1

From Causal Relations to Causal Constraints

THIS BOOK EXAMINES the family of causal notions and causal constraints employed in fundamental science, and analyzes some of the conceptual relations between them. It argues that the concepts of determinism, locality, stability, and symmetry, as well as conservation laws and variation principles, constitute a complex web of constraints that circumscribe the causal structure of our world. It argues, further, that mapping out the various links between these causal constraints is an indispensable, though neglected, aspect of the project of understanding causation. The book thus seeks to shift our attention from causal relations between individual events (or properties of events) to the more general causal constraints found in science, and the relations between them. In so doing, it does not purport to replace causal relations with causal constraints in every context, but rather to suggest a broader perspective on causation and a new research program for the philosophy of causation.[1]

Philosophical analysis of complex concepts usually begins with definitions. The exploration of causality is no exception. Enormous effort

1. I take "causation" and "causality" to be synonymous, generally using the latter when referring to writings that use this term, and the former otherwise. As I explain, I extend the application of these terms beyond a relation between individual events, and hence they cannot always be associated with specific cause-events or effect-events. The same caveat applies to the notion of causal *constraint*, which is the focus of this book.

has been devoted to formulating the "right" definition of causation and defeating rival definitions. Regularity theories, counterfactual analyses, interventionist/manipulation accounts, probabilistic theories, transmission accounts, and explanation-oriented accounts are regular contenders in this ongoing competition, which comprises much of the literature on causation. Each of these definitions captures important characteristics of the notion of cause, but also raises difficulties that advocates of competing conceptions are quick to seize on. Needless to say, the contending accounts are never conclusively defeated by such difficulties; their advocates find ways to patch them as necessary. Nevertheless, the "attack/patch" cycle has an adverse cumulative impact as the difficulties pile up. More generally, concern over definitions and their weaknesses has led philosophers to devote a great deal of attention to intriguing yet marginal "hard cases." Adding "epicycles" may salvage a threatened definition of causation, but sheds little light on the ways in which causal notions are actually used, and in particular, on how they're used in scientific contexts. Science seeks to identify constraints that distinguish what may happen, or is bound to happen, from what is excluded from happening. Hence the notion of causal *constraint*, which is broader than the notion of cause, is at the center of my analysis. Even when searching for individual cause-events (effect-events), awareness of the framework of constraints that these individual events must satisfy is vital. And because there is no single causal constraint that is operative in science, but rather several different constraints, a study of the relationships between the various constraints is called for. I do not take the notion of cause to be reducible to any one of the constraints in question, or to a particular combination of them. The evolution of causal constraints—and thus of our understanding of causation—is as open-ended as the evolution of science in general. The difference between the "causal constraints" approach to causation, and the traditional approach, will become sharper as the book proceeds.

I will not review the current philosophical accounts of causation, and the difficulties they pose, in any detail. The *Oxford Handbook of Causation* (Beebee, Hitchcock, and Menzies 2009) gives an admirably balanced account of the literature. But it will be useful to briefly identify the main contending proposals, and the key issues they bring to the fore, as these key issues underscore my claim that it is time for a different approach to causation.[2]

- **Regularity theories**, also known as Humean theories, reduce causation to lawful behavior and therefore assimilate causation to determinism.[3] Nonetheless, it is not laws that constitute causes, but the events that fall under them. Roughly, Hume's definition of causation set down three conditions: contiguity in space, succession, and constant conjunction of the same event types.[4] The succession condition can in turn be divided into a condition of contiguity in time, and an asymmetry condition to the effect that the cause must precede the effect. Physics has had to discard the spatial and temporal contiguity requirement due to its emptiness in continuous spacetime, and thus in mathematical theories that involve a continuum (such as theories employing differential and partial differential equations). But the remaining conditions are independent of the contiguity requirement and permit extension of the cause–effect relation to distant events: *any* event *c* that is regularly or lawfully followed by an event *e* can be considered the cause of *e*.

2. The few references in the following sections are only examples of relevant literature. Beebee, Hitchcock, and Menzies (2009) provides a wealth of bibliographic information.

3. The concept of determinism is touched on later in this chapter and examined in more detail in chapter 2.

4. Hume devotes extensive sections of *A Treatise of Human Nature* and *An Enquiry concerning Human Understanding* to causation. This brief summary ignores many interpretative issues and the vast literature thereon. Furthermore, it extends the meaning of "Humean" beyond anything Hume himself would have recognized. Even metaphysically extravagant accounts such as David Lewis's are often considered Humean, to say nothing of accounts, such as Mackie's and Davidson's, that deviate less radically from Hume's own account. Davidson's view is particularly interesting in this respect, as it is committed to the existence of laws that have no scientific utility. I would argue that Davidson's main point about causation (the distinction between causal and explanatory contexts; see below) is independent of the Humean commitment to laws. Other issues discussed under the rubric of "Humean causation" include the question of "Humean supervenience"—are there natural laws or only natural facts?—and questions about the status of laws—what sense, if any, can we give to the metaphor of laws "governing" the physical world? See, e.g., Maudlin (2007). I do not address these problems.

The connection Hume established between causation and lawful behavior has had a lasting impact on the philosophy of science. Yet regularity theories, though still among the leading accounts of causation, have also garnered much criticism. Objections target the very connection between regularity and causation, denying either the necessity or the sufficiency of regularity for causation. The claim that regularity is unnecessary for causation entails the acceptability of singular—that is, nonrepeatable—cause–effect relations. Although I do not deny the feasibility of such singular causal relations, their existence is peripheral to my primary concern. When focusing on science rather than, say, human actions, it is largely possible to remain within the boundaries of lawful causation.[5] The converse claim—that regularity is insufficient for causation—is backed by several arguments. For one thing, regularities, even when they appear to be lawlike, may reflect accidental rather than causal connections. For another, regularities as such lack the asymmetry typical of causal relations.[6] There are also examples like the tower and its shadow, which, despite the nonaccidental nature of the regularity in question, speak against the identification of lawful regularities with causation. The height of the tower and the length of the shadow are correlated by laws, but we see the shadow's length as caused (and explained) by the tower's height, and not the other way around. The example suggests a distinction between causal

5. The possibility of lawless causal relations and lawless actions is discussed near the end of chapter 7; the constraints examined up to that point are all lawful. I see no objection to considering singular events such as the Big Bang or the extinction of dinosaurs to be causes of later developments, but ascription of causal roles to singular events is not pivotal for my treatment of causation. There is an interesting exchange of letters, from 1920, between Einstein and Schlick on the question of causality without regularity. Einstein initially maintained that regularity was unnecessary, and suggested a hypothetical scenario in which such singular causal relations had to be posited. Schlick then convinced him that without regularities we could not even take measurements, and that our scientific notion of cause thus presupposes the existence of regularities. But Einstein continued to maintain that once we have formulated the concept of lawful causality, we should be able to identify singular causal relations (Albert Einstein Archive, 21: 576; Schlick 1920).

6. This asymmetry is touched on at the end of the chapter.

and noncausal regularities, where only the former are truly explanatory.[7] The concern that regularity falls short of causation often motivates the requirement that causal connections, unlike mere regularities, must be embodied in concrete mechanisms.[8] While the distinction between lawful and accidental regularities is crucial for science (and in that sense the critique of regularity theories is warranted), a connection between causation and concrete mechanisms is often lacking. When it comes to very general causal constraints, such as the relativistic limit on the speed of interaction, the search for an underlying mechanism is futile. The distinction between laws and mere regularities can be supplemented by a hierarchy that differentiates lower-order constraints, constraints on facts, from higher-order constraints, constraints on laws. Although constraints on laws do not fit the Humean scheme, they should be seen as causal constraints; see chapters 5 and 6. But even if we were able to fend off all the standard objections to the regularity account of causation, it would, from the perspective of this book, remain inadequate. Except for determinism, the constraints that make up the family of causal concepts (henceforth, causal family) cannot be expressed in the language of regular succession of individual events.

- The **counterfactual account** championed by David Lewis (1973) analyzes the causal relation between event *c* (the cause) and event *e* (the effect) in terms of the counterfactual "had *c* not occurred, *e* would not have occurred."[9] In addition to the formidable problem of analyzing counterfactuals,[10] and the

7. Arguing for a pragmatic approach to explanation, Van Fraassen (1980, 132) attempts to destabilize this intuition by depicting a case in which the length of the shadow is the motive behind construction of the tower. I would argue that even in this contrived case, it is the tower that causes and explains the shadow, but in any event, the example is not an instance of standard scientific explanation.

8. See Glennan (2009) and the literature cited there.

9. As Lewis notes, this account can be found in Hume's writings alongside the regularity account, giving rise to interpretative questions about Hume's "true" analysis of causation.

10. We will encounter some of these problems, in particular, sensitivity to description, in chapter 2.

metaphysical assumptions this analysis mandates, the counterfactual account faces the challenge of overdetermination. Recall the standard example (a typical hard, though marginal, case) of two desert travelers who set out, separately, to murder a third, one pouring poison into the victim's water flask, the other puncturing it, so that the victim will die of either poisoning or dehydration. The counterfactual conditional "Had *x* not punctured the water flask, *z* would not have died," fails to identify the cause, for (assuming the cause of death to be dehydration) it would not be true that had the water flask not been punctured, the victim would not have died. I should stress that I do not adduce these problems to critique counterfactual considerations in general, but to critique their adequacy as definitive criteria of causation.[11] I take counterfactuals to be indispensable for reasoning, and will use them extensively in chapter 2. But counterfactuals are also used in contexts that have nothing to do with causation: "If I were you, I would accept the offer"; "This triangle (pointing, for example, to a triangle with sides 3 cm, 4 cm, and 6 cm in length) is not right-angled—if it were, it would satisfy the Pythagorean theorem." Because of their broader applicability, counterfactuals cannot be relied on to pick out causal relations.

- **Process accounts** of causation focus on prolonged progressions rather than instantaneous events, and tie causation to a particular process, such as energy transfer from one system or state to another (Fair 1979; Salmon 1984; Dowe 1992, 2000). This approach can handle the case of Jane's happiness being due to John's response to her message but has difficulty with the seemingly parallel case of Jane's unhappiness being due to John's failure to respond. In other words, the process approach is unable to account for failures and omissions.
- **Probabilistic accounts** of causation (Suppes 1970; Kvart 1986) have the great advantage of extending causal discourse to non-

11. Critique of counterfactual reasoning is quite common; see, e.g., James ([1884] 1956).

deterministic contexts. From the probabilistic perspective, a cause need only raise the probability of the effect-event's occurring; there is no need for it to determine or induce its occurrence. On the other hand, however, probabilistic accounts engender paradoxes of their own—namely, probabilistic correlations that do not seem to reflect causal relations, or events that seem entitled, from the intuitive point of view, to be considered causes of certain "effect" events, yet appear to lower, rather than raise, the probability of their occurrence.

- **Interventionist or manipulation accounts** of causation (also known as agency accounts) have been in vogue for several decades (Menzies and Price 1993; Woodward 2003; Hitchcock 2007). Here, causes are identified as the factors that, when manipulated, change the result that would have ensued in the absence of that intervention. In identifying causes as necessary conditions of effects, the manipulation account has much in common with the counterfactual account. In focusing on human intervention, however, it has given rise to the objections of anthropocentrism, and—since intervention is actually a causal concept—circularity. One merit of the interventionist criterion is that it distinguishes causal relations from correlations that are merely accidental. Together with insights from the counterfactual and probabilistic accounts, it has stimulated elegant work on causal networks and their graphic representations (Pearl 2000). Causal networks are highly valuable in a variety of practical contexts—legal, medical, economic, policy-making, and so on—where distinguishing effective from ineffective intervention is essential. Nevertheless, the manipulation account, I contend, is far too limited to provide a comprehensive understanding of causal processes in the world. This point requires elaboration.

Consider, first, the contrast between the regularity account and the manipulation account. In a standard case, where natural laws and initial conditions determine a certain result (a certain chemical reaction, say), we can often manipulate the initial conditions, but not the laws. On the manipulation account, therefore, the initial conditions constitute the only relevant factor for a causal account of the process. By contrast, the regularity theorist ascribes a crucial role to the laws—to what we *cannot* manipulate.[12] The initial conditions can be considered causes only because they are invariably followed by the same trajectories, that is, they are considered causes only because of the existence of laws that are non-manipulable constraints. And although (for the reasons mentioned above) I do not consider the regularity account, as it stands, fully satisfactory, there is something fundamentally correct about the intuition that constraints which we cannot manipulate are an inherent feature of causal descriptions and explanations. Furthermore, invoking laws is by no means the only context in which we ascribe a causal role to the non-manipulable. The electric charge of the electron, which though non-manipulable is causally efficacious, is a case in point.[13] The constraints we will examine in this book are typically not subject to human intervention, but they enable us to grasp and predict the dynamics of unfolding events, and exclude infinitely many alternatives to what actually transpires. In this sense, they also constrain our interventions—manipulations and interventions are always carried out *within* a general framework of constraints. The manipulation account of causation thus already *presupposes* the preconditions of possible manipulation, preconditions that a complete causal story must take into account and render explicit.

12. Manipulation theorists such as Woodward (2003) make it clear they do not restrict manipulation to procedures and actions that humans can actually carry out, but allow for a broader class of manipulations that are possible in principle. Laws of nature, however, are beyond control even in this weaker sense.

13. A more controversial example is Newtonian space, which plays a causal role in Newton's theory—acceleration relative to absolute space has genuine physical effects—yet this space and its geometric structure are fixed and cannot be manipulated. This interpretation is in line with Einstein's view of the matter (see Mutuality of Causal Relations at the end of this chapter). See, however, DiSalle (1995) for a critique of this causal interpretation of space.

In view of the difficulties that beset each of these accounts, I have relinquished the search for *the* definition of causation, instead taking causation to be a cluster concept comprising a broad range of causal notions. My primary focus will be the causal notions employed in science, which include the much-discussed notion of determinism, but also notions such as stability and locality, which philosophers tend to neglect. In turning to pluralism, I am not, of course, alone: a pluralistic attitude to causation has been advocated (sometimes only in passing) by Reichenbach (1956); Anscombe (1971); Cartwright (1983, 2004); and Godfrey-Smith (2009), not to mention Aristotle, who introduces his presentation of the four causal categories by noting that the number of causes matches that of the things "comprehended under the question 'why' " (*Physics* II, 198a15–16). Skyrms (1984) speaks aptly of an "amiable jumble" of causal notions that can, but need not, work together.

But the declaration of pluralism is only a starting point. To do justice to causation, recognition of the variety of causal notions must be augmented with detailed investigation of their usage, especially in fundamental science. Science, like daily life, presents us with a spectrum of causal notions and constraints. These scientific constraints are often in some way "descended" from the more intuitive constraints of daily life, though differing from them significantly in precision and scope. The invocation of causal notions in scientific contexts is particularly noteworthy in view of arguments that challenge causation's place in fundamental science, or relegate it to "folk science" (Russell 1913; Norton 2007). I will return to the Russell-Norton position further on, but for now, let me point out that the argument I make in this book is that as soon as we shift our attention from the familiar paradigms of breaking a glass or tickling a baby to determinism, locality, stability, and conservation laws, it becomes evident that causal notions permeate fundamental science.

Critique of causal discourse also comes from another direction. These critics concede causation's place in fundamental physics, but deny it elsewhere, arguing that higher-level realms, such as the realms of biological, mental, and social events, are causally inert. Alleged causal relations on these levels, or between higher-level events and events at the fundamental level, are (according to this view) all reducible to the causal relations of physics. This argument is rebutted in chapter 7.

Our first encounter with causal notions does not come from fundamental science. We acquire the notion of cause early in life in relatively simple situations, such as sucking, pulling, pushing, holding, biting, and so on. These actions involve several elements of the cluster concept of cause, so that many of the features explicated in the aforementioned accounts of causation are operative. A child pulls the string of a toy that plays a tune. The interaction is both regular—whenever the string is pulled the tune plays—and local—no action at a distance; it instantiates both the manipulative and the counterfactual conditional accounts of causation—had the string not been pulled, no tune would have been played; and there is no creation ex nihilo—the power was provided by tugging the string, energy was transferred from the child's hand to the string, and from the string to the musical instrument. Infants are unable to articulate these concepts, but may acquire a rudimentary grasp of some of them, and learn to associate the features in question with each other, so as to form a more complex sense of causation. Later on, with exposure to less paradigmatic causal nexuses, and to science, intuitive ideas give way to more explicit notions, occasionally undoing the automatic associations, or establishing new ones. A child who is at one stage prone to "magical" thinking, for instance, believing that merely wishing someone ill suffices to actually bring about harm, will, with age and experience, likely revise this conception.

A recurrent concern about causality, dating back to Hume, derives from empiricism: do we ever *observe* causal connections? And if not, ought we not either renounce causation or reduce it to observable features of the world? Anscombe's influential *Causality and Determination* (1971) argues, contra the Humeans, that we do indeed observe and experience numerous *instances* of causal connection: pushing, breaking, burning, and so on. I agree with Anscombe. Granted, there are also many less evident cases, where the causal connection is not observable, but the same goes for other relations; they too are manifest in paradigmatic situations and remote from immediate experience in others. Motherhood, for instance, is not in general a directly observable relation, but when we happen to be present during a delivery, we can witness it directly. And we can now adduce empirical evidence of motherhood in DNA sequences. This evidence does not render motherhood directly observable, but establishes it beyond a reasonable doubt, making it as objective as other empirical relations. Moreover, it has long been acknowledged that science does not restrict itself to the directly observable; it is empirical only in the sense that it expects its nonobservable concepts and laws to have observable implications. In this respect, causal thinking is in line with scientific thinking in general. My perspective on causation is realist, and I take causal constraints to be objective. This realism should not be construed as a commitment to the existence of causes as metaphysical entities that exist in addition to the entities they relate. Even for those who do not embrace his account in its entirety, Hume's critique of the traditional conception of causation has discredited this picture of causes as hidden "arrows" between events. The characterization of realism in mathematics as a commitment to the objectivity of statements rather than the existence of mathematical objects (Kreisel 1958) can be applied, mutatis mutandis, in the context of causation: realism about causation means that causal claims are objectively true or false.

I should, perhaps, note that the empiricist status of causation has undergone an ironic transformation. Hume deemed spatial and temporal relations legitimate from the empiricist point of view, using them as the basis for his definition of causation in terms of constant conjunction. But the relationship between the causal and temporal orders turned out to be quite different from that which he envisaged. According to the special theory of relativity (STR), the temporal relations between events are only well defined in regions of spacetime charted by light signals representing (and limiting) the *possibility of causal interaction*. When events are separated by space-like distances, there can be no causal interaction between them, and consequently, their temporal order is not invariant, but varies with the coordinate system. Rather than being reducible to spatiotemporal relations, causality now appears to be the basis for the very structure of spacetime. Causal relations are thus at least as fundamental as temporal relations, and arguably (as suggested, for example, in Reichenbach 1956), conceptually prior to temporal relations.

Typically, definitions allow us to replace the defined term (*definiendum*) with the terms that define it (*definiens*). When, for instance, Mackie (1965) argues that a cause is (at least) an INUS condition, he is suggesting that "the short-circuit caused the fire" could be replaced with "the short-circuit was an INUS condition of the fire."[14] By contrast, causal *constraints*, being necessary but insufficient conditions, do not replace the notion of cause in this way. Even so, they constitute an essential aspect of using and understanding causal discourse. In accepting Dan's alibi to the effect that at the time his Cornwall cottage was set on fire, he was in Oxford giving a talk, the court takes it for granted that there is no action at a distance. Although this causal assumption constrains our identification of causal connections, it does not allow us to replace locutions signifying causal constraints with the term *cause*, nor does it identify the cause of the fire. Similarly, when physicists refer to the limit on the speed of interaction as "relativistic causality," they are using the term *causality* to refer to a constraint—a necessary condition—that neither defines causality nor points to the cause of a particular process or event. The same caveat applies to other constraints, such as symmetries, conservation laws, and variation principles: they circumscribe our causal thinking but do not provide us with synonyms for the term *cause* or coextensional alternative locutions. It appears that no one constraint constitutes a condition that is both necessary and sufficient and could thus serve as an adequate definition of the notion of cause, a definition that covers all its applications. Hence there is also no general "causal principle" (more on this later). I have, therefore, relinquished the quest for such a definition and adopted a pluralistic approach, taking the notion of cause to be an irreducible cluster concept covering various constraints imposed by the theories we employ. The cluster's com-

14. An INUS condition is a condition that in itself is neither necessary nor sufficient for the occurrence of the effect, but constitutes a necessary component of a complex that is sufficient but unnecessary for bringing about the effect. INUS stands for Insufficient but Necessary (in a cluster that is) Unnecessary but Sufficient.

ponent concepts do not apply across the board; where they apply, and to what level of precision, is an empirical question.

The fact that the causal constraints examined in this book do not presuppose individual cause-events or effect-events, or that such putative events cannot always be picked out, is no drawback: in scientific contexts, cause–effect language (and ontology) is not always helpful. Indeed, it is generally cause–effect ontology, as opposed to causal discourse per se, that is targeted by critiques of the concept of cause. We can understand why a certain chemical reaction occurs in the direction it does by pointing to symmetry principles and conservation laws. From the perspective adopted here, this would be a perfectly adequate causal explanation, though it does not single out any individual event as the cause of the outcome. I am not claiming that we can always dispense with identification of individual cause-events, or that we ought to do so.[15] Nor do I want to question the utility of the ontology of events: we certainly do wish to apply the general constraints to individual systems, states, and events. But typically, these applications do not pinpoint any single event as a cause, and all the more so, as *the* cause. There are contexts, however, not only in daily discourse, but also in science, in which ascription (or denial) of causal roles to individual events becomes essential. In assessing the implications of relativistic causality, for example, it is sometimes crucial (as we will see in chapter 4) to identify causal relations and the transmission of information (or lack thereof) between individual events. But even when such ascription of a causal role to specific events is irrelevant or meaningless (as in the example of Newtonian absolute space), it does not follow that there is no causal story to tell.

In the literature, considerable effort is devoted to distinguishing causes from conditions, and singling out *the* cause of an event from other events that stand in a causal relation to it but lack some feature that would make them the sole cause. Speeding might be singled out as the cause of an accident, while the curve in the road, the weather, or the design of the vehicle, are deemed mere "conditions." The distinction is

15. I do not deny the existence of causality in singular cases, see note 5 earlier in this chapter.

generally considered pragmatic and thus context-dependent. The idea is that whereas, from the logical point of view, many factors stand in a causal relation proper with the effect-event, from the pragmatic point of view, it is legitimate to distinguish just one of them as its cause. Pragmatic criteria include, for example, deviation from the regular course of events, and human intervention (Hart and Honoré 1959; Mackie 1965). Both these criteria distinguish the speeding from the other conditions that played a causal role in bringing about the accident. But in a different context, say that of evaluating the road's safety, the curve in the road might be the focus of blame and the target of intervention, whereas speeding drivers are deemed background conditions. Such pragmatic considerations are not foreign to scientists, who invoke them routinely when planning experiments and analyzing their results, but with regard to the causal constraints addressed in this book, they can be set aside.

What, then, is the role of causal notions in science? Causal notions and constraints, I suggest, are employed to describe, predict, and explain change. They tell us which processes and changes in the physical world are possible, and which are not. This characterization gives us a far broader picture of causation than the picture painted by portraying only cause–effect relations. Causal notions in this broader sense, though not the only explanatory notions, are unique in explaining *change*. Logical and mathematical notions may also play an explanatory role, and, in the realm of human action, reasons, arguably different from causes, fulfill a central explanatory function. But our understanding of change in the physical world is not, and cannot be, complete without causal notions. Typically, causal notions involve changes that occur over time, a characteristic that distinguishes them from the nontemporal relations found in mathematics. And they also involve matter—masses, forces, fields, and their interactions—which again sets them apart from the purely mathematical. Thus even when expressed in mathematical language, causal relations and constraints go beyond purely mathematical constraints; they are (at least part of) what we add to mathematics to get physics.

Causal relations in the physical world have not always been properly distinguished from necessary connections in the logical-mathematical realm. In Spinoza's system, for instance, logical necessity and causal necessity are on a par. But the disentanglement of these kinds of necessity has become common, if not mandatory, in modern science. Time, matter, and the possibility of change are crucial to maintaining a distinction between physics and mathematics. To reiterate my characterization, a causal constraint is any constraint that delimits change, distinguishing changes that are sanctioned by science from those that are ruled out. The test of legitimacy is empirical: instances of legitimate change are detectable in the physical world, excluded changes are not. This conception of the difference between physics and mathematics generates a natural account of causation in terms of temporal change and the constraints that such change must satisfy. One alternative to this natural account is to erase the distinction between physics and mathematics altogether and embrace a hyperrationalist picture of the world, such as Spinozism, or a hypermathematical one, such as Pythagoreanism. I find this alternative unappealing.

Lange (2017) argues for another alternative.[16] In addition to causal explanations, on the one hand, and purely mathematical explanations, on the other, he introduces a third category of explanations, which are neither causal nor mathematical. Interestingly, he focuses on the notion of constraint in this context, referring to this third category as "explanation by constraint." Lange's attentiveness to constraints is commendable, but whereas the constraints I speak of are *causal*, he deems explanation by constraint noncausal, implying that the notions of cause and constraint exclude each other. The rationale for this exclusion seems to be recognition of a hierarchy of laws, some being more general—and hence, in Lange's view, more necessary—than others. Lange takes Newton's second law of motion to be higher up in this hierarchy than the law of universal gravitation or Coulomb's law, because Newton's second law applies to forces in general and would presumably apply to new forces, were any to be discovered. The hierarchical picture is apt—some causal constraints do indeed apply to *laws* rather than events—but why reserve causal status for lower-level laws? Imposing this restriction

16. I thank an anonymous reader for calling my attention to this recent book.

involves Lange in extended discussions of what does and doesn't count as causal explanation, the sort of analysis that scientists do not engage in, and I am trying to avoid.[17] The inclusive concept of causation as comprising *any* constraint on change, regardless of its place in the hierarchy of laws, affords a better understanding of the function of causal notions in science.

Chapters 5 and 6 provide illustrations of the causal role of higher-level principles such as symmetries and variation principles, and of the difference between mathematical and physical constraints. Symmetries, for instance, are expressed in mathematical language—the language of group theory—which gives them a formal, even a priori, appearance. But there are facts about the world that, despite their expression in this group-theoretical language, cannot be considered mathematical facts. A physical process may be invariant under spatial translation, a feature that is reflected in the mathematical formulation of the process and the laws it obeys, a formulation that is independent of specific coordinates. But the fact that this is the correct formulation of the law, viz., that this symmetry is reflected in reality, is not a mathematical fact—we could envisage physical processes and laws that are not invariant in this way (as was actually the case in Aristotelian physics). Furthermore, among the symmetry considerations adduced by physicists, we find Curie's principle, according to which (roughly) we cannot get asymmetry from symmetry. Rather than being a purely mathematical theorem, Curie's principle (discussed in chapter 5) identifies changes we can expect to find in the physical world; that is to say, it is a causal principle. The causal constraints encountered in fundamental science go well beyond the intuitive causal concepts we grew up with, and well beyond the examples that recur in the philosophical literature on causation. Inclusion of symmetry considerations in the causal family is a prime example of the extension that is called for when we move from causal relations between

17. E.g., Lange elaborates on a distinction between cases in which the law of conservation of energy functions as a causal explanation, and cases in which it functions as an explanation by constraint (2017, chap. 2). I see both cases as clear instances of causal explanation. Note also the contrast between his account of Pauli's exclusion principle (183) and my causal account of this principle in chapter 5.

individual events to a much more general understanding of physical change.

It might be objected that temporality no longer distinguishes physics from mathematics; the time required for a calculation, for instance, is a major parameter in computation theory. But this objection is misguided. The temporal terminology in computational science is premised on the realization that computations are carried out by physical systems—human beings or machines—that are constrained by physical possibility and cannot perform instantaneous calculations. But this realization is not part of mathematics. The fundamental notion of computational theory—the number of steps required—is indeed mathematical. But though the length of a calculation in terms of the number of steps it takes is a mathematical consideration, its length in terms of the time it takes is not—it involves assumptions about the physical world. Similarly, the causal properties ascribed to algorithms such as those of cellular automata are also a figure of speech. In describing John Conway's Game of Life (Berlekamp, Conway, and Guy 1982; Gardner 1970), we might say, for example, that step n is the cause and step $n + 1$ the effect, but this formulation tends to conflate the algorithm with the computer that implements it. The computer does indeed operate in a causal manner, using electric circuits and the like, so that each of its states is causally related to earlier ones. The algorithm, however, is temporal and causal only in a metaphorical sense.

To fully understand change, we must be able to understand not just what happens, but also what *fails* to happen or is *excluded* from happening. By the same token, it is just as causally relevant to learn that a system is *in*sensitive to a certain parameter as to learn that it is sensitive to it. Traditional accounts of causation do not fully address this negative aspect of change and causation. It should be noted, first, that there are different kinds of exclusion. When we think of an event c as bringing about an effect e, it is implied that whereas e must, given c, occur, every other outcome is excluded. This kind of exclusion is specific to a particular state of affairs; an event that is *ruled out* under one set of initial conditions may be permitted, or even mandatory, under different initial conditions. There are, however, types of events—certain chemical or nuclear reactions—that *never* occur, regardless of the initial conditions. From the physicist's perspective, what does not happen has just as much causal significance as what does. Regarding such absolute exclusions, the working assumption is that there must be some underlying principle that explains them. Symmetry principles and conservation laws provide explanations of this kind and are sometimes explicitly formulated in terms of exclusion rather than affirmatively, in terms of what they mandate. Pauli's exclusion principle (discussed in detail in chapter 5) is a case in point. As noted, it is also important to distinguish between constraints that bar specific event types and constraints on the general form of laws. Symmetries exemplify the latter type of constraint as well. They are therefore often considered to rank higher in the hierarchy of laws of science than ordinary laws.

We can think of physical constraints on legitimate change in terms of an analogy that invokes two different models of legality. On the first, usually deemed applicable to state officials, everything that they are not mandated by law to do is prohibited.[18] On the second, usually deemed applicable to citizens, everything that is not prohibited by law is permitted. Similarly, we can think of the constraints imposed on natural processes either as necessitating everything that happens, or as excluding certain occurrences, but leaving a considerable amount of freedom: whatever is not excluded may happen. On the former model, science is expected to show that the occurrence of anything that happens is determined by law, while the occurrence of anything else is excluded by law. This expectation places severe restrictions on what will be considered an adequate scientific theory. On the latter, freedom-granting model, science is only expected to show compatibility with the law. That is, it suffices for science to formulate laws that are not violated, or to put it differently, laws that permit, rather than determine, what happens. In principle, the two models could converge. That is, it could be the case

18. The analogy is incomplete: first, because in the legal realm laws are normative and are often violated in practice, whereas in science they are descriptive and, to the extent that they are true, cannot be violated. Second, the freedom-excluding model is too extreme; officials, even in their capacity as such, have some liberties. Overall, however, for officials, the freedom-excluding model is the default account.

that we begin by listing the proscriptions, with the intent to delimit the freedoms that remain, but by the time we have the complete list of proscriptions, we discover that there is no room left for freedom, and everything that does happen is in fact determined by law. Many scientists strive to demonstrate the reality of this rigid scenario, seeking to make the list of exclusions sufficiently comprehensive to eliminate any freedom whatsoever. Such an ambition was voiced by Einstein, who pondered the question of whether God had any choice when creating the world. On the no-choice picture, brute facts, contingencies that are inexplicable by fundamental laws, are an embarrassment to science. The theory of the Higgs field, and the search for the Higgs boson that confirms it, are motivated by a desire to derive particles' masses from general principles rather than accept their values as contingent and inexplicable parameters.

The freedom-excluding scenario is not, in fact, realized in contemporary physics, where the two models coexist. The freedom-excluding model has obvious links to determinism, but recall that even when laws are deterministic, the question of freedom might still be applicable to the initial conditions.[19] On the other hand, quantum mechanics (QM), its indeterminism notwithstanding, invokes more symmetry principles than are recognized in classical mechanics, taking them to be strictly (rather than probabilistically) obeyed.[20] As a rule, symmetry principles fit the freedom-granting model; they exclude certain processes, certain nuclear reactions, say, but leave room for more than a single possible outcome. A particularly interesting combination of freedom and necessity can be found in Feynman's picture of quantum mechanics, where freely moving particles seem, nonetheless, to behave as if they were exemplifying the rule that "everything that can happen does happen."[21]

19. Newton, e.g., thought that the solar system's initial conditions were not determined by the laws of physics, but ensued from God's benevolent choice; these initial conditions were later derived from mechanics. The general question about initial conditions, however, is still open, and is often a matter of controversy, as in statistical mechanics. See Albert (2011).

20. For the moment, I ignore approximate symmetries.

21. The question of the relation between this tenet and the traditional causal principle(s) merits examination, but I will not take it up here.

Cox and Forshaw (2011) adduce this tenet as the theme of their Feynmanian exposition of quantum mechanics. A similar view was debated in the seventeenth century, though with the theological gloss prevalent at the time. Leibniz, in disparaging this view, argued that to expect God to realize every possibility, regardless of its merit, is comparable to the expectation that a poet should "produce all possible verses, good and bad" (Strickland 2006, 137). Feynman's version of QM, and its implications for causation, are discussed in chapter 6.

Up to this point I have given two reasons for broadening our conception of causation beyond its familiar philosophical habitat. First, the causal notions and constraints explored here are all required for a comprehensive understanding of changes that take place in the world, and are the tools scientists employ to acquire such an understanding. Second, omissions and exclusions, which are integral to any account of causation in science, but constitute notorious stumbling blocks for most philosophical accounts of causation, fit smoothly into the picture suggested here. There are two further considerations that support the broader approach to causation. One is historical: the causal constraints of contemporary science are the progeny of intuitions and assumptions that have been associated with causation for as long as memory serves, and are therefore rooted in a long tradition of causal discourse. The other is conceptual, and pertains to the links between different constraints. Viewing causality as a manifold enables me to bring to the fore questions about the relationships between determinism and locality, determinism and stability, stability and symmetry principles, stability and variation principles, and so on. These questions, which have received little philosophical scrutiny, can be tackled from a general conceptual viewpoint or from the perspective of a particular scientific theory. Such an investigation will yield answers that support my reading of causation as a family of interrelated concepts. But let me add that these interrelations are worthy of scholarly attention regardless of the validity of my claim that all the constraints in question are in fact members of the causal family, and are needed for explication of the notion of causation.

I have been moving freely between the context of causation and that of causal explanation as if they were interchangeable. To be more precise, we should follow Davidson in recognizing a crucial difference between the two contexts. In his celebrated "Causal Relations" ([1967] 1980), Davidson observes that while truth values of singular causal statements are independent of the descriptions used to refer to the related events, explanatory contexts, much like other intensional contexts, are description sensitive. Compare, first:

1. Lord Kelvin made significant contributions to thermodynamics.
2. Sharon knows that Lord Kelvin made significant contributions to thermodynamics.

The truth value of (2), unlike that of (1), may change when "William Thomson" is substituted for "Lord Kelvin," for Sharon may not know that William Thomson is Lord Kelvin. Davidson points to an analogous difference between (3) and (4):

3. The reaction caused the explosion.
4. The reaction explains the explosion.[22]

According to Davidson, singular causal relations are extensional—the truth values of sentences affirming (denying) them do not change when we refer to the same entities by means of different descriptions. By contrast, explanatory contexts, being sensitive to the descriptions of the events in question, are referentially opaque. This opacity is due to the fact that explanations comprise laws that connect *types* of events rather than individual events. To explain an event by subsuming it under a law (or set of laws), we must, therefore, refer to it by means of the right description—namely, the description matching the event type specified by the relevant law(s). Davidson's insight has often been overlooked, but is crucial for a proper understanding of the notions of determinism and stability. As we will see in chapter 2, Davidson's point is particularly relevant for the assessment of Russell's critique of causation qua determinism. Moreover, description sensitivity is characteristic of probabilistic explanations as well. Statistical mechanics provides an

22. My examples differ slightly from Davidson's. (3) and (4), in particular, are short for what Davidson formulates as: "That the reaction occurred caused it to be the case (explains) that. . . ." Sidney Morgenbesser is known to have made the same point. See also Steiner (1986).

instructive example of the scientific significance of descriptive categories. Here, the fact that macrostates vary enormously in the number of microstates they comprise is of crucial explanatory importance. Since we ourselves define macrostates, the explanatory import of statistical mechanics hinges on description-sensitive facts. Description sensitivity, however, does not breed subjectivity, as has been alleged. Once a description is chosen, the claims made in terms of this description can be objectively true or false. There are, of course, natural and unnatural, useful and useless descriptions, and finding the most helpful descriptions is far from trivial. But these obstacles do not entail any conflict between description sensitivity and objectivity. These points are elaborated on in chapters 2 and 3.

As mentioned, the role of causation in science has been a matter of controversy. Russell (1913) dismissed causation, arguing that mature sciences, physics in particular, consist of differential equations that do not invoke the notions of cause and effect.[23] More recently, John Norton (2007) has revived this negative attitude, arguing that the notion of cause can be tolerated in "folk science," but not in fundamental science. This position is referred to as the "eliminativist" or "error" theory of causation. Alluding to Russell's famous remark that the law of causality survives, "like the monarchy, only because it is erroneously supposed to do no harm" (1913, 1), Norton and other members of the dismissive camp are sometimes also referred to as "republicans" (Price and Corry 2007). Norton draws an analogy to science, where advanced theories typically recover the results of less advanced theories in some limited way—the predictions of classical mechanics, say, are derived from those of the special theory of relativity for velocities much lower than that of light. He thus seeks to recover the causal structure of our commonsense picture of the world from the more accurate depiction generated by fundamental science, which he takes to be completely free of causal notions. Despite their concurrence vis-à-vis causality, Russell's objection to determinism is quite different from Norton's: whereas

23. Quine too noted that "the notion of cause itself has no firm place in science" (1966, 229). Interestingly, Russell (1948) rehabilitates causality, espousing a view that is closer to process accounts than to regularity accounts.

Russell maintains that determinism is empty and trivially satisfiable by any theory, Norton argues that determinism is *false* even in the context of the theory considered its safest harbor—classical mechanics. I discuss Russell's position in chapter 2. Note, however, that in their critique of causation, both Russell and Norton actually focus on *determinism*, which is obviously a narrower concept of causation than that which I am recommending here. Their arguments, even if accepted, leave intact other causal constraints' applicability and usefulness in fundamental science.

The republican critique actually has two targets: the notion of cause and the causal principle. Although these concerns are not identical, both Russell and Norton connect them. Russell, as we saw, derides the causal principle, but also claims that "the word 'cause' is so inextricably bound up with misleading associations as to make its complete extrusion from the philosophical vocabulary desirable" (Russell 1913, 1). Norton links the notion and the principle even more directly: "Centuries of failed attempts to formulate a principle of causality, robustly true under the introduction of new scientific theories, have left the notion of causation so plastic that virtually any new science can be made to conform to it" (Norton 2007, 12). I grant this premise—there may be no "principle of causality" whose truth is secured a priori or established beyond reasonable doubt by experience. Indeed, there is not even an agreed-on formulation of the traditional principle. But the redundancy of the concept of causation does not follow from the demise of the causal principle. (Would it not be an overreaction to give up the *concept* of justice just because we are unable to formulate an overarching *principle* of justice?) Combining critique of the principle of causality with critique of the concept of cause (as Russell and Norton do) is, perhaps, understandable if one identifies causality with determinism and takes determinism to imply a very general principle about reality, such as "every event has a cause."[24] The failure of this general principle is then taken to imply the futility of the very concept of causation. But in my view, the principle as an assertion about the world (or our best theory of the

24. For the moment, this ancient version of the causal principle will do; more accurate formulations follow below and in chapter 2.

world) should still be distinguished from the concept. After all, it could be the case that some systems or processes obey deterministic laws though others do not, in which case the concept would be applicable despite the fact that the general principle fails. Thus, from the broader perspective adopted in this book, an open-minded attitude to the principle is appropriate. Rather than aspiring to a consensus regarding a universal causal principle, we must make do with a family of causal constraints that, like other natural laws, are subject to repeated testing and refinement.[25] And the same goes for the family of causal notions—they too must prove their value to science through their scientific applications.

In more recent publications (Frisch 2009a, 2009b; Norton 2009), the controversy over causation in science has shifted from the question of whether there is a meaningful principle of causality to the question of whether there is an *independent* principle of causality, that is, a principle leading to results that could not have been reached by any other physical principle. But if, as I maintain, conservation laws and variation principles *are* causal principles, then any result derived from them—and such results abound in physics—is derived from a causal principle or causal constraint (even if not from the sort of singular causal principle that Russell and Norton are so dismissive of). Moreover, seeking to establish the existence of an "independent" principle of causality is uncalled for. We might just as well ask whether there is an independent notion of a family, that is, if there is a family relation over and above, and independent of, being a daughter, brother-in-law, cousin, and so forth. Clearly, family relations can be subdivided, but does this make the notion of family redundant? I would argue that it doesn't, but the status of the general concept (of family and cause alike) is not the main issue. There may be no significant difference between the two pictures of causation—a single concept made up of several components, and a cluster of distinct concepts that are closely interrelated. If so, the debate over the term *causation* dissipates into a minor verbal disagreement. I want to stress, however—and this goes beyond the merely verbal—that gen-

25. Some of these constraints are quite general; see the discussion of Curie's principle in chapter 5.

uine questions remain about the relations between the various subconcepts. Regardless of whether we deem the notion of family redundant, we should be able to answer the question of whether cousins, say, can also be brothers. Analogously, regardless of whether we deem the notion of causation redundant, we should be able to answer the question of whether determinism implies locality or stability. This book is written from a pro-causation perspective, but the project it tackles—analysis of the relationships between members of the causal family—should, I believe, engage "republicans" as well.

To familiarize ourselves with the causal family, let me briefly introduce some of its members, emphasizing their connections to earlier traditions and intuitions about causation.[26]

DETERMINISM. The most prominent member of the causal family, determinism is frequently taken to be the core meaning of causation. It is also the meaning most closely associated with the so-called causal principle. Although the term *determinism* was coined in the nineteenth century, the ideas associated with determinism, such as exclusion of chance, go back to antiquity, and have been widely discussed ever since, under a range of rubrics, in particular causality and necessity. Determinism calls to mind two earlier principles: the universality principle, according to which nothing happens without a cause, and the regularity principle, according to which the same (type of) cause invariably leads to the same (type of) effect. In themselves, these principles are neither equivalent nor coextensional—a world satisfying one of them can violate the other. If, however, regularity is considered constitutive of causality (as it is in Hume's analysis), then a world satisfying the universality principle also satisfies the regularity principle. The converse does not follow. Despite the fact that "determinism" is often invoked as a feature of reality (for example, when debating the problem of human freedom), it is preferable to think of it as a property of theories. On the contemporary understanding, a theory is deterministic when it implies (roughly) that the entire trajectory of a closed physical

26. Clearly, specific laws such as Newton's laws can also be thought of as causal constraints, but I think of the causal family as including general rather than specific constraints.

system is determined by its initial conditions (or indeed, its conditions at any particular moment). When this stipulation is met, both regularity and universality obtain.

The contemporary definition of determinism thus combines the features that were traditionally thought to characterize the causal nexus, thereby linking causation and determinism. The differential equations of theoretical physics highlight this connection. Einstein put it as strongly as this: "The differential law is the only form which completely satisfies the modern physicist's demand for causality" ([1927] 1954, 255). Surprisingly, though, in restricting itself to closed systems, the contemporary definition of determinism creates new problems for some accounts of causation. By definition, a closed system cannot be interfered with. If one conceives of causality along the lines of the manipulation account, then, as Stachel (1969) has convincingly argued, determinism and causation are incompatible; the former can only be satisfied in closed systems, the latter in open ones. From the perspective of this book, however, neither the identification of determinism with causation, nor the claim that they are incompatible, is justified. Determinism is but one type of causal constraint, one member of the causal family. Its subtle relations with other constraints will be explored in detail in the coming chapters.

LOCALITY. Although in many contexts, the concept of determinism is taken to be synonymous with that of causality, there are also contexts—in particular, the context of the special theory of relativity (STR) and its relation to quantum mechanics (QM)—where it is the term *locality* that is typically used interchangeably with *causality* (or *relativistic causality*). A descendant of the traditional "no action at a distance" constraint, as well as the earlier *Natura non facit saltum* principle, locality is a constraint that excludes spatial or temporal gaps in physical interaction. The idea underlying the term *locality* is that changes in the physical world follow local "instructions" from the immediate environment rather than instantaneous ones from distant locations. Satisfying the desideratum of locality is one of STR's advantages over Newtonian mechanics, which involved instantaneous gravitational interaction between distant masses. The fact that both *determinism* and *locality* are used interchangeably with *causality* may lead us to assume that these terms are closely related, or at least coextensional. As we will see in chapter 4, however, locality and determinism are distinct concepts that figure in various intricate relations in different theories. Their interrelation is particularly intriguing in the framework of QM, where entangled states exhibit nonlocal correlations that have been alleged to pose a threat to QM's compatibility with STR. To preempt this threat, the relativistic constraint of locality has been narrowed down to no-signaling. That is, nonlocal correlations are legitimate as long as they do not allow the transmission of information between distant (though correlated) events. We will see in chapter 4 that indeterminism is the key to peaceful coexistence between QM and STR.

STABILITY. A stable state is a state to which a system tends to return after having been slightly perturbed. Stability might be the phenomenon we seek to explain: explaining the stability of atoms, for example, was one of the problems that led to the discovery of quantum mechanics. But stability is also an important explanatory notion adduced to understand the prevalence of one type of state, say equilibrium, over another type of state known to be erratic or short-lived. Unlike determinism and locality, the notion of stability does not rest on classical intuitions about causation. This might reflect the fact that, despite its explanatory import, stability does not constitute a general causal constraint. Depending on various factors, such as the nature of the relevant boundary conditions and the kind of perturbation involved, the same laws are compatible with the existence of both stable and less stable states. A system obeying deterministic laws can thus reach stable or unstable states, and the same is true of stochastic systems. Stability must therefore be carefully distinguished from determinism. In chapter 2, I argue that the conflation of these concepts, which is not uncommon, leads to serious blunders, and in particular, to imputing teleology to nonteleological processes. A better understanding of the notion of stability can serve to obviate teleology in a variety of contexts: history, evolutionary theory, mechanics, and statistical mechanics. The terminology used in these contexts may differ from that used in physics. Analysis of the concept of stability will therefore be accompanied by explication of related notions such as necessity, contingency, robustness, and resilience, all of which suffer from vagueness and ambiguity. The notion of stability is also invoked to elucidate the relationships between different physical levels, quantum and classical mechanics, classical and statistical mechanics. Exploration of the concept of stability is thus edifying vis-à-vis debates over reduction and emergence, examined in chapter 7.

CONSERVATION LAWS. That some physical quantities are conserved, whereas others are not, can explain why certain interactions are commonly observed, and others, never encountered. Like determinism and locality, conservation laws are constraints on possible change, and as such, they articulate our understanding of causation. The belief that nature allows neither genuine creation nor annihilation originated in antiquity; it is expressed in principles such as *nil posse creari de nihilo* and *causa aequat effectum*. In face of the experience of change, proponents of these ideas sought to uncover underlying constituents of reality that remained constant. The ultimate explanation of change, on this approach, is that change is only apparent. Among modern thinkers, Emil Meyerson is notable for advocating a kind of Parmenidean view on which change is illusory and "identity constitutes the essence of our understanding" ([1908] 1930, 402). Even when change is not altogether denied, it is generally believed to be constrained by some parallelism between earlier and later states, between cause and effect, between input and output. Descartes, who discovered (an early version of) the conservation of linear momentum, asserts:

> Now, it is manifest . . . that there must at least be as much [reality] in the efficient and total cause as in the effect of that cause. For where, I ask, could the effect get its reality from, if not from the cause? And how could the cause give it to the effect unless it possessed it? ([1641] *Meditations* III: 40; 1985 2: 28)[27]

Modern science has elaborated on these rudimentary intuitions about conservation in various ways. Classical mechanics led to the dis-

27. The translation (1985) is based on the original Latin text published in 1641; the brackets in this edition indicate insertions from the French version published three years later.

covery of the conservation of energy and linear and angular momentum, and QM has added further conservation laws. (Note that a theory can be indeterministic, like QM on the standard interpretation, and still impose strict causal constraints through its conservation laws.) In view of the fact that conservation laws are rooted in traditional ideas about causality, it is not surprising that the term *causality* has been used to refer to the applicability of conservation laws. For Niels Bohr, *causality* means the conservation of energy and momentum. In his oft-repeated claim that causal descriptions and spatiotemporal descriptions are complementary (that is, the accuracy of their joint application is restricted by Heisenberg's uncertainty relations), the term *causal description* should be understood in this way (and not, for instance, as connoting determinism). Explaining complementarity, Bohr states: "We have thus either space-time description or description where we can use the laws of conservation of energy and momentum. They are complementary to one another. We cannot use them both at the same time" ([1928] 1985, 6: 369).[28] Conservation laws and symmetries are inseparable members of the causal family. The causal function of conservation laws therefore also has bearing on the causal function of symmetry principles.

SYMMETRIES. Physicists place symmetry principles, which constrain the form of lower-level laws and guide theory construction, at the top of the hierarchy of physical laws. Symmetry considerations appear to be backed by a priori reasoning that resembles mathematics rather than physics. Their epistemic status is thus a matter of controversy. The connection with conservation laws, however, suggests that, to the extent that conservation laws are empirical laws that flesh out the causal structure of the world, so are symmetries. Although the connection between symmetries and conservation laws had been recognized earlier, it was proved by Emmy Nöther, who showed that, under a wide range of conditions, every continuous symmetry is correlated with a conserved quantity (Nöther 1918). In some cases, the

28. Quantum phenomena such as crossing a potential barrier seem to violate the conservation of energy and momentum. But ascribing definite energy and momentum values to the crossing particle would preclude its localization in space and time, hence we cannot "catch" it in the act of violation. This generates the complementarity Bohr invokes here.

connection between particular symmetries and other causal constraints is obvious. In the framework of STR, for instance, the principle of relativity, which is a symmetry principle, and the limit on the speed of signal transmission, which is a causal constraint, are closely linked. In other cases, gauge symmetries in particular, the connection is less obvious, and even debatable. I argue in chapter 5 that as a rule, symmetry principles function in the same manner as other causal constraints, and illustrate this claim by examining Pauli's exclusion principle. The connection between causation and symmetry is also conspicuous in Curie's principle, according to which symmetries manifested by a cause are inherited by its effects.

VARIATION PRINCIPLES. These principles single out the specific trajectories taken by physical systems. They determine, for instance, that light moves along the trajectory that takes the least time, that a particle follows the trajectory of least action, and that a freefalling body moves along a geodesic. Like symmetry principles, variation principles have a privileged status—they too are considered to be among the most general constraints on the form of theories. At first glance, variation principles appear to be teleological, and were indeed seen, when first discovered, as a demonstration of divine wisdom and benevolence. Over time, the teleological interpretation of these principles has given way to a causal understanding. Nevertheless, vestiges of the purposive impression seemed to linger. I will argue that, surprisingly, it was only in the context of quantum mechanics that the futility of the teleological interpretation could finally be established.

The foregoing list of causal constraints in physics introduces the constraints that will be examined in the coming chapters; it does not purport to be exhaustive. In addition, let me note two constraints that will not be thoroughly examined.

ASYMMETRY OF THE CAUSAL RELATION. Like determinism and locality, asymmetry is often considered an essential characteristic of the causal relation, and thus often referred to as "causality."[29] At the same

29. Frisch (2014) concentrates on this aspect of causation.

time, causal asymmetry has been contested on various grounds, especially its incompatibility with the fundamental laws of physics. The problem of whether and how causal asymmetry is related to temporal asymmetry is also much debated. Despite its centrality in the intuitive picture of causation, the asymmetry condition must be added "manually" in some of the leading accounts of causation, for example, the regularity and probabilistic accounts. On my pluralistic approach, the need to add this asymmetry to the other members of the causal family does not pose a problem. Moreover, causal asymmetry can be posited when focusing on individual processes and ignored when considering the general constraints imposed by conservation laws, symmetry principles, and variation principles. As they have no built-in asymmetry, these constraints play a causal role in controlling change, but not in controlling its direction. This tolerant strategy, I contend, is methodologically apt. Tolerance would be inappropriate, however, were the objection regarding the incompatibility between causal asymmetry and the fundamental laws of physics valid. But is it valid?

The incompatibility argument draws on the time-reversal symmetry of the fundamental laws of physics. A law is said to be time-reversal symmetric if whenever it allows a trajectory from event *c* to event *e*, it also allows the time-reversed trajectory from *e* to *c*. A common analogy is a film played backward: under time-reversal symmetry, we are unable to tell which film represents the actual course of events and which is the reversed film depicting a fictitious (though possible) course of events. The argument against causal asymmetry is that under the regime of time-reversal-symmetric laws, there is no observable difference between the two evolutions, and thus no reason to deem some events causes and others effects.[30] Consider a transition from an event *c* to an event *e*, and the following questions:

1. Do the fundamental laws allow us to retrodict the occurrence of *c* from the occurrence of *e* in the same way that they enable us to predict the occurrence of *e* from the occurrence of *c*?

30. When asymmetry is taken to be constitutive of causation, the argument targets the concept of causation tout court.

2. Do the fundamental laws allow the time-reversed transition from the occurrence of *e* to the occurrence of *c*?
3. Is there a sense in which *c* caused *e* but *e* did not cause *c*?

These questions are, in my view, distinct. Let us first consider questions (1) and (2). Laws that are deterministic and time-reversal symmetric yield an affirmative answer to both these questions, but this does not mean that the questions are equivalent. Had the laws been deterministic but not time-reversal symmetric, they would not necessarily sanction the reversed process, but could still allow retrodiction. On the other hand, under conditions of utter randomness, time-reversal symmetry could obtain despite the failure of prediction and retrodiction. As far as the incompatibility argument is concerned, however, the crucial point pertains to the relation between the first two questions and the third. From the affirmative answer to questions (1) and (2), the incompatibility argument concludes that question (3) must be answered in the negative. That is, it contends that if the time-reversed process can be predicted and is in fact allowed, there is no reason to take *c* to be the cause of *e* rather than take *e* to be the cause of *c*. But why should the laws' time-reversal symmetry exclude cause–effect asymmetry in the individual case? Over the last two weeks, I lost 3 pounds, but it would also have been possible for me to gain 3 pounds. (Indeed, given precise information about my diet and energy expenditure, these changes could have been predicted.) Does this mean that there is no fact of the matter as to what actually happened? Losing and gaining weight is a complex macroprocess involving much more than the fundamental laws of physics, but in principle, the point also applies to microprocesses. The fact that the laws of physics *allow* a process to unfold in opposed directions is compatible with the fact that, on any particular occasion, only one of these possibilities is realized. As it stands, therefore, the incompatibility argument, popular though it seems to be, does not refute causal asymmetry. More direct support for this asymmetry can be drawn from the discussion of Curie's principle in chapter 5.[31]

31. See Hemmo and Shenker (2012a) for an argument that anchors temporal asymmetry in the concept of velocity, and hence in fundamental physics.

MUTUALITY OF CAUSAL RELATIONS. Newton's third law states that if a body *a* exerts a force *F* on another body *b*, then *b* in turn exerts on *a* a force –*F* equal in size and opposite in direction to *F*. This law is violated in some physical theories (for instance, by the electromagnetic force), and is certainly not generally accepted outside physics. In philosophy, the idea of mutual action has sometimes been expressed more vaguely, requiring that if *a* can causally affect *b*, it must also be possible for *b* to causally affect *a*. Such stipulations appear, for example, in debating the mind-body problem. As I said, though, they are rarely encountered in science. A notable exception is Einstein's argument in support of the dynamic spacetime of the general theory of relativity (GTR). "It is contrary to the mode of thinking in science to conceive of a thing (the space-time continuum) which acts itself, but which cannot be acted upon" (Einstein 1922, 55–56). Here the mutuality constraint motivates the most revolutionary aspect of the new theory. According to Einstein, Newtonian mechanics violates our causal intuitions, for it allows space to act on matter, but does not countenance the reverse action, that is, the action of matter on space. GTR, according to which spacetime is shaped by the distribution of matter, while also determining this distribution, corrects this deficiency.[32] Einstein's use of the concept of causality in this context is somewhat idiosyncratic, but illustrates the fecundity of a concept of causality that is richer and more varied than the thin notion of causation debated in the philosophical literature.

This introduction has outlined the motivations for the book as a whole. Each chapter is largely self-standing, with the occasional slight overlap. Chapter 2 analyzes the determinism-stability relation as manifested in everyday contexts; chapter 3 analyzes it as manifested in physics. Both chapters show that the notions of determinism and stability are often conflated, giving rise to teleological thinking. Chapter 4 focuses on the

32. Note that this requirement of mutual causal influence does not involve the identification of individual cause-events and effect-events. As mentioned in note 13 above, DiSalle (1995) challenges Einstein's causal interpretation of the relation between matter and spacetime.

relation between determinism and locality, particularly in the context of quantum mechanics, where subtle payoff relations between these constraints are manifested. Chapter 5 examines symmetry principles and conservation laws. It illustrates how symmetry principles—despite their a priori appearance—function as causal constraints on a par with other members of the family of causal concepts. The "least action" principle is explored in chapter 6, which returns to the illusion of teleology, arguing that only within the QM framework is the principle's teleological appearance finally dispelled. Chapter 7 uses some of the results reached in earlier chapters to examine the relations between different levels of causality. It discusses reduction, emergence, and the intriguing possibility of lawless events in a deterministic world.