press, Oxford.

Cartwright, N.: 1989, *Nature's Capacities and their Measurement*, Oxford University Press, Oxford.

Cartwright, N.: 1999, *The Dappled World: A Study of the Boundaries of Science*, Cambridge University Press.

Cartwright, N.: 2000, Against the completability of science, *in* M. W. F. Stone and J. Wolff (eds), *The Proper Ambition of Science*, Routledge, chapter 11, pp. 209–222.

Cartwright, N.: 2007, *Hunting Causes and Using Them: Approaches in Philosophy and Economics*, Cambridge University Press.

Cartwright, N.: 2017, Causal Powers: Why Humeans can't even be instrumentalists, *in* J. Jacobs (ed.), *Causal Powers*, Oxford University Press, chapter 2.

Cartwright, N. and Pemberton, J.: 2013, Aristotelian powers: Without them, what would modern science do?, *in* R. Groff and J. Greco (eds), *Powers and capacities in philosophy: the new Aristotelianism*, Routledge, chapter 5, pp. 93–112.

Chakravartty, A.: 2007, *A Metaphysics for Scientific Realism: Knowing the Unobservable*, Cambridge University Press.

Choi, S.: 2002, Causation and Gerrymandered World Lines: A Critique of Salmon, *Philosophy of Science* **69**, 105–117.

Choi, S.: 2008a, Dipsositional Properties and Counterfactual Conditionals, *Mind* **117**, 795–841.

Choi, S.: 2008b, Dispositional Properties and Counterfactual Conditionals, *Mind* **117**, 795–841.

Choi, S. and Fara, M.: 2016, Dispositions, *in* E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, spring 2016 edn, Metaphysics Research Lab, Stanford University.

Cohen, J. and Callender, C.: 2009, A better best systems account of lawhood, *Philosophical Studies* **145**, 1–34.

Corry, R.: 2009, How is Scientific Analysis Possible?, *in* T. Handfield (ed.), *Dispositions and Causes*, Oxford: Clarendon Press, chapter 6.

Craver, C.: 2007, *Explaining the Brain: mechanisms and the mosaic unity of neuroscience*, Oxford Clarendon Press.

Creary, L.: 1981, Causal Explantation and the Reality of Natural Component

Forces, *Pacific Philosophical Quarterly* **62**, 148–157.
da Costa, N. and French, S.: 1998, Models, Theories, and Structures: Thirty Years on, *Philosophy of Science* **67, Supplement**(S116-S127).
Daly, C.: 2005, So where's the explanation?, *in* H. Beebee and J. Dodd (eds), *Truthmakers: The Contemporary Debate*, Oxford Clarendon Press, chapter 6, pp. 85–104.
Davidson, D.: 1967, Causal Relations, *Journal of Philosophy* **64**(21), 691–703.
Davidson, D.: 1970, Mental Events, *in* L. Foster and J. W. Swanson (eds), *Experience and Theory*, Humanities Press, chapter 20.
Davidson, D.: 1976, The Individuation of Events, *in* M. Brand and D. Walton (eds), *Action Theory*, Reidel, chapter 22.
Davidson, D.: 1985, Reply to Quine on Events, *in* E. Lepore and B. McLaughlin (eds), *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, Blackwell.
Davidson, D.: 1995, Laws and Cause, *Dialectica* (263-279).
De Clercq, R., Lam, W. and Shang, J.: 2014, Is there a problem with the causal criterion of event identity?, *American Philosophical Quaterly* **51**(2), 109–120.
de Finetti, B.: 1937, La Prévision: Ses Lois Logiques, Ses Sources Subjectives, *Annales de l'Institut Henri Poincaré* **7**, 1–68.
Demarest, H.: 2017, Powerful properties, powerless laws, *in* J. Jacobs (ed.), *Causal Powers*, Oxford University Press, chapter 4.
Devitt, M.: 2009, "Ostrich Nominalism" or "Mirage Realism"?", *Putting Metaphysics First: Essays on Metaphysics and Epistemology*, Oxford University Press, pp. 13–19.
Dipert, R.: 1997, The Mathematical Structure of the World: The World as a Graph, *Journal of Philosophy* **94**(7), 329–358.
Dowe, P.: 1995, Causality and Conserved Quantities: A Reply to Salmon, *Philosophy of Science* **62**, 321–333.
Dowe, P.: 2000, *Physical Causation*, Cambridge University Press, New York.
Dray, W.: 1957, *Laws and Explanation in History*, Oxford University Press.
Dretske, F.: 1977, Laws of Nature, *Philosophy of Science* **44**, 248–68.
Drewery, A.: 2006, The logical form of universal generalizations, *Australian Journal of Philosophy* **83**(3), 373–393.
Dummett, M.: 1978, The Social Character of Meaning, *Truth and Other Enigmas*, Harvard University Press, chapter 23, pp. 420–430.
Dupré, J.: 1993, *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*, Harvard University Press.
Earman, J.: 1986, *A Primer on Determinism*, D. Reidel Publishing Company.
Earman, J.: 2011, Sharpening the Electromagnetic Arrow(s) of Time, *in* C. Callender (ed.), *The Oxford Handbook of Philosophy of Time*, Oxford University Press.
Earman, J. and Roberts, J.: 1999, There Is No Problem of Provisos, *Synthese* **118**(3), 439–478.
Easwaran, K.: 2014, Why Physics Uses Second Derivatives, *British Journal of Philosophy of Science* **65**, 845–862.
Eberhardt, F. and Scheines, R.: 2007, Interventions and Causal Inference,

Philosophy of Science **74**, 981–995.

Edgington, D.: 2011, Causation First: Why Causation is Prior to Counterfactuals, *in* C. Hoerl, T. McCormack and S. R. Beck (eds), *Understanding Counterfactuals, Understanding Causation: Issues in Philosophy and Psychology*, Oxford University Press.

Eels, E.: 1991, *Probabilistic Causality*, Cambridge University Press, MA.

Ehring, D.: 1997, *Causation and Persistence*, Oxford University Press, New York.

Einstein, A.: 1934, On the Method of Theoretical Physics, *Philosophy of Science* **1**, 163–9.

Einstein, A. and Infeld, L.: 1971, *The Evolution of Physics*, Cambridge University Press.

Elby, A.: 1992, Should we explain the EPR correlations causally?, *Philosophy of Science* **59**, 16–25.

Elga, A.: 2000, Statistical Mechanics and the Asymmetry of Counterfactual Dependence, *Philosophy of Science* **68**(3), S313–S324.

Elgin, C.: 2004, True enough, *Philosophical Issues* **14**, 113–131.

Ellis, B.: 2002, *The Philosophy of Nature: A Guide to the New Essentialism*, Acumen.

Evans, G.: 1982, *The Varieties of Reference*, Oxford University Press.

F., M.: 2005, Lewis's Animadversions on the Truthmaker Principle, *Truthmakers: The Contemporary Debate*, Oxford University Press, pp. 117–40.

Fair, D.: 1979, Causation and the Flow of Energy, *Erkenntnis* **14**, 219–250.

Fales, E.: 1990, *Causation and Universals*, Routledge, London.

Fechner, G. T.: 1864, *Ueber die Physikalische und philosophische Atomenlehre*, H. Mendelssohn.

Fenton-Glynn, L.: 2011, A Probabilistic Analysis of Causation, *British Journal of Philosophy of Science* **62**, 343–392.

Fenton-Glynn, L.: 2016, A Proposed Probabilistic Extension of the Halpern and Pearl Definition of 'Actual Cause', *British Journal of Philosophy of Science* p. DOI: 10.1093/bjps/axv056.

Fenton-Glynn, L.: 2017, Imprecise Best System Chances, *EPSA15 Selected Papers*, Springer, chapter 24, pp. 297–308.

Feynman, R.: 2011, Electromagnetism, *Feynman Lectures in Physics (Millenium Edition)*, Vol. 2, Basic Books, chapter 1.

Fine, K.: 1994, Essence and Modality, *in* J. Tomberlin (ed.), *Philosophical Perspectives*, Vol. 8, Ridgeview Publishing Company, pp. 1–16.

Fine, K.: 2002, Varieties of Necessity, *in* S. G. Szabo and J. Hawthorne (eds), *Conceivability and Possibility*, Oxford University Press, pp. 253–281.

Fine, K.: 2012, Guide to Ground, *in* F. Correia and B. Schneider (eds), *Metaphysical grounding: Understanding the structure of reality*, Cambridge University Press, pp. 37–80.

Fodor, J.: 1974, Special Sciences (Or: The Disunity of Science as a Working Hypothesis), *Synthese* **28**(2), 97–115.

Fodor, J.: 1991, You Can Fool Some of The People All of The Time, Everything, Else Being Equal: Hedged Laws and Psychological Explanations, *Mind* **100**(1), 19–34.

Frege, G.: 1950, *The Foundations of Arithmetic*, Blackwell.

Friend, T.: 2016, Laws are conditionals, *European Journal for Philosophy of Science* **6**, 123–144.

Frigg, R.: 2010, Models and fiction, *Synthese* **172**, 251–268.

Frisch, M.: 2005, *Inconsistency, Asymmetry and Non-locality: A Philosophical Investigation of Classical Electrodynamics*, Oxford University Press.

Frisch, M.: 2014, *Causal Reasoning in Physics*, Cambridge University Press.

Gettier: 1963, Is Justified True Belief Knowledge?, *Analysis* **23**(6), 121–123.

Giere, R.: 1988, Laws, Theories, and Generalizations, *in* Grünbaum and W. Salmon (eds), *The Limits of Deductivism*, University of California Press, pp. 37–46.

Giere, R.: 1995, Skeptical Perspective: Science without Laws, *in* F. Weinert (ed.), *Laws of Nature: Essays on the Philosophical, Scientific and Historical Dimensions*, Walter de Gruyter & Co., pp. 120–138.

Giere, R.: 2010, An agent-based conception of models and scientific representation, *Synthese* **172**, 269–281.

Glennan, S.: 1996, Mechanisms and the nature of causation, *Erkenntnis* **44**, 49–71.

Glynn, L.: 2015, Imprecise chance and the best system account.

Goldman, A.: 1967, A causal theory of knowing, *Journal of Philosophy* **64**(12), 357–372.

Goodman, N.: 1947, The Problem of Counterfactual Conditionals, *Journal of Philosophy* **44**, 113–28.

Goodman, N.: 1954 [1983], *Fact, Fiction and Forecast*, Harvard University Press.

Goodman, N.: 1976, *Languages of Art*, Hackett Publishing.

Gorman, M.: 2005, The essential and the accidental, *Ratio* **18**(3), 276–289.

Grasshoff, G. and May, M.: 2001, Causal Regularities, *in* W. Spohn, M. Ledwig and M. Esfeld (eds), *Current Issues in Causation*, Mentis, pp. 85–114.

Hájek, A.: 1997, "Mises redux"—redux: fifteen arguments against finite frequentism, *Erkenntnis* **45**, 209–227.

Hájek, A.: 2012, Interpretations of probability, *in* E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, winter 2012 edn, Metaphysics Research Lab, Stanford University.

Hall, N.: 2004a, Causation and the Price of Transitivity, *in* J. Collins, N. Hall and L. Paul (eds), *Causation and Counterfactuals*, MIT Press, chapter 4.

Hall, N.: 2004b, Two Concepts of Causation, *in* J. Collins, N. Hall and L. Paul (eds), *Causation and Counterfactuals*, MIT Press, chapter 10.

Hall, N.: 2004c, Two Mistakes about Credence and Chance, *Australian Journal of Philosophy* **82**, 93–111.

Hall, N.: 2007, Structural Equations and Causation, *Philosophical Studies* **132**, 109–136.

Hall, N.: 2015, Humean Reductionism About Laws Of Nature, *in* B. Loewer and J. Schaffer (eds), *A Companion to David Lewis*, Wiley, chapter 17.

Halpern, J. and Pearl, J.: 2001, Causes and Explanations: A Structural-Model

Approach—Part I: Causes, *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence* pp. 194–202.

Hanson, N. R.: 1958, *Patterns of Discovery: an inquiry into the conceptual foundations of science*, Cambridge University Press.

Harré, R.: 1993, *Laws of Nature*, Duckworth.

Hausman, D.: 1998, *Causal Asymmetries*, Cambridge University Press.

Heil, J.: 2003, *From an Ontological Point of View*, Oxford Clarendon Press.

Heil, J.: 2005, Dispositions, *Synthese* **144**(3), 343–356.

Heil, J.: 2012, Causal Relations, *in* A. Marmodoro and D. Yates (eds), *Metaphysics of Relations*, chapter 8, pp. 127–137.

Heil, J.: 2015, Ordinary Objects, talk given at the Power Structuralism project, Corpus Christi, Oxford University.

Helmholtz, H.: 1896, Über das Ziel und die Fortschritte der Naturwissenschaft, *Eröffnungsrede für die Naturforscherversammlung zu Innsbruck* .

Hempel, C.: 1945, Studies in the Logic of Confirmation I, *Mind* **54**(213), 1–26.

Hempel, C.: 1965, *Aspects of Scientific Explanation*, Free Press, New York.

Hempel, C. and Oppenheim, P.: 1948, Studies in the Logic of Explanation, *Philosophy of Science* **15**, 135–175.

Hendry, R. and Rowbottom, D.: 2009, Dispositional essentialism and the necessity of laws, *Analysis* **69**(4), 668–677.

Hesse, M.: 1963, *Models and Analogies in Science*, Sheed and Ward.

Hicks, M. T. and van Elswyk, P.: 2015, Humean laws and circular explanation, *Philosophical Studies* **172**(2), 433–443.

Hitchcock, C.: 1995, Discussion: Salmon on Explanatory Relevance, *Philosophy of Science* **62**, 304–320.

Hitchcock, C.: 2001a, A Tale of Two Effects, *Philosophical Review* **110**(3), 361–396.

Hitchcock, C.: 2001b, The Intransitivity of Causation Revealed in Equations and Graphs, *Journal of Philosophy* **98**, 273–299.

Hitchcock, C.: 2007a, What Russell Got Right, *in* H. Price and R. Corry (eds), *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*, Oxford University Press.

Hitchcock, C.: 2007b, What's Wrong with Neuron Diagrams?, *in* J. Campbell, M. O'Rourke and S. H. (eds), *Causation and Explanation*, MIT Press, chapter 4, pp. 63–92.

Hitchcock, C.: 2016, Probabilistic causation, *in* E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, winter 2016 edn, Metaphysics Research Lab, Stanford University.

Hitchcock, C. and Knobe, J.: 2009, Cause and norm, *Journal of Philosophy* **11**, 587–612.

Horsten, L.: 2010, Impredicative Identity Criteria, *Philosophy and Phenomenological Research* **80**(2), 411–439.

Huemer, M. and Kovitz, B.: 2003, Causation as Simultaneous and Continuous, *The Philosophical Quaterly* **53**(213), 556–565.

Hulswit, M.: 2002, *From Cause to Causaton. A Piercean Perspective*, Kluwer Academic Publishers.

Hume, D.: 1738 [1978], *A Treatise of Human Nature*, Oxford Clarendon Press.
Hume, D.: 1777 [1993], *An Enquiry Concerning Human Understanding*, 2nd edn, Hackett Publishing, Indianapolis.
Humphreys, R. F. and Beringer, R.: 1950, *First Principles of Atomic Physics*, Harper.
Hüttemann, A.: 2004, *What's Wrong with Microphysicalism?*, Routledge, London.
Hüttemann, A.: 2007, Naturgestze, *in* A. Bartels and M. Stökler (eds), *Wissenschaftstheorie*, pp. 135–153.
Hüttemann, A.: 2014, Ceteris Paribus Laws in Physics, *Erkenntnis* **79**, 1715–1728.
Illari, P. and Williamson, J.: 2013, In Defence of Activities, *Journal for General Philosophy of Science* **44**(1), 69–83.
Ismael, J.: 2008, Raid! Dissolving the Big, Bad Bug, *Noûs* **42**(2), 292–307.
Jackson, F.: 1977, A Causal Theory of Counterfactuals, *Australian Journal of Philosophy* **55**, 3–21.
Jeffrey, R.: 1971, Statistical Explanation vs. Statistical Inference, *in* W. Salmon, R. Jeffrey and J. Greeno (eds), *Statistical Explanation and Statistical Relevance*, University of Pittsburgh, pp. 19–29.
Johnston, M.: 1992, How to Speak of the Colors, *Philosophical Studies* **68**, 221–263.
Kant, I.: 1998, *Critique of Pure Reason*, Cambridge University Press, New York.
Keinänen, M.: 2011, Tropes—The Basic Constituents of Powerful Particulars?, *Dialectica* **65**(419-50).
Kim, J.: 1976, Events as Property Exemplifications, *in* M. Brand and D. Walton (eds), *Action Theory*, Reidel, chapter 23.
Kirchhoff, G. R.: 1876, *Vorlesungen über mathematische Physik: Mechanik. I*, B. G. Teubner.
Kneale, W.: 1950, Natural Laws and Contrary-to-Fact Conditionals, *Analysis* **10**(6), 121–125.
Kolin, A.: 1950, *Physics: its laws, ideas and methods*, McGraw-Hill.
Kripke, S.: 1980, *Naming and Necessity*, Harvard University Press.
Kvart, I.: 1991, Transitivity and Preemption of Causal Impact, *Philosophical Studies* **64**, 125–160.
Kvart, I.: 2004, Causation: Probabilistic and Counterfactual Analyses, *in* J. Collins, N. Hall and L. Paul (eds), *Causation and Counterfactuals*, MIT Press, chapter 15.
Lange, M.: 1993a, Natural Laws and the Problem of Provisos, *Erkenntnis* **38**, 233–248.
Lange, M.: 1993b, When Would the Natural Laws Have Been Broken?, *Analysis* **53**(4), 262–9.
Lange, M.: 2000, *Natural Laws in Scientific Practice*, Oxford University Press.
Lange, M.: 2004, A Note on Scientific Essentialism, Laws of Nature, and Counterfactuals, *Australian Journal of Philosophy* **82**(2), 227–241.
Lange, M.: 2008, Could the Laws of Nature Change?, *Philosophy of Science*

75, 69–92.

Lange, M.: 2009, *Laws and Lawmakers: Science, Metaphysics, and the laws of Nature*, Oxford University Press, Oxford.

Lange, M.: 2013, Grounding, scientific explanation, and Humean laws, *Philosophical Studies* **164**, 255–261.

Lange, M.: forthcoming, Transitivity, self-explanation, and the explanatory circularity argument against Humean accounts of natural law, *Synthese* pp. 1–17.

Le Roy, E.: 1901, Un Positivisme Nouveau, *Revue de Métaphysique et de Morale* **9**(2), 138–153.

Leslie, A. and Keeble, S.: 1987, Do Six-Month-Old Infants Perceive Causality?, *Cognition* **25**, 265–88.

Lewis, D.: 1973a, Causation, *Journal of Philosophy* **20**, 11–13.

Lewis, D.: 1973b, *Counterfactuals*, Blackwell Publishing.

Lewis, D.: 1983, New Work for a Theory of Universals, *Australian Journal of Philosophy* **61**(4), 343–377.

Lewis, D.: 1984, Putnam's Paradox, *Australian Journal of Philosophy* **62**, 221–236.

Lewis, D.: 1986a, Causation: postscripts, *Philosophical Papers II*, Oxford University Press, Oxford, chapter 21.

Lewis, D.: 1986b, Counterfactual Dependence and Time's Arrow, *Philosophical Papers II*, Oxford University Press, Oxford, chapter 17.

Lewis, D.: 1986c, Counterfactual Dependence and Time's Arrow: postscripts, *Philosophical Papers II*, Oxford University Press, Oxford, chapter 17.

Lewis, D.: 1986d, Events, *Philosophical Papers II*, Oxford University Press, chapter 23.

Lewis, D.: 1986e, *Introduction*, Oxford University Press, Oxford.

Lewis, D.: 1986f, *On the Plurality of Worlds*, Blackwell.

Lewis, D.: 1986g, *Philosophical Papers II*, Oxford University Press, Oxford.

Lewis, D.: 1994, Humean Supervenience Debugged, *Mind* **103**, 473–490.

Lewis, D.: 1997, Finkish Dispositions, *The Philosophical Quaterly* **47**(187), 143–158.

Lewis, D.: 2000, Causation as Influence, *in* J. Collins, N. Hall and L. Paul (eds), *Causation and Counterfactuals*, Vol. 97, MIT Press, chapter 3, pp. 182–197.

Lewis, D.: 2004, Void and Object, *Causation and Counterfactuals*, MIT Press, chapter 10.

Loewer, B.: 1996, Humean Supervenience, *Philosophical Topics* **24**, 101–127.

Loewer, B.: 2012, Two accounts of laws and time, *Philosophical Studies* **160**(1), 115–137.

Lowe, E. J.: 1989, *Kinds of Being*, Oxford University Press.

Lowe, E. J.: 2006, *The Four-Category Ontology: A Metaphysical Foundation for Natural Science*, Oxford University Press.

Lowe, E. J.: 2009, *More Kinds of Being: A Further Study of Individuation, Identity, and the Logic of Sortal Terms*, Wiley-Blackwell.

Lowe, E. J.: 2012, There are (Probably) No Relations, *in* A. Marmodoro and

D. Yates (eds), *Metaphysics of Relation*, Oxford University Press, chapter 6, pp. 100–112.

Mach, E.: 1900, *Die Principien der Wärmelehre. Historisch-kritisch entwickelt*, Barth.

Mach, E.: 1905, *Erkenntnis und Irrtum. Skizzen zur Psychologie der Forschung*, Barth.

Machamer, P., Darden, L. and Craver, C.: 2000, Thinking about Mechanisms, *Philosophy of Science* **67**, 1–25.

Mackie, J.: 1973, *Truth, Probability and Paradox*, Oxford Clarendon Press.

Mackie, J.: 1980, *The Cement of the Universe*, Oxford Clarendon Press.

Malthus, T. R.: 1798, *An essay on the principle of population, as it affects the future improvement of society with remarks on the speculations of Mr. Godwin, M. Condorcet and other writers*, London: Printed for J. Johnson.

Malzkorn, W.: 2000, Realism, Functionalism and the Conditional Analysis of Dispositions, *The Philosophical Quaterly* **50**(201), 452–469.

Manley, D. and Wasserman, R.: 2007, A gradable approach to dispositions, *The Philosophical Quaterly* **57**(226), 68–75.

Manley, D. and Wasserman, R.: 2008, On Linking Dispositions and Conditionals, *Mind* **117**(465), 60–84.

Manley, D. and Wasserman, R.: 2011, Dispositions, Conditionals, and Counterexamples, *Mind* **120**(480), 1191–1227.

Marshall, D.: 2015, Humean laws and explanation, *Philosophical Studies* **172**, 3145–3165.

Martin, C. B.: 1994, Dispositions and Conditionals, *The Philosophical Quaterly* **44**, 1–8.

Martin, C. B.: 2008, *The Mind in Nature*, Oxford Clarendon Press.

Maudlin, T.: 2004, Causation, counterfactuals and the third factor, *in* J. Collins, N. Hall and L. Paul (eds), *Causation and Counterfactuals*, MIT Press, chapter 18.

Maudlin, T.: 2007, *The Metaphysics within Physics*, Oxford University Press, New York.

McDermott, M.: 1995, Redundant Causation, *British Journal of Philosophy of Science* **46**, 523–544.

McDowell, J.: 1998, On the Sense and Reference of a Proper Name, *Meaning, Knowledge and Reality*, Harvard University Press, Cambridge, MA.

McKenzie, K.: 2013, How (Not) To Be a Humean Structuralist, *in* V. Karakostas and D. Dieks (eds), *EPSA11 Perspectives and Foundational Problems in Philosophy of Science*, Vol. 2, Springer, pp. 307–318.

McKitrick, J.: 2005, Are Dispositions Causally Relevant?, *Synthese* **144**, 357–371.

Mellor, D. H.: 1995, *The Facts of Causation*, Routledge.

Mellor, D. H.: 2004, For Facts as Causes and Effects, *in* J. Collins, N. Hall and L. Paul (eds), *Causation and Counterfactuals*, MIT Press, chapter 12.

Menzies, P.: 1989, Probabilistic Causation and Causal Processes: A Critique of Lewis, *Philosophy of Science* **56**, 642–663.

Menzies, P.: 1996, Probabilistic Causation and the Pre-emption Problem, *Mind* **105**, 85–117.

Menzies, P.: 1998, How Justified are the Humean Doubts about Intrinsic Causal Links?, *Communication & Cognition* **31**, 339–364.

Michotte, A.: 1963, *The Perception of Causality*, Methuen, London.

Miller, E.: 2015, Humean scientific explanation, *Philosophical Studies* **172**(5), 1311–1332.

Mitchell, S.: 2000, Dimensions of Scientific Law, *Philosophy of Science* **67**(2), 242–265.

Molnar, G.: 2003, *Powers: A Study in Metaphysics*, Oxford University Press, Oxford.

Morgan, M. and Morrison, M.: 1999, Introduction, *Models as Mediators*, Cambridge University Press, chapter 1.

Mumford, S.: 1998, *Dispositions*, Oxford University Press, Oxford.

Mumford, S.: 2004, *Laws in Nature*, Routledge.

Mumford, S. and Anjum, R. L.: 2011, *Getting Causes from Powers*, Oxford University Press.

Needham, J.: 1951, *The Grand Titration: Science and Society in East and West*, George Allen and Unwin Ltd, London.

Nickel, B.: 2010, Ceteris Paribus Laws: Generics and Natural Kinds, *Philosophers' Imprint* **10**(6).

Norton, J.: 2009, Is There an Independent Principle of Causality in Physics?, *British Journal of Philosophy of Science* **60**(3), 475–86.

Nozick, R.: 1981, *Philosophical Explanations*, Harvard University Press.

Papineau, D.: 1992, Reliabilism, Induction and Scepticism, *The Philosophical Quaterly* **42**(166), 1–20.

Papineau, D.: 2001, The rise of physicalism, *in* C. Gillet and B. Loewer (eds), *Physicalism and its Discontents*, Cambridge University Press.

Paul, L.: 2000, Aspect Causation, *Journal of Philosophy* **97**, 235–256.

Pearl, J.: 1988, *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann.

Pearl, J.: 2000, *Causality*, Cambridge University Press, New York.

Pereyra, G. R.: 2002, *Resemblance Nominalism: A Solution to the Problem of Universals*, Oxford Clarendon Press, Oxford.

Pettigrew, R.: 2012, Accuracy, Chance, and the Principal Principle, *Philosophical Review* **121**(2), 241–275.

Pettigrew, R.: 2013, A New Epistemic Utilitity Argument for the Principal Principle, *Episteme* **10**, 19–35.

Pietroski, P. and Rey, R.: 1995, When Other Things aren't Equal: Saving Ceteris Paribus Laws from Vacuity, *British Journal of Philosophy of Science* **46**, 81–110.

Planck, M.: 1937, *Vom Wesen der Willensfreiheit*, Leibzig: Barth.

Poincaré, H.: 1905, *Science and Hypothesis*, Walter Scott Publishing.

Price, H. and Weslake, B.: 2009, The Time-Asymmetry of Causation, *in* H. Beebee, P. Menzies and C. Hitchcock (eds), *The Oxford Handbook of Causation*, Oxford University Press, pp. 414–43.

Psillos, S.: 2002, Salt does dissolve in water, but not necessarily, *Analysis*

62, 255–57.

Putnam, H.: 1977, Realism and Reason, *Proceedings and Addresses of the American Philosophical Association* **50**, 483–498.

Putnam, H.: 1980, Models and Reality, *The Journal of Symbolic Logic* **45**, 464–482.

Quine, W. V. O.: 1948, On what there is, *Review of Metapysics* **2**(1), 21–38.

Quine, W. V. O.: 1973, *The Roots of Reference*, Open Court Publishing, Chicago.

Quine, W. V. O.: 1985, Events and Reification, *in* E. Lepore and B. McLaughlin (eds), *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, Blackwell, pp. 162–171.

Ramsey, F.: 1950 (2^{nd} edition), *Foundations of Mathematics*, Routledge and Kegan Paul.

Ramsey, F.: 1978 [1928], Universals of Law and Fact, *in* D. H. Mellor (ed.), *Foundations: Essays in Philosophy, Logic, Mathematics and Economics*, Routledge and Kegan Paul.

Reichenbach, H.: 1956, *The Direction of Time*, University of California Press.

Reichenbach, H.: 1958, *The Philosophy of Space and Time*, Dover, New York.

Reutlinger, A.: 2011, A Theory of Non-Universal Laws, *International Studies in the Philosophy of Science* **25**, 97–117.

Roberts, J.: 2008, *The Law-Governed Universe*, Oxford University Press.

Rodriguez-Pereyra, G.: 2005, Why Truthmakers?, *in* H. Beebee and J. Dodd (eds), *Truthmakers: The Contemporary Debate*, Oxford Clarendon Press, chapter 2, pp. 17–32.

Rosenberg, A.: 2001, How is biological explanation possible?, *British Journal of Philosophy of Science* **52**, 735–760.

Rueger, A.: 2006, Connection and Influence: A Process theory of Causation, *Journal for General Philosophy of Science* **37**, 77–97.

Russell, B.: 1912, On the Notion of a Cause, *Proceedings of the Aristotelian Society* **13**, 1–26.

Salmon, W.: 1977, An 'At-At' Theory of Causation, *Philosophy of Science* **44**, 215–224.

Salmon, W.: 1984a, *Scientific Explanation and the Causal Structure of the World*, Princeton University Press.

Salmon, W.: 1984b, Scientific Explanation: Three basic conceptions, *Proceedings of the Biennial Meeting of the Philosophy of Science Association* **2**, 293–305.

Salmon, W.: 1994, Causation without counterfactuals, *Philosophy of Science* **61**, 297–312.

Salmon, W.: 1998a, A New Look at Causality, *Causation and Explanation*, Oxford University Press, New York, chapter 1.

Salmon, W.: 1998b, *Causation and Explanation*, Oxford University Press, Oxford.

Salmon, W.: 1998c, Causation without Counterfactuals, *Causation and Explanation*, Oxford University Press.

Sánchez-Dorado, J.: 2017, Methodological lessons for the integration of philosophy of science and aesthetics: The case of representation, *in* O. Bueno, G. Darby, S. French and D. Rickles (eds), *Thinking about Science, Reflecting on Art*, Routledge, chapter 1.

Schaffer, J.: 2000, Causation by disconnection, *Philosophy of Science* **67**(2), 285–300.

Schaffer, J.: 2001, Causes as probability raisers of processes, *Journal of Philosophy* **98**, 75–92.

Schaffer, J.: 2003, Causes Need Not Be Physically Connected to Their Effects: The Case for Negative Causation, *in* C. Hitchcock (ed.), *Contemporary Debates in Philosophy of Science*, Basil Blackwell, London, chapter 10.

Schaffer, J.: 2005a, Contrastive Causation, *Philosophical Review* **114**, 297–238.

Schaffer, J.: 2005b, Quidditistic Knowledge, *Philosophical Studies* **123**, 1–32.

Schaffer, J.: 2012, Causal Contextualisms, *in* M. Blaauw (ed.), *Contrastivism in Philosophy: New Perspectives*, Routledge, New York.

Schlottman, A. and Shanks, D.: 1992, Evidence for a Distinction between Judged and Perceived Causality, *Quarterly Journal of Experimental Psychology* **44A**, 321–342.

Schrenk, M.: 2007, *The metaphysics of ceteris paribus laws*, Ontos, Frankfurt.

Schrenk, M.: 2014, Better Best Systems and the Issue of CP-Laws, *Erkenntnis* **79**, 1787–1799.

Schrödinger, E.: 1951, *Science and Humanism*, Cambridge University Press.

Schurz, G.: 2002, Ceteris Paribus Laws: Classification and Deconstruction, *Erkenntnis* **57**(3), 351–372.

Schwarz, W.: forthcoming, Proving the Principal Principle, *in* A. Wilson (ed.), *New Essays on Time and Chance*, Oxford University Press.

Sellars, W.: 1948, Concepts as Involving Laws and Inconceivable Without Them, *Philosophy of Science* **15**(4), 287–315.

Serway, R. and Jewett, J.: 2006, *Principles of Physics: A Calculus-Based Text*, 4^{th} edn, Thomson Brooks/Cole.

Shoemaker, S.: 1997, Causality and Properties, *in* D. H. Mellor and A. Oliver (eds), *Properties*, Oxford University Press, chapter 15.

Simon, H.: 1953, Causal ordering and identifiability, *in* W. C. Hood and T. C. Koopmans (eds), *Studies in Econometric Method*, Wiley.

Skow, B.: 2016, Explanations, Why-Questions, Reasons, and Causes.
URL: *http://web.mit.edu/bskow/www/research/levels-of-reasons.pdf*

Spirtes, P., Glymour, C. and Scheines, R.: 2000, *Causation, Prediction, and Search*, MIT Press.

Spohn, W.: 2002, Laws, Ceteris Paribus conditions, and the Dynamics of Belief, *Erkenntnis* **52 (special issue)**, 373–394.

Stalnaker, R.: 1968, A Theory of Conditionals, *Studies in Logical Theory, American Philosophical Quaterly* pp. 98–112.

Stanley, J. and Szabo, Z. G.: 2000, On Quantifier Domain Restriction, *Mind and Language* **15**, 219–261.

Steel, D.: 2006, Homogeneity, selection, and the faithfulness condition, *Minds and Machines* **16**(3), 303–317.

Storer, T.: 1951, On Defining 'Soluble', *Analysis* **11**(6), 134–137.

Strevens, M.: 2007, Mackie remized, *in* J. Campbell, M. O'Rourke and S. H. (eds), *Causation and explanation, Volume 4 of topics in contemporary philosophy*, MIT Press, Cambridge, MA.

Strevens, M.: 2008, *Depth: An Account of Scientific Explanation*, Harvard University Press.

Suárez, M.: 2010, Scientific Representation, *Philosophy Compass* **5**(1), 91–101.

Suppes, P.: 1970, A Probabilisitic Theory of Causality, *Acta Philosophica Fennica* **24**.

Swoyer, C.: 1982, The Nature of Laws of Nature, *Australian Journal of Philosophy* **60**(3), 203–223.

Tahko, T.: 2015, The Modal Status of Laws: In Defence of a Hybrid View, *The Philosophical Quaterly* **65**(260), 509–528.

Tooley, M.: 1987, *Causation*, Oxford Clarendon Press.

Van Fraassen, B.: 1989, *Laws and Symmetry*, Oxford Clarendon Press.

Van Fraassen, B.: 1993, Armstrong, Cartwright, and Earman on Laws and Symmetry, *Philosophy and Phenomenological Research* **53**, 431–44.

Van Fraassen, B.: 2008, *Scientific Representation*, Oxford University Press.

Vetter, B.: 2015, *Potentiality: From Dispositions to Modality*, Oxford University Press.

Ward, B.: 2002, Humeanism without Humean supervenience: A Projectivist Account of Laws and Possibilities, *Philosophical Studies* **107**(3), 191–218.

Ward, B.: 2007, The Natural Kind Analysis of Ceteris Paribus Law Statements, *Philosophical Topics* **35**(1 and 2), 359–380.

Ward, B.: 2009, Cartwright, Forces, and Ceteris Paribus Laws, *Southwest Philosophy Review* **25**(1), 55–62.

Weisberg, M.: 2007, Three Kinds of Idealization, *Journal of Philosophy* **104**(12), 639–659.

Williamson, T.: 2007, *The Philosophy of Philosophy*, Blackwell Publishing.

Wilson, J.: 2010, What is Hume's Dictum, and Why Believe It?, *Phenomenological Research* **80**, 595–637.

Wilson, J.: 2014, No Work for a Theory of Grounding, *Inquiry* **57**(5-6), 535–579.

Wittgenstein, L.: 1961, *Tractatus Logico-Philosophicus*, Routledge (1922), London.

Woodward, J.: 2003, *Making Things Happen*, Oxford University Press.

Woodward, J.: 2013, Simplicity in the Best Systems Account of Laws of Nature, *British Journal of Philosophy of Science* **0**, 1–33.

Woodward, J.: 2014, Interventionism and Causal Exclusion, *Philosophy and Phenomenological Research* .

Woodward, J.: 2015, Methodology, ontology, and interventionism, *Synthese* **192**, 3577–3599.

Woodward, J. and Hitchcock, C.: 2003, Explanatory Generalizations Part 1: A Counterfactual Account, *Noûs* **37**(1), 1–24.

Woody, A.: 2015, Re-orienting discussions of scientiâŘčc explanation: A functional perspective, *Studies in History and Philosophy of Science* pp. 1–9.

Zhang, J. and Spirtes, P.: 2011, Intervention, determinism, and the causal minimality condition, *Synthese* **182**(3), 335–347.

Zilsel, F.: 1942, The Genesis of the Concept of Physical Law, *The Philosophical Review* pp. 245–79.