The Underdetermination Debate: How Lack of History Leads to Bad Philosophy

Wolfgang Pietsch[1], Philosophy of Science, Carl von Linde-Akademie, Technische Universität München, Germany

Abstract:
Two distinct versions of underdetermination have shaped the debate on this issue during the 20th century. The first can be attributed to Pierre Duhem, the eminent historian of physics; the other is due to W.V.O. Quine, who by contrast showed little interest in the history of science. My aim in this essay is to trace in their respective renderings of the underdetermination thesis the influence of their different stances towards the history of science and to establish that the historical perspective really makes a difference in that Duhem's version of the thesis is more plausible as well as more useful for understanding the scientific enterprise. Towards the end of the paper, several somewhat novel arguments for underdetermination will be developed that are informed by the historical perspective.

1. Introduction

Over the course of a century, the debate on underdetermination has produced an abundance of versions of the thesis that evidence does not uniquely determine scientific theories. Almost everybody agrees that some weak transitory underdetermination is a historical reality while several strong renderings are clearly implausible. Thus, the real challenge of the debate consists in formulating the underdetermination thesis in a way that strikes the right balance between the extremes. Such a formulation reaches beyond the trivial observation that theories are underdetermined if relevant evidence is missing. It should be methodologically useful both for the working scientist and for the historian of science while evading the common objections.

We will show in this essay that as a guideline to a philosophically viable conception of underdetermination the historical perspective proves to be essential. Conversely, attempts at an ahistorical formulation of the thesis will lead to a distorted notion of underdetermination that is open to all kinds of objections. The story will rely on a simplified construal of the debate on underdetermination in the 20th century. It will contrast the views of Pierre Duhem and W.V.O. Quine, respectively the most prominent proponents of a historical and an ahistorical rendering of underdetermination. We will show that the common objections against underdetermination are fatal only to Quine's version. Thus, in this case historical ignorance indeed led to bad philosophy.

Ultimately, the underdetermination thesis should be able to make sense of those episodes in the history of science, where a situation of underdetermination was diagnosed by at least some of the leading scientists involved. Examples are plenty, of which I will mention three. One that is much discussed concerns the transition from Euclidean to non-Euclidean physics at the turn from the 19th to the 20th century. There was a fairly short period, when several of the leading geometers and physicists of the time, including Hermann von Helmholtz, Henri Poincaré, and Albert Einstein, held the choice between the different axiomatizations of geometry to be of conventional nature, each axiomatization corresponding to a different formulation of the laws of physics (e.g. Helmholtz 1870, Poincaré 1902, Einstein 1921). The demand for simplicity of those laws would somewhat narrow down sensible choices of the axiomatization of geometry.

Another often-cited example is the equivalence between wave and matrix mechanics in the early years of quantum theory as pointed out in particular by Erwin Schrödinger: "Considering the extraordinary differences between the starting-points and the concepts of Heisenberg's quantum mechanics and of the theory which has been designated 'undulatory' or 'physical' […] it is very

---

[1] pietsch@cvl-a.tum.de

strange that these two new theories agree with one another with regard to the known facts, where they differ from the old quantum theory. […] That is really very remarkable, because starting-points, presentations, methods, and in fact the whole mathematical apparatus, seem fundamentally different." (Schrödinger 1926, p. 45) As Schrödinger stresses, matrix mechanics emphasizes the discontinuous nature of matter, while wave mechanics emphasizes the continuous aspects. Consequently, wave mechanics is naturally formulated in terms of differential equations, while matrix mechanics introduced its own peculiar algebraic language into physics. As in most historical cases of underdetermination, the exact interpretation of this situation has been quite controversial both among the scientists directly involved and later on among historians and philosophers of quantum theory (Muller 1997, Perovic 2008).

While a considerable philosophical literature exists on the equivalence of matrix and wave mechanics and even more on the conventionality of geometry, another pertinent example has evaded almost completely the attention of modern philosophers of science.[2] Underdetermination occurs in the transition from Newtonian action-at-a-distance theories to field theoretic formulations in classical physics, most notably in the case of electrodynamics (Pietsch 2010 and unpublished). Reading Maxwell's *Treatise on Electricity and Magnetism*, nothing short of being the *Principia* for classical electrodynamics, most modern readers will be surprised that a large part of the preface is devoted to an analysis of the relation between Faraday's field electrodynamics and the action-at-a-distance electrodynamics of Coulomb or Ampère. Maxwell emphasizes the equivalence of both theories, not only in terms of empirical adequacy, but also of non-empirical epistemic criteria: "In a philosophical point of view, moreover it is exceedingly important that two methods should be compared, both of which have succeeded in explaining the principal electromagnetic phenomena, and both of which have attempted to explain the propagation of light as an electromagnetic phenomenon, and have actually calculated its velocity, while at the same time the fundamental conceptions of what actually takes place, as well as most of the secondary conceptions of the quantities concerned, are radically different." (Maxwell 1873, p. xii)

My criteria for formulating the 'correct' version of the underdetermination thesis are of pragmatic nature, that it is of methodological importance for actual scientific theorizing and that it proves a useful tool for reconstructing certain episodes in the history of science. Of course, it should also evade all objections that have been raised against the thesis. With an eye on actual scientific practice both in its historical and methodological dimensions, the underdetermination thesis that we will embrace will be largely in Duhemian spirit. By contrast, the logical and linguistic remarks in Quine's rendering will turn out largely irrelevant to our formulation of the thesis. Basically, the historical version of underdetermination that we are eventually going to defend is the following: *In the history of science, especially of physics, we repeatedly encounter situations where several theories are equally strong in terms of their empirical consequences and with regard to epistemic virtues but rely on different metaphysics that provide the scientists with different instructions what to do next and what to expect from nature. There are no convincing reasons to exclude the possibility of such situations in the future.* None of the known objections is fatal to this formulation of the underdetermination thesis, while it nevertheless holds important implications for scientific method.

Section 2 will provide an overview of the most important versions of the underdetermination thesis. In Section 3, we will compare Duhem and Quine's versions of underdetermination and link the differences to their respective attitudes towards the history of science. The attitude towards history can also serve as a guideline for detecting further differences beyond those that are usually discussed in the literature. In the next step of the argument, it will be pointed out in Section 4 how the neglect of historical perspective has been detrimental to the underdetermination thesis since the common objections are fatal only to the ahistorical view. Section 5 will briefly summarize from the

---

2 Complaints about an alleged lack of examples are quite frequent in the literature (e.g. Norton 2008, p. 25). However, the fact that the underdetermination in classical electrodynamics has barely been discussed suggests that examples were searched for in the wrong place and that there are many more to be unearthed from the history of science (see also Section 3, item iv).

literature the two most important arguments in favour of the underdetermination thesis and point out their weaknesses. Concluding from Sections 4 and 5 we are confronted with a stalemate, where neither the arguments for nor those against underdetermination turn out ultimately convincing. Section 6 will try to ameliorate this situation by searching for further arguments in support of underdetermination that are informed by the historical perspective. We will show that denying underdetermination can seriously hinder progress in certain instances of the evolution of science. Also, denying underdetermination leads to an implausible conception of scientific theories, in which conventional elements can be unequivocally separated from non-conventional elements. The relevance of a historical perspective for thinking about scientific method will be briefly addressed in Section 7.

2. Underdeterminations

More than a century after the publication of the *Aim and Structure of Physical Theories*, Duhem's *opus magnum* on the methodology of physics, the often heated debate concerning underdetermination is difficult to disentangle. Various renderings of the underdetermination thesis are around and more than once critics have pointed out that some fairly uncontroversial version is being defended while far-reaching conclusions are then drawn from another much more dubious version (e.g. Laudan 1990, p. 324). Conversely, defendants of underdetermination might argue that all objections concern only some distorted rendering of the underdetermination thesis, while the real thing stays undefeated. In this vein, we will argue in this paper.

In any case, it is crucial to distinguish the different versions of the underdetermination thesis that have been proposed in the literature. The common denominator in all of them is that *in some manner evidence underdetermines theory*. Various claims result from specifying the three central concepts, i.e. underdetermination, evidence, and theory. The most fundamental distinction concerns the methodological toolbox by which underdetermination is established—resulting in two generic types of underdetermination (Laudan, 1990): *deductive underdetermination*, sometimes also referred to as Humean underdetermination, and *ampliative underdetermination*. In case of the former, underdetermination is established by means of a pure hypothetico-deductive method, where theories are evaluated solely on the basis of their observable, deductive consequences. By contrast, ampliative underdetermination takes into account further criteria. Thus, ampliative underdetermination is deductive underdetermination plus underdetermination with respect to X, where X might be some non-empirical epistemic virtues like simplicity, fruitfulness, scope or consistency.[3] For genuine inductivists, X would incorporate sophisticated inductive methods, e.g. eliminative induction or Bayesian probabilism.

Another distinction arises when specifying the extent of evidence that is taken into account. Most importantly, do we refer to the actual evidence in a specific historical context, or to all possible evidence, i.e. all observation statements implied by a theory. Both approaches have been proposed in the literature: the former goes by the name *transient underdetermination,* the latter is often called *permanent underdetermination* (Stanford 2009; see also Sklar 1975, pp. 380-381). In the first case, the underdetermination predicament usually goes away when further evidence accumulates; in the second, no evidence can decide between rival theories. Sometimes, transient underdetermination is taken to imply that there necessarily exists a piece of evidence which will decide between competing approaches. We will not require that. In our usage of the terms, a case of transient underdetermination, i.e. with respect to actual evidence, can also constitute a case of permanent underdetermination, i.e. with respect to possible evidence.

Transient underdetermination has recently been examined extensively in the work of Kyle Stanford (2001, 2006, 2009). Stanford has formulated a novel epistemic challenge to scientific method, his so-called *problem of unconceived alternatives,* which might also be termed the problem of *recurrent transient underdetermination*. His claim is that underdetermination can be a challenge

[3] A comment on notation: To adequately mirror the distinction between deductive and ampliative underdetermination, I distinguish between empirical adequacy and non-empirical epistemic virtues.

to realism not only if it is permanent but also if it is transient, while recurrent. If there continue to be currently unthought-of alternatives even to our best-established theories then underdetermination clearly constitutes a challenge to scientific realism. According to Stanford, the history of science provides ample evidence that the threat of unconceived alternatives is real.

Finally, one can distinguish versions of underdetermination by specifying what exactly one understands by theory. A conservative proponent of underdetermination will insist on some, however fuzzy dividing line between observation statements, theoretical hypotheses, and analytic statements. Given this assumption, underdetermination concerns only some esoteric ‘isles’ of interconnected empirical hypotheses in abstract scientific theories. By contrast, a more radical proponent might deny any meaningful distinction between observation statements, theoretical hypotheses, and analytic statements. From this perspective, one will quickly conclude that not only some restricted areas, but *all* our knowledge is underdetermined. While familiar from the literature, these brands of underdetermination have not yet been given concrete names. We will call them *isolated* and *ubiquitous* underdetermination. Further distinctions have been proposed, but the mentioned three are the most important ones or at least the most relevant for our purposes.

Following this systematic exposition let us now locate Duhem and Quine on this matrix. For this purpose we rely on Duhem’s *The Aim and Structure of Physical Theory*, especially chapter IV of part one[4], and on *Two Dogmas of Empiricism*, the *locus classicus* for Quine’s version of underdetermination. Duhem turns out a careful advocate of ampliative underdetermination taking into account both non-empirical epistemic virtues, which he summarizes as economy of thought, and inductive methods, whose effectiveness he criticizes. To account for actual theory choice in cases where rational criteria fail, Duhem proposes his theory of good sense. For him, theories are evaluated according to empirical adequacy, non-empirical epistemic virtues and inductive support. He clearly leans toward transient underdetermination in that he generally considers the historical context with a specific situation of evidence. This can be most clearly perceived from the fact that all his examples are drawn from the history of science, especially from physics. Finally, he clearly endorses isolated underdetermination by denying that phenomenological sciences like physiology are affected by underdetermination at all. According to Duhem, underdetermination concerns only theoretical hypotheses in advanced (“symbolic”) scientific theories.

Quine is in many ways more radical than Duhem. He is also less clear about what his views really are, at least in the influential Quine (1951). For example, he wavers between deductive and ampliative underdetermination. Still, most of his examples indicate that he has the deductive rendering in mind, e.g. when he writes: “Even a statement very close to the periphery [of the web of knowledge] can be held true in the face of recalcitrant experience by pleading hallucinations or by amending certain statements of the kind called logical laws.” (1951, p. 43) Clearly, underdetermination is established here in terms of observable deductive consequences only. This impression is also supported by his radical denial of the analytic-synthetic distinction, rendering both epistemic virtues and inductive methods revisable and thus ultimately incapable for deciding between competing approaches. In principle, for Quine even deductive logic is open to revision, leading to an extreme relativism where everyone can claim whatever pleases him. Detailed discussions of scientific method are remarkably absent from Quine’s main writings on underdetermination (i.e. 1951 and 1975). There is no critique of inductive methods and nothing comparable to Duhem’s theory of good sense.[5] At least from his most explicit formulations, the reader gains the impression that Quine (1951) leans toward the more radical options of permanent underdetermination (“Any statement can be held true come what may”; p. 43) and ubiquitous underdetermination (“The totality of our so-called knowledge or beliefs, from the most casual matters of geography and history to the profoundest laws of atomic physics or even of pure mathematics and logic, is a man-made fabric which impinges on experience only along the edges”;

[4] The better known chapter VI of part two is much less telling about Duhem’s stance concerning underdetermination.

[5] Quine elsewhere discusses the epistemic virtues simplicity, familiarity, scope, and fecundity (1955). However, his elaboration remains rather detached from scientific practice mixing examples from everyday life with some superficial expositions of examples from physics.

p. 42)

3. Chasing Duhem, fleeing Quine

In this part of the argument we will show how the main differences between Duhem and Quine's versions of underdetermination derive from their respective stances toward history. We will dub Duhem's viewpoint the historical and Quine's the logical rendering of underdetermination. Even a superficial comparison between Duhem's *The Aim and Structure* and Quine's *Two Dogmas of Empiricism* reveals the fundamentally different perspectives. Duhem's text is rich with examples from the history of physics, while Quine—"indifferent as Americans often are concerning history" (Vuillemin 1979, p. 598)—indulges in fundamental considerations of logic and language that remain detached from any historical reality of science. Of course, this corresponds to their philosophical heritage: Duhem was a working physicist and an acclaimed historian of physics, while Quine was a logician and philosopher of language.

It is quite straightforward why someone with the historical interests of Duhem should impose careful and sensible limitations on the underdetermination thesis. Summarized from the last section, the main issues that distinguish Duhem's account from Quine's are[6]: i) the limited extent of holism; ii) that only theories involving symbolic representations are affected; iii) the invocation of *good sense* to account for actual theory change; iv) that the thesis is illustrated by means of historical examples; v) that the thesis is framed in terms of actual evidence pertaining to a specific historical situation; vi) that complete empirical equivalence is not necessarily required, but rather equal strength in terms of empirical adequacy and non-empirical epistemic virtues like simplicity, fruitfulness or coherence; vii) that underdetermination has immediate implications for the methodology of physics.

Let us now point out in a somewhat tedious exercise, how each of these differences can be traced back to Duhem's historical and Quine's logical outlook. i) A crucial distinction concerns the extent of holism that underpins the underdetermination thesis. Confirmational holism means the idea that scientific hypotheses are not vulnerable to experiments in isolation, but only as a group (Duhem 1991, pp. 183-188). Whenever a prediction turns out wrong, there is considerable ambiguity which hypotheses to abandon. In the words of Duhem, "the validity of [the scientist's] conclusion is as great as the validity of his confidence [in the accuracy of all other propositions he has used in addition to the examined hypothesis]" (ibd., p. 185). While Duhem advances a holism that is restricted to limited groups of hypotheses within physics, Quine's holism comprises everything from mathematics and logic to the purely phenomenological sciences: "The unit of empirical significance is the whole of science." (1951, p. 42)

How is this difference connected with the presence or absence of historical perspective? While it may in principle be true that all knowledge is connected via logic and mathematics, the implications of Quine's holism are far too unrealistic for any serious scientist or historian of science to accept. A physicist at CERN formulating a hypothesis about the Higgs boson will hopefully never be worried by the fact that you are reading my paper in this very moment, although these two statements are in principle connected via the laws of logic. An astronomer concerned with the mechanics of the solar system will hopefully never dare to question the principles of mathematics, if he finds Mercury to behave unexpectedly. Quine's unlimited holism bears no insight as to why, in the past, certain hypotheses were rejected and others kept.

By contrast, Duhem's holism is historically accurate. It is a historical truism that unexpected experimental results in physics never lead to prompt and unanimous conclusions by the scientists involved. Rather, "[n]o absolute principle directs this inquiry, which different physicists may conduct in very different ways without having the right to accuse one another of illogicality" (Duhem 1991, p. 216). Often dissenters remain who do not accept the consensus reached by the majority. While Quine's holism can account for the existence of dissenters in principle, it is

---

6 The differences between a Quinean and a Duhemian rendering of underdetermination are addressed in among others Veuillemin (1978), Ariew (1984), Quine (1986), and Gillies (1993).

inaccurate about where, when, and to what extent dissent surfaces. For example, unexpected trajectories of Mercury will never produce dissenters concerning the principles of deductive logic. Duhem's holism can account for that, Quine's cannot.

ii) Somewhat relatedly, Duhem restricts the range of underdetermination to theories that involve symbolic representations and do not reason directly on facts. Accordingly, only mature theories involving an abstract mathematical layer that is not accessible to direct observation are affected by underdetermination. Only when observation requires sophisticated scientific instruments, we encounter the kind of theory-ladenness that eventually leads to holism and underdetermination. Thus, underdetermination will occur in physics but not in physiology (Duhem 1991, p. 180). By contrast, Quine (1951) famously frames his discussion in an outright denial of the analytic/synthetic distinction per se, which for him constitutes one of the dogmas of empiricism. He thus renders the distinction between theoretical and observational statements, between symbolic representation and statements of facts largely meaningless. To be fair, Quine does introduce a distance-measure indicating how far away a statement is from the sensory periphery (1951, p. 43). But his wording suggests that for him the distinction between observation statements and theoretical hypotheses is only a matter of degree and not of qualitative nature. It therefore cannot yield the conceptual basis for restricting underdetermination to specific areas in the web of experience.

It is not difficult to connect this difference with Duhem's interest and Quine's disinterest in the history of science. Someone interested in a historically adequate account of physical methodology must allow for a distinction between observational and theoretical statements that is sufficiently robust to ground qualitative differences between them. From the pragmatic, historical point of view it would just be outrageous to treat in the same manner highly abstract theoretical statements, e.g. concerning the properties of quarks and strings, and pure observation statements like 'The needle of my measuring device points to 5'. Contrary to Quine, this is not just a matter of degree.

iii) The historical fact that in spite of underdetermination most physicists eventually agree on the implications of experiments is accounted for by Duhem's theory of 'good sense'. It has sometimes been suggested that good sense should be understood entirely in terms of inductive methods and of non-empirical epistemic virtues. Such an interpretation must be rejected since Duhem insists that good sense is not rationally reconstructable; rather, it refers to "reasons which reason does not know" (1991, p. 217), "that confused collection of tendencies, aspirations, and intuitions" which cannot be further analysed and which cannot be rigorously formulated (ibd., p. 104). Presumably, Duhem is aware of the large gap between actual practice in physics and toy models of scientific inference as discussed in philosophy of science. While Duhem acknowledges the utility of inductive frameworks, his holism points to the ambiguities of these frameworks and to subjective elements involved.

Duhem's account of good sense, as unsatisfying as it may be, is historically accurate, since it accounts for the eventual convergence of opinions, while nevertheless acknowledging that there is no universally accepted inductive method in physics and that leading scientists like Einstein have repeatedly stressed the power of intuitions and creativity in the development of scientific theories. On the other side, it is not surprising that Quine with his lack of interest in history does not come up with a concept comparable to good sense.

Issues i) to iii) are frequently cited in comparisons between Duhem and Quine (e.g. in Gillies 1993). However, once it is understood that the historical perspective is crucial, further differences can be detected which are otherwise easily overlooked. iv) One rather obvious aspect concerns the choice of examples that are used to illustrate the underdetermination thesis. Duhem's examples are all taken from the history of physics, while Quine (1951) starts an unfortunate chain of examples that are construed from existing theories involving redefinition of terms, abandonment of logic, hallucinations and the like: Brutus may not have killed Caesar, if 'killed' happened to have the sense of 'begat' (p. 36); an allegedly failed prediction can be held true if pleading hallucination or changing the laws of logic (p. 43).[7] None of these ‘examples’ will convince working scientists that underdetermination actually constitutes an interesting epistemological problem with relevance for scientific practice. Admittedly, Quine’s examples have an advantage as well, in that they can be constructed starting from our currently best theories. By contrast, Duhem’s examples are today largely outdated—in fact most of them already were in Duhem’s days. Working with Duhem’s examples, one could argue that underdetermination is only a problem for science in its immature stages—irrelevant to our modern theories. So, why didn’t Duhem worry? Presumably, he was illustrating a methodological point about certain epistemic situations that can arise in the evolution of physics. Implicit in his choice of examples is the claim that such situations may always recur in the future. This shows Duhem to be one of the ancestors of Stanford’s recurrent transient underdetermination leading to the problem of unconceived alternatives (Stanford 2001, 2006, 2009).

v) Another obvious consequence from the historical perspective is that underdetermination must be considered with respect to the actual evidence in a specific historical context. For the historian the most interesting situations are those, where theories are equally confirmed by past evidence but through different metaphysics provide different research programs for the future. By contrast, Quine insisted on rendering underdetermination in terms of possible evidence, abstracting from a specific historical situation: “natural science is empirically under-determined […] not just by past observation but by all observable events” (Quine 1975, p. 313). Duhem’s historical viewpoint is largely indifferent concerning the question if some future evidence might decide between competing approaches. Plausibly, given the commitment and the ingenuity of the scientists involved, the approaches are often potentially equivalent with regards to future evidence.

vi) In actual historical contexts, one should not expect exact empirical equivalence between rival theories, not even with respect to past evidence. For example, different approaches will deviate from each other in the periphery, the domain of application may be somewhat different, or they will differ with regard to the accuracy with which they can describe certain phenomena. In actual historical contexts, competing theories will be largely on a par in terms of epistemic virtues and inductive support, but they won’t necessarily be empirically equivalent. By contrast, Quine prefers the logically clean formulation of empirically equivalent theories with respect to possible evidence. In general, actual history fails to provide such rivals.

vii) Finally, there is a difference in aim: Duhem’s historical perspective mirrors an interest in scientific method that can already be deduced from the title *Aim and Structure of Physical Theory* and that is lacking in Quine’s account. If Duhem succeeds in making sense of certain historical episodes by employing the underdetermination thesis, this will also reveal the function of underdetermination for the methodology of physics. In this way, historical case studies of underdetermination can provide methodological insights that are relevant for contemporary physics as well.

In summary, we have seen in this section that all differences between Duhem’s and Quine’s renderings of underdetermination can be traced back to an interest or lack of interest in the history of science, respectively.

4. All objections refuted

In the next step of the argument we will now show that the familiar objections against underdetermination from the literature are fatal only to Quine’s logical rendering of the thesis and not to Duhem’s historical formulation. The historical perspective thus turns out essential for formulating a defensible version of the underdetermination thesis. Conversely, the lack of historical perspective has opened up the underdetermination thesis to easy criticism. Two objections are

---

[7] Quine (1975) later corrects his stance and denies that these are genuine examples of underdetermination (cp. also Laudan 1990, pp. 332-335). He demands that underdetermination “needs to be read as a thesis about the world” (1975, p. 324) concerning theories that are “equally sustained by all experience, equally simple, and irreconcilable by reconstrual of predicates” (ibd., p. 328). By that time, however, Pandora’s box was already open.

usually given particular weight: the argument from an impoverished account of scientific method, which has most prominently been voiced in Larry Laudan's influential *Demystifying Underdetermination* (1990, p. 346), and the identical rivals objection (Norton 2008; cp. also Quine 1975, Magnus 2003, Frost-Arnold & Magnus 2009). We will briefly discuss some further objections against the underdetermination thesis that play a more minor role in the literature: Grünbaum's point that there is no general argument proving the existence of alternative theories, Laudan and Leplin's claim that empirical equivalence is contextual, and finally the objection from scientific import.

According to the identical rivals objection, in all alleged cases of underdetermination we are actually dealing with different formulations of one and the same theory. John Norton puts it in the following way: "The very fact that observational equivalence can be demonstrated by arguments brief enough to be included in a journal article means that we cannot preclude the possibility that the theories are merely variant formulations of the same theory." (2008, p. 17) The proponent of the historical rendering need not be worried since a closer look reveals that this attacks only the logical version of underdetermination. First, Norton evokes equivalence with respect to possible evidence rather than actual evidence. Also, he requires only that both theories are observationally equivalent and not that they are on a par regarding non-empirical epistemic virtues. Finally, good examples of historical underdetermination are decidedly those that can*not* be demonstrated—in full detail—within a journal article. Rather, such rival approaches differ in all important aspects, in particular ontology, mathematical framework as well as experimental practices (cp. the quotes by Schrödinger and Maxwell in the introduction). As Stanford rightly points out genuine examples of underdetermination require "the sort of difficult conceptual achievement that demands the sustained efforts of real scientists over years, decades, and even careers" (2006, p. 15)—an achievement that cannot be laid out in a few paragraphs.

The crux of evaluating the identical rivals objection lies in the quest for good criteria, in which situations we are dealing with variant formulations of theories and in which we are not (Magnus 2003). Norton's somewhat ingenious suggestion—referring to the possibility of formulating the equivalence within a journal article—is quite helpful as a first guess, but clearly fails as a reliable criterion. Quine proposes as a criterion that the theories can be translated by means of a reconstrual of predicates, i.e. essentially by a redefinition of terms (1975, p. 320). However, this is neither necessary nor sufficient. It is not sufficient, since it ensures only empirical equivalence but not equivalence with respect to epistemic virtues. It is not necessary, since even in reformulations sometimes interesting shifts in meaning can occur. For example, a remarkable shift in the fundamental ontology occurs when Euclidean geometry is reformulated in non-Euclidean terms. It is not altogether clear why we should speak of equivalent theories in this case (Magnus 2003, p. 1258).[8]

While there may not be a universal criterion for the identity of theories, the historical perspective can provide some insights. Genuine cases of underdetermination are those in which the theories are potentially equivalent in terms of their empirical consequences but also with regard to epistemic virtues, but where they differ enough to provide the scientists with different outlooks on the world suggesting different research programs. In summary, the identical rivals objection does not expose serious problems of the underdetermination thesis, but disqualifies those ahistorical examples where the competing accounts are easily mapped onto each other, in particular Quine's suggestions referring to redefinition of terms, hallucination and the like.

Taking examples from the history of science also immunizes the underdetermination thesis against the second main objection, namely that the underdetermination thesis allegedly presupposes an impoverished hypothetico-deductive account of scientific method. This objection belongs to the standard repertoire of arguments against underdetermination and has been most prominently voiced in Laudan (1990, p. 346). Laudan correctly emphasizes that deductive underdetermination does not

8 Magnus (2003, p. 1263) suggests that there is a sufficient condition for non-identity of theories, namely empirical inequivalence. However, this criterion falls prey to Laudan and Leplin's point about the contextuality of empirical equivalence to be discussed below.

warrant ampliative underdetermination. The underlying claim is that underdetermination is usually established on the grounds of observable deductive consequences only, leaving aside all the wonderful inductive tools belonging to more sophisticated accounts of scientific method. Surely, deduction cannot tell us which of the several assumptions that enter into the derivation of a false prediction is wrong. However, sophisticated inductive accounts like eliminative induction or Bayesian probabilism are supposedly able to provide such advice, or so the objection goes.

It is certainly correct to criticize the lack of detailed discussions of inductive methods in most arguments for underdetermination. Still, the objection does not succeed against the historical rendering of underdetermination that we are defending here. Most importantly, if the historical examples would really fall prey to this objection, it would imply that scientists like James Clerk Maxwell, William Thomson, or Erwin Schrödinger—all of whom acknowledged underdetermination at some point—worked with an impoverished scientific method. This would clearly be an absurd consequence given that the work of such leading physicist should provide role models for scientific method. We can therefore conclude that the objection from an impoverished account of scientific method does not affect the historical rendering of underdetermination while strongly discrediting the strict logical rendering which indeed relies on equivalence of observable consequences alone.

Of course, several accounts of induction exclude underdetermination to some degree, as for example Laudan (1990, p. 332) and Norton (2008, pp. 29-32) show in some detail. But both seem to be fighting windmills here, at least if one assumes Duhem to be one of the targets. Nowhere in *The Aim and Structure* does Duhem claim that when a prediction turns out wrong, we are given free choice, which of the assumptions to abandon that entered the derivation of the prediction. Rather, Duhem admits that often scientists will readily agree on this issue. He also credits induction with a fruitful role for scientific inference: “induction may indicate to some extent the path leading to certain hypotheses” (1991, p. 259). The essential disagreement between Duhem and critics of underdetermination like Norton consists not in the fact that one allows for inductive methods and the other does not. Rather, Duhem is much more sceptical than Norton concerning the reach of induction: “[n]o system of hypotheses can be obtained by experimental induction alone” (ibd.). Duhem offers a general critique of inductive methods, examining Newton and Ampère’s claims that they deduced their theories uniquely from experience (pp. 190-200). While Norton seems to imply that scientists’ choices are always rationally reconstructable in terms of inductive methods, Duhem insists that there are elements involved which defy a full rational reconstruction. Such elements Duhem subsumes under his notion of good sense.

We are now on much more elusive grounds. Eventually, an opponent of underdetermination would have to show that (i) induction as used in physics is indeed *fully* formalizable, and also (ii) that the various inductive methods always lead to the same results in the same situations, or alternatively that only one of these inductive methods is correct. Furthermore, he would have to prove (iii) that there are no subjective elements involved in the inductive process which could easily destroy any consensus between scientists, e.g. he would have to exclude such accounts of induction as subjective Bayesianism.

Neither Norton nor Laudan’s survey of inductive methods establishes any of this. In the end, there is a large gap between showing that some toy models of scientific inference do not exhibit underdetermination and establishing that underdetermination is not an issue for real science. That scientific inference in the real world is a much trickier business can be deduced from the fact that no working scientist has ever been able to come up with a full-fledged account of scientific method. And those that have tried to develop a universal inductive method like Francis Bacon or John Stuart Mill have eventually been ridiculed for their alleged naiveté. Also, some of our most esteemed scientists like Albert Einstein have repeatedly stressed the role of intuitions or even of a decisively creative element in the development of physical theories.

Still, the objection from an impoverished account of scientific method exposes a weakness in the holist argument for underdetermination which is embraced both by Quine and by Duhem and which will be discussed in more detail in the next section. It constitutes therefore an objection against one of the main arguments in support of underdetermination rather than against the underdetermination thesis itself. Certainly, undermining an argument for a thesis does not necessarily refute the thesis. In the same vein, Laudan acknowledges that there may be ampliative underdetermination even though "[t]he fact that a theory is deductively underdetermined [...] does not warrant the claim that it is ampliatively underdetermined" (1990, p. 346).

This point turns out closely related to an objection against underdetermination that has been voiced prominently by Adolf Grünbaum. He stresses that from the impossibility of an *experimentum crucis* one cannot deduce the necessary existence of alternative formulations: "Duhem *cannot* guarantee on any general logical grounds the deducibility of O [empirical findings] from an *explanans* constituted by the conjunction of H [empirical hypothesis] and some revised *non*-trivial version R of A [auxiliary assumptions]." (Grünbaum 1960, p. 75) Granted, but neither can one exclude the existence of alternative formulations. Grünbaum's argument seems to imply a stalemate between proponents and opponents of underdetermination.

Let us briefly address some further worries that have sometimes been voiced against underdetermination but which once again turn only on the logical rendering. The notion of empirical equivalence, which is central to the logical view, has rightly been criticized by Laudan and Leplin, who argue in an influential paper that empirical equivalence is "both contextual and defeasible" (1991, p. 454). Indeed, judgments of empirical equivalence have sometimes been revoked, when the context changed. For example, while in the end of the 19th century Euclidean and non-Euclidean geometry were deemed empirically equivalent, many hold today that physical geometry has been empirically proven to be non-Euclidean. As pointed out in the last section, underdetermination requires the commitment of the scientists to develop competing frameworks in a way that they remain empirically equivalent. Thus, the contextuality of empirical equivalence is just a further argument for rendering underdetermination with an eye to the historical context.

Another objection which is frequently voiced by working scientists could be termed objection from scientific import. If fully equivalent theories existed, then science would not lose much if we just chose an arbitrary one and discarded all others. Maybe, underdetermination with full empirical equivalence would still be interesting for the metaphysician, who could learn from underdetermination that different metaphysics are able to account for the same observations. According to this view, underdetermination could serve as a mine for arguments in favour of or against metaphysical realism while the working scientist would just shrug his shoulders and reply that the debate is of little import for the practice of science. However, this picture is mistaken about the notion of empirical equivalence which we have just seen to depend on the continuous commitment of the scientists. It also misconstrues the role of metaphysics for science. Metaphysics is not irrelevant for scientific research; rather it serves as a necessary and indispensable guideline for the inductive business of science. It leads the scientist to consider what theoretical problems to tackle next or which experiments to do.

In the example of underdetermination between field theory and action at a distance in classical electrodynamics, each perspective had its merits linked to the specific metaphysics of the approaches. The field ontology led to the discovery of many phenomena that concern the 'medium' between charges and currents, for example the theories of dielectrics and diamagnetism (Faraday) or the unification of optics with electrodynamics (Maxwell). On the other hand, the action-at-a-distance ontology proved fruitful for finding the Newtonian force laws governing electrostatics and magnetostatics (Coulomb) or the unification of electrostatic and electrodynamic interaction (Weber). Progress in electrodynamics would have been seriously hampered by an exaggerated dogmatism concerning ontology as well as by an outright denial of underdetermination in scientific method.

In summary, underdetermination is really about equally strong theories with different metaphysics that provide the scientist in a specific historical context with different instructions what to do next and what to expect from nature. If this historical perspective on underdetermination is presupposed then none of the known objections from the literature is fatal to the underdetermination thesis.

5. 'Arguments' for underdetermination

The plausibility of the underdetermination thesis suffers less from the objections leveled against it, which were shown not to be very persuasive, but from the weakness of the arguments for it (cp. also Norton 2008, pp. 21-26). In the following we will outline both Duhem and Quine's defense of underdetermination. Both authors rely heavily on confirmational holism. A second strain of arguments for underdetermination is of (meta-)inductive nature enumerating examples. Finally, a third strain will be developed in the next section. There, we will argue that a denial of underdetermination leads to a crippled account of scientific method. We will show that in essential periods of scientific evolution at least some scientists must acknowledge underdetermination (as Maxwell and Schrödinger in the quotes from the introduction) while a universal denial of underdetermination would seriously block progress. Also, we will show that a denial of underdetermination implies an unrealistic conception of scientific theories, misconstruing the roles of conventions and research ideals.

The most widely accepted argument for underdetermination relies on confirmational holism. It can be found clearly articulated in the argument against *crucial experiments* of Duhem's *Aim and Structure* (1991, p. 183-190). According to Duhem, in abstract sciences like physics it is impossible to test hypotheses without the help of auxiliary assumptions, which usually include the basic tenets of several physical theories. For instance, a hypothesis about the nature of light can be verified or falsified only by presupposing some core assumptions of optics, thermodynamics, and mechanics that are used in the construction of the respective scientific instruments and in the design of the experimental set-up. Consequently, the test of a hypothesis is only as good as the confidence in those other tenets. In principle, one can always hold on to any hypothesis and in the case of recalcitrant evidence blame some of the auxiliary assumptions. As Duhem states provocatively physicists could for example have saved the particle nature of light in spite of Foucault's experiment, if they had only attached some value to this task (p. 187).

The holist argument also builds the backbone of Quine's defense of underdetermination. The *locus classicus* is *Two Dogmas of Empiricism*, in which he famously attacks the analytic/synthetic distinction as well as the reductionist claim that statements can be confirmed or disconfirmed in isolation. This leads him directly to confirmational holism, from which underdetermination is then derived. According to Quine, changing a statement in the interior of the conceptual net of a theory can always be compensated by other adjustments in the interior, leaving the edge of the net, representing the empirical results of the theory, unchanged. Of course, this involves the assumption that such compensatory adjustments are always possible for which a rigorous argument is missing. In a later essay, Quine points out that mature scientific theories always encompass extra 'stuffing' which is not determined by the infinite conjunction of relevant observation statements. Even by the best standards of rationality, a certain freedom of choice remains implying underdetermination (1975, p. 324).

While the holist argument is widely accepted, it clearly relies on a purely hypothetico-deductive method and therefore only establishes deductive underdetermination, i.e. identity of the observational consequences (cf. the argument from an impoverished account of scientific method in the last section). Regarding the more interesting ampliative underdetermination, the argument can provide good reasons neither for nor against it (Laudan 1990, p. 346). Unfortunately then, there seems to be no shortcut to establishing ampliative underdetermination.

Do we have to resort to the (meta-)inductive justification of underdetermination on a case-by-case basis? Proponents of underdetermination should point out a sufficient number of examples in order to conclude that underdetermination is something everyone should worry about. By contrast, opponents of the thesis would have to show that these examples are few and that the few are trivial or belong to immature theories. In any case, the inductive argument for underdetermination (if it is to establish ampliative underdetermination) involves detailed historical studies, a profound knowledge of the respective sciences and is in general beyond the reach of a single philosophy paper. Another weakness of inductive justification is that it can establish underdetermination in general only for the narrow realm of the special science from which the examples are taken. Given the vast differences in outlook and methodology, it is not clear why an inductive justification in, say, biology should immediately carry over to physical theories.

Ultimately, the inductive argument results once again in a stalemate between opponents and proponents of underdetermination, where one side will claim that there are a large number of examples for underdetermination and the other side will deny these examples on a piecemeal basis. Proponents will produce historical evidence that theories will always be troubled by underdetermination. Opponents will try to establish that mature theories are not implicated. Given the difficulties in determining what exactly a mature theory is, the prospects of both an inductive justification or a piecemeal confutation of underdetermination are rather dim. Godfrey-Smith (2008) has pointed out a related symmetry of arguments for and against underdetermination, where supporters will claim that there are always several theories that can account for a given body of evidence, while opponents will respond that for every two rival theories there always exists some piece of evidence that can discriminate between them.

Taking stock, the holist and the inductive arguments do not provide sufficient grounds for establishing the more interesting ampliative underdetermination.

6. Arguments from history and scientific method

In the previous two sections we have detected a lack of good arguments for and against underdetermination. The common objections turned out either to be directed against a misconstrued version of the thesis or to concern only arguments for the thesis. On the other side, neither the holist nor the inductive argument proved particularly strong for establishing underdetermination. Thus, we are confronted with a situation, where the arguments are neither strong enough to establish nor to discount the underdetermination thesis.

In view of this situation, the present section will be devoted to developing additional arguments in order to make a more convincing case for underdetermination. The main trick will consist in a change of perspective away from treating underdetermination as a crucial argument in the realism-antirealism debate towards considering the implications of underdetermination for scientific method. I will present arguments to the effect that denying underdetermination leads to a crippled scientific method. First, a methodology excluding underdetermination would hinder progress in crucial episodes during the evolution of science. Second, a denial of underdetermination would imply an implausible role for conventions and research ideals in scientific theories. These methodological arguments for underdetermination are all informed by the historical perspective.

a) Underdetermination and scientific progress

In this section, we argue that progress in science would be seriously hindered if some scientists would not allow for underdetermination. Throughout the history of science we encounter repeatedly episodes where leading scientists have acknowledged underdetermination—as in the quotes by Maxwell and Schrödinger from the introduction. We will see that their attitude contributed considerably to the progress in their respective scientific discipline.

First of all, proponents of underdetermination are in a good position for integrating virtues of competing programs. Their creative work in building coherent theories is much facilitated by the fact that they have at their disposal a plethora of possibly useful analogies between rival approaches. By contrast, an opponent of underdetermination will insist on there being only one true theory, rendering pointless any elaboration of analogies between different programs. If only one of the approaches eventually tells the true story and if by consequence all other programs are wrong, why should one expect scientific progress from developing competing programs in parallel and adapting them to each other. Even if a combination of several programs will turn out the correct theory, the elaboration of analogies should not be helpful since where one theory is right the other(s) must necessarily be wrong. Thus, accepting underdetermination is a necessary premise for fruitfully exploiting analogies between competing programs.

Furthermore, establishing underdetermination is a helpful manoeuvre to facilitate the transition between different paradigms during a scientific revolution. It enables a comparably smooth and undogmatic paradigm change that allows proponents of unsuccessful approaches to save face to a certain extent. After all, given underdetermination, the acceptance of a new paradigm does not render wrong what someone once believed in and preached to students.

Let us illustrate these considerations by means of the rivalry between action-at-a-distance and field electrodynamics (Pietsch 2010 and unpublished). The acknowledgment of underdetermination allowed William Thomson and James Clerk Maxwell to develop analogies between both frameworks and thereby contribute to progress in electrodynamics. As an example, the electric potential φ and the vector potential $\boldsymbol{A}$ were developed to be effectively employed in both programs, while however designating different things. In field theory they describe the state of a continuous entity, while in action at a distance they describe the relation between particles. As Olivier Darrigol writes: "Thomson forged multi-purpose concepts that transcended cultural barriers and individual theoretical preferences [...] Physicists conversant with French [action-at-a-distance] electrostatics could easily express the potential in terms of electric fluid densities. The followers of Faraday's views, if any, could draw the lines perpendicular to the equipotential surfaces and call them force lines." (2000, p. 136) Also, the eventual transition from the dominant action-at-a-distance paradigm to the outsider view of Faraday would have been much more difficult without a previous demonstration of some equivalence between the approaches. Similar arguments can be given for almost any case of historical underdetermination, e.g. Schrödinger's assessment of equivalence between matrix and wave mechanics.

If physicists like Maxwell or Schrödinger had not acknowledged underdetermination, they would have dogmatically insisted on their preferred formulation—as admittedly many scientists did in the same situation, for example Faraday on the field view or Heisenberg on matrix mechanics. Nobody would have invested in developing the links between the different approaches and thus a split in the scientific community would most probably have ensued and would have taken up much intellectual esprit in essentially pointless debates. In short, if noone would have been prepared to take the stance of underdetermination, progress in physics would have been seriously hindered.

Admittedly, these considerations do not establish that scientists should necessarily and always endorse underdetermination. Rather, a scientific method *excluding* underdetermination might have its heuristic merits as well, as for example Kuhn emphasizes: "[The] invention of alternates is just what scientists seldom undertake except during the pre-paradigm stage of their science's development and at very special occasions during its subsequent evolution. So long as the tools a paradigm supplies continue to prove capable of solving the problems it defines, science moves fastest and penetrates most deeply through confident employment of those tools. The reason is clear. As in manufacture so in science—retooling is an extravagance to be reserved for the occasion that demands it." (1996, p. 76) In a sense, we are confronted with two different modes of conducting scientific research, one allowing for and the other denying underdetermination. Each has its respective merits and therefore its *raison d'être*. However, if the possibility of underdetermination is excluded in total, progress in science will at some point be severely hindered.

b) Conventions and underdetermination

An intimate relationship between conventional elements in scientific theories and the underdetermination thesis has been pointed out by several authors (e.g. Brown 1989, pp. 50-56). Conventions are not determined empirically but rather by pragmatic considerations, they are in many ways relative to us who decide on them. Thus, there exists no unique, correct choice for conventions. Since, broadly speaking, different choices of conventions yield different theories, scientific theories are underdetermined with respect to these choices. An argument for underdetermination results, if we can show that scientific theories contain *non-trivial* conventional elements leading to non-trivial underdetermination.

For the opponent of underdetermination, two different responses to this argument are feasible. First, he could argue that scientific theories do not necessarily contain conventional elements.[9] Second, he could claim that the kind of underdetermination that results from conventional choices is trivial in the sense that it can easily be detected and isolated. To counter these rejoinders, we will now show that every scientific theory necessarily contains conventions, e.g. in the units of measurable quantities, in fundamental constants and in symmetries and invariances, and that it is far from trivial to determine conventional elements within scientific theories.

Over the course of history, metrology as a scientific discipline (not met*eo*rology!) has involved some of the finest minds in science. Metrology is concerned with measuring in general; it is the science that defines the basic units of all those quantities necessary for the description of the physical world. In every major country there is at least one large research institution that is exclusively concerned with metrology, e.g. the National Institute of Standards and Technology in the United States or the Physikalisch-Technische Bundesanstalt in Germany. No doubt, metrology deals mainly with conventions, e.g. that we measure length with meters, time with seconds, mass with kilograms etc. In a way, the existence of metrology institutes with a large scientific staff and with enormous budgets is already proof that the fixing of conventions is far from trivial.

There is obviously no shortage of conventional elements in scientific theories, at least when they incorporate measurable continuous quantities. It is beyond question that scientific theories are underdetermined with respect to the choice of the fundamental units of those quantities. For example, physical geometry is underdetermined with respect to the choice of foot or meter as the fundamental unit of length. Of course, an opponent of underdetermination will be quick to point out that such underdetermination is trivial and therefore cannot have the profound methodological implications that we attributed to the underdetermination thesis.

Everything depends on the question if conventional elements can always be identified and isolated as easily as in the case of meter vs. foot. If so, then the opponent of underdetermination can rightly claim that the choice of units only leads to trivial examples. However, the prospects of such an endeavour seem rather dim. After all, the conventional choice of a measure for a continuous quantity does not always concern a trivial factor between different units as in: 1 meter = 3.28 feet. One can imagine much more complicated relations, for example a length measure being a complex function of the position. This possibility cannot be ruled out a priori for reasons of simplicity. After all, whichever measure one regards as fundamental, the other is complex in relation to it.

As first pointed out by Hermann von Helmholtz (1870), the physical equations have to be adjusted in order to compensate for different choices of spatial measure. Thus, when we consider more complicated choices, the conventional element cannot be easily isolated and separated from presumably non-conventional elements like the fundamental laws. Indeed, the very formulation of the fundamental laws depends on the exact choice of measure. Eventually, a spatial measure should be chosen that renders the physical equations as *simple* as possible. Consequently, the underdetermination due to conventional elements is limited by considerations of simplicity. However, simplicity is a far too malleable concept to fully rule out underdetermination arising from the choice of spatial measure.

An illustration of the complex issues involved in the choices of measures for physical quantities can be found in Hasok Chang's admirable book *Inventing Temperature*. Chang terms epistemic iteration the complex process how a reliable measure for temperature is constructed in parallel with scientific progress in the theory of heat. Chang's historical case study provides extensive evidence that the choice of measure for fundamental quantities cannot be isolated from the theory itself, that conventional elements cannot be isolated from non-conventional elements. Certainly, Chang does not embrace a "simplistic type of conventionalism" which would allow for

---

9 Brown's conventionalist twist has been criticized in Okasha (2000, pp. 289-290) on the grounds that conventionalism is "just one possible response—and one which, with the exception of Poincaré, has usually appealed more to philosophers than to scientists". This is historically incorrect. Geometric conventionalism was in various shades accepted by a considerable number of leading physicists at the turn from the 19th to the 20th century—among them von Helmholtz (1870) and Einstein (1921).

arbitrary choices of measures. Rather, considerations of simplicity narrow down the choices. However, it is highly unlikely that simplicity can single out a unique true theory. In any case, no convincing argument has ever been given in that respect. The sophisticated conventionalism of Poincaré is closely related with the coherentism advocated by Chang. While arbitrary choices of conventions are ruled out, there is no ultimate empirical justification for the choice of measures for fundamental physical quantities (Chang 2004, p. 223).

Fundamental constants also involve conventional elements that are often tied to the choices of measures of fundamental quantities. Fundamental constants provide another example how difficult it is to determine the boundary between conventional and empirical elements further establishing that conventions cannot be easily isolated and that therefore underdetermination resulting from conventional choices is not trivial. Consider the velocity of light c. The exact status of this constant continues to be debated. Following the establishment of special relativity, some have held c to simply result from the erroneous assumption that space and time are distinct concepts. Rather, we supposedly live in a unified space-time where space and time should be measured with the same units and consequently the velocity of light should be one. Others have questioned this viewpoint insisting on the conceptual distinctness of space and time. Then, the velocity of light is of much more pronounced empirical nature, for instance it may even undergo change over time.[10]

There are further somewhat less tractable conventions in scientific theories. Consider for example the three quantities force $\boldsymbol{F}$, mass m, and acceleration $\boldsymbol{a}$ connected via Newton's relation $\boldsymbol{F} = m\,\boldsymbol{a}$. Several interpretations are feasible. First, we could interpret the equation as providing a definition for one of the quantities $\boldsymbol{F}$, m, or $\boldsymbol{a}$. In this case only two of the quantities are independently measurable. Whichever two we take to be fundamental depends on a conventional choice that is not necessitated by empirical facts. This being an ontological convention concerning attributes of fundamental entities, we are thus led to ontological underdetermination. Essentially, the underdetermination of action-at-a-distance and field electrodynamics is a sophisticated case of such an ontological underdetermination that results from a reformulation of the corresponding physical theory and a different choice about what is considered fundamental and what is not. Second, we could interpret the equation $\boldsymbol{F} = m\,\boldsymbol{a}$ as an empirical law connecting three quantities that are independently measurable. However, it is a tricky question how one can establish that there exist independent measurement procedures for quantities connected by a deterministic law. As of today, there seems to be no common agreement on the exact status of $\boldsymbol{F} = m\,\boldsymbol{a}$, whether it is an empirical law or a definition and, in case of the latter, which quantity is being defined. Once again, the boundary between conventional and empirical elements is quite blurred.

A further argument for the existence of non-trivial conventional elements in scientific theories can be based on symmetries and invariances. Every invariance of a scientific theory directly implies a conventional choice. Once again, trivial examples are readily available. Homogeneity of space implies the conventional choice of the origin of the coordinate system in physical geometry. Isotropy of space implies the conventional choice of the direction of the axes of the coordinate system. Gauge invariance in classical electrodynamics implies a certain conventional choice of the potentials φ and $\boldsymbol{A}$. Lorentz invariance in the theory of relativity implies the conventional choice of the reference system, i.e. the velocity of the observer. The historical fact that invariances and symmetries have often been debated, or that invariances that were long thought trivial have eventually been drawn into question, as left-right mirroring in particle physics, points to the fact that the conventions implied by symmetries and invariances are not trivial in the sense that they can be easily isolated.

[10] A variable speed of light has been proposed in various physical contexts, most notably in cosmology. For example, Arnot (1941) suggests it to account for the Hubble expansion, a possibility discussed also by Popper (1940). In the more modern literature, a varying speed of light has been proposed to solve various problems of big bang theory, for example the flatness, the horizon, and the cosmological constant problems (e.g. Barrow 1999). The literature on changing natural constants in physics is rich and controversial. Over the years it has involved some of the finest minds in science; cp. for example Dirac's widely-discussed suggestion of a varying gravitational constant (1937).

In general, the structure of the bare theory itself does not allow us to determine which parts are conventional and which parts are empirical. Rather, we also have to take into account how propositions are treated by the people working with them. Often, a proposition becomes a convention only if it is treated as such. This is the reason why without detailed historical stories like Chang's exposition of the 'invention' of temperature, the argument for underdetermination based on conventions cannot be explicated.

c) Research ideals and underdetermination

If one allows that the broader metaphysical world-view can influence scientific theorizing and if metaphysical pluralism is accepted in the sense that several world-views are possible, then underdetermination is a plausible consequence. Metaphysical propositions which exert influence on scientific theorizing will be called research ideals in the following.

Consider as an example the issue of determinism vs. indeterminism, i.e. the question whether every event is determined by events prior to it or not. Arguably, this is not an empirical question. After all, to prove determinism one would have to show that every event is determined by some other events prior to it. Given the finiteness of our experience, such an endeavour is surely impossible. To prove indeterminism one would have to show that there are events which are not fully determined by prior causes. Given that there are arbitrarily many candidates for such prior causes and again the finiteness of our experience, indeterminism cannot be established either. Only if further restrictions are admitted, e.g. locality, determinism and indeterminism become empirically distinguishable.

Different choices of research ideals can eventually lead to underdetermination. For example, the choice between determinism and indeterminism lies at the root of the underdetermination between orthodox non-relativistic quantum theory and Bohmian mechanics, the former based on an indeterministic and the latter on a deterministic metaphysics. Other cases of underdetermination originating in different choices of research ideals are readily available, e.g. concerning the choice between a fundamentally discrete or a continuous nature of matter. The underdetermination between field and action-at-a-distance electrodynamics falls broadly into this category. As Schrödinger (1926, p. 45) points out, matrix and wave mechanics also originate in different conceptions of matter, the former stressing the discontinuous and the later the continuous aspects.

An opponent of underdetermination would have to deny the role sketched for research ideals in this section. He would either have to question that metaphysics exerts an influence on science or claim that there is only one correct metaphysics. Both standpoints are implausible if judged from the history of science.

7. Conclusion

The historical perspective provides a reliable guide to formulating the underdetermination thesis. It helps to avoid some common distortions of the thesis which have in the past opened it up to the usual objections. If history is taken as a guide then underdetermination will automatically be construed in a way that it can account for certain episodes in which underdetermination was explicitly acknowledged by leading scientists. The historical perspective will also allow to comprehend in which ways acknowledging underdetermination can enable scientific progress. In this way, underdetermination turns out a powerful tool both for the historian to make sense of certain episodes in the history of science and in the hands of the working scientist to enable progress in certain contexts. In a slight variation of a well-known quote by Duhem[11] we conclude that to give the history of underdetermination is at the same time to make a (methodo-)logical analysis of it.

---

[11] „To give the history of a physical principle is at the same time to make a logical analysis of it." (Duhem 1991, p. 269)

Acknowledgements
I thank the editors of this volume, Tad Schmaltz and Seymour Mauskopf, for helpful comments and criticism. I am also grateful to Mauricio Suárez for suggestions on an earlier version of this article as well as to audiences at EPSA09 in Amsterdam and at &HPS2 in Notre Dame.

References

Ariew, Roger. 1984. The Duhem Thesis. *British Journal for the Philosophy of Science* 35: 313-325.

Arnot, Frederick L. 1941. *Time and the Universe*. Sydney: Australasian Medical Publishing.

Barrow, John D. 1999. Cosmologies with Varying Speed of Light. *Physical Review D* 59, 043515.

Brown, James R. 1989. *The Rational and the Social*. London: Routledge.

Chang, Hasok. 2004. *Inventing Temperature. Measurement and Scientific Progress*. Oxford: OUP.

Darrigol, Olivier. 2000. *Electrodynamics from Ampère to Einstein*. Oxford: OUP.

Dirac, Paul A. M. 1937. The Cosmological Constants. *Nature* 139: 323.

Duhem, Pierre. 1990. Logical Examination of Physical Theory. *Synthese* 83: 183-188.

Duhem, Pierre. 1991. *The Aim and Structure of Physical Theory*. Princeton, NJ: Princeton University Press. Translation of *La théorie physique. Son objet et sa structure.* Paris: Chevalier & Rivière, 1906.

Einstein, Albert. 1921. Geometry and Experience. In *Sidelights on Relativity*, 25-56. Mineola, NY: Dover, 2010.

Frost-Arnold, Greg and P. D. Magnus. 2009. The Identical Rivals Response to Underdetermination. *PhilSci archive*. http://philsci-archive.pitt.edu/archive/00003390/. Accessed 15 February 2010.

Gillies, Donald. 1993. *Philosophy of Science in the Twentieth Century.* Oxford: Blackwell. Excerpt reprinted in *Philosophy of Science. The Central Issues*, ed. M. Curd and J. A. Cover, 302-319. New York: Norton, 1998.

Godfrey-Smith, Peter. 2008. Recurrent Transient Underdetermination and the Glass Half Full. *Philosophical Studies* 137 (1): 141-148.

Grünbaum, Adolf. 1960. The Duhemian Argument. *Philosophy of Science* 27 (1): 75-87.

Helmholtz, Hermann von. 1870. Über den Ursprung und die Bedeutung der geometrischen Axiome. In *Populäre Wissenschaftliche Vorträge von H. Helmholtz. Drittes Heft*, 21-54. Braunschweig: Vieweg und Sohn. 1876.

Kuhn, Thomas. 1996. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Kukla, André. 1996. Does Every Theory Have Empirically Equivalent Rivals? *Erkenntnis* 44: 137–166.

Laudan, Laurens. 1965. Discussion: Grünbaum on “The Duhemian Argument”. *Philosophy of Science* 32 (3/4): 295-299.

Laudan, Larry. 1990. Demystifying Underdetermination. In *Scientific Theories*, ed. C. Wade Savage, 267-297. Minneapolis: University of Minnesota Press. Reprinted in *Philosophy of Science. The Central Issues*, ed. M. Curd and J. A. Cover, 320-353. New York: Norton, 1998.

Laudan, Larry and Jarrett Leplin. 1991. Empirical Equivalence and Unterdetermination. *The Journal of Philosophy* 88 (9): 449-472.

Magnus, P. D. 2003. Underdetermination and the Problem of Identical Rivals. *Philosophy of Science* 70: 1256-1264.

Maxwell, James C. 1873. *A Treatise on Electricity and Magnetism.* Oxford: Clarendon Press.

Muller, F. A. 1997. The Equivalence Myth of Quantum Mechanics. *Studies in the History and Philosophy of Modern Physics* 28: 35-61, 219-247.

Norton, John D. 2008. Must Evidence Underdetermine Theory? In *The Challenge of the Social and the Pressure of Practice: Science and Values Revisited*, ed. M. Carrier, D. Howard, and J. Kourany, 17-44. Pittsburgh: University of Pittsburgh Press.

Okasha, Samir. 2000. The Underdetermination of Theory by Data and the 'Strong Programme' in the Sociology of Knowledge. *International Studies in the Philosophy of Science* 14 (3): 283-297.
Perovic, Slobodan. 2008. Why Were Matrix Mechanics and Wave Mechanics Considered Equivalent? *Studies in History and Philosophy of Modern Physics* 39: 444-461.
Pietsch, Wolfgang. 2010. On Conceptual Issues in Classical Electrodynamics: Prospects and Problems of an Action-at-a-distance Interpretation. *Studies in History and Philosophy of Modern Physics* 41 (1): 67-77.
Pietsch, Wolfgang. Unpublished. Transient Underdetermination and 19th-Century Electrodynamics.
Poincaré, Henri. 1902. *La Science et l'Hypothèse*. Paris: Flammarion.
Popper, Karl R. 1940. Interpretation of Nebular Red-Shifts. *Nature* 145-3663: 69-70.
Quine, Willard V. O. 1951. Two Dogmas of Empiricism. Reprinted in *From a Logical Point of View: 9 Logico-Philosophical Essays*, 2nd ed., 20-46. Cambridge, MA: Harvard University Press, 1980.
Quine, Willard V. O. 1955. Posits and Reality. Reprinted in *The Ways of Paradox and Other Essays*, 2nd ed., 246-254. Cambridge, MA: Harvard University Press.
Quine, Willard V. O. 1975. On Empirically Equivalent Systems of the World. *Erkenntnis* 9: 313-328.
Quine, Willard V. O. 1986. Reply to Jules Vuillemin. In *The Philosophy of W. V. Quine*, ed. Lewis E. Hahn and Paul A. Schilpp, 619-622. La Salle: Open Court.
Schrödinger, Erwin. 1926. On the Relation between the Quantum Mechanics of Heisenberg, Born, and Jordan, and that of Schrödinger. In *Collected Papers on Wave Mechanics,* 45-61. New York, NY: Chelsea, 1982.
Sklar, Lawrence. 1975. Methodological Conservatism. *Philosophical Review* 84: 384-400.
Stanford, P. Kyle. 2001. Refusing the Devil's Bargain: What Kind of Underdetermination Should we Take Seriously? *Philosophy of Science* 68: S1-S12.
Stanford, P. Kyle. 2006. *Exceeding Our Grasp.* Oxford: Oxford University Press.
Stanford, P. Kyle. 2009. Underdetermination of Scientific Theory. In *Stanford Encyclopedia of Philosophy (Winter 2009 Edition),* ed. Edward N. Zalta. http://plato.stanford.edu/archives/win2009/entries/scientific-underdetermination/.
Vuillemin, Jules. 1979. On Duhem's and Quine's Theses. *Grazer philosophische Studien* 9: 69-96. Reprinted in *The Philosophy of W. V. Quine,* ed. Lewis E. Hahn and Paul A. Schilpp, 595-618. La Salle: Open Court.