Title
Tea Leaves and Telescopes: Creating future fundamental theory in physics

Author
Schneider, Michael D.

Publication Date
2020

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Tea Leaves and Telescopes: Creating future fundamental theory in physics

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Philosophy

by

Michael D. Schneider

Dissertation Committee:
Professor James Owen Weatherall, Chair
Chancellor's Professor Jeffrey A. Barrett
Professor JB Manchak
Professor P. Kyle Stanford

2020

DEDICATION

To Hannah, wherever we go from here

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I thank Jim Weatherall, all my committee members, Hannah Rubin, Elliott Chen, Kino Zhao, and the Philosophy of Physics Reading Group at LPS, for their counsel, sanity checks, and camaraderie throughout the writing of this dissertation. I also thank George Smith— and, long before him, Christos Galanopoulos— for setting me along this path. Finally, I thank my parents and my sister: my earliest advisors, reviewers, editors, and audience.

In regards to more material things, I would like to thank the Templeton Foundation, the School of Social Sciences at UCI, and the Newkirk Center for Science and Society at UCI for their various grant funding, which allowed me to pursue my research over the past several years largely unencumbered. A good deal of that research is reflected in this dissertation (though not all).

I also should thank the journals Philosophy of Science and British Journal for the Philosophy of Science (Univeristy of Chicago Press and Oxford University Press, respectively) for their author's permissions policies that allow me to incorporate my published work into this dissertation (as chapters 2 and 4, and 3, respectively). And I thank the reviewers and editors associated with these journals, who provided invaluable feedback along the way.

Finally, I thank Otten (2012) for providing the core LaTeX template used to produce this final document, in accordance with the required formatting.

VITA

Michael D. Schneider

EDUCATION

Doctor of Philosophy in Philosophy, with Physics Emphasis **2020**
University of California (Logic and Philosophy of Science) *Irvine, CA*

Master of Arts in Philosophy **2018**
University of California (Logic and Philosophy of Science) *Irvine, CA*

Bachelor of Science in Physics and Philosophy **2015**
Tufts University *Medford, MA*

SINGLE-AUTHORED PUBLICATIONS

"Creativity in the social epistemology of science" **Accepted**
Philosophy of Science

"Betting on future physics" **2019**
British Journal for the Philosophy of Science

"What's the problem with the cosmological constant?" **2018**
Philosophy of Science

ABSTRACT OF THE DISSERTATION

Tea Leaves and Telescopes: Creating future fundamental theory in physics

By

Michael D. Schneider

Doctor of Philosophy in Philosophy

University of California, Irvine, 2020

Professor James Owen Weatherall, Chair

I consider how current theory in fundamental physics matters to the development of future theory in the discipline. As discussed in the Introduction (Chapter 1), my aim is to better understand how we may achieve transformational research in fundamental physics, particularly concerning quantum gravity, when it is our current theoretical understanding that nonetheless functions as our guide. Toward that end, Chapter 2 provides a methodology within contemporary quantum gravity research that is consistent with the focus given to the 'cosmological constant problem' in practice. The method I sketch there is contrary to a picture of theory development that is otherwise spelled out in terms of explorations of theory space. Instead, it centers the importance of specifying the technical machinery that is needed to articulate empirical problems for the future theory, in advance of that theory's development. Chapter 3 then argues that this method can be epistemically well founded. In the case of the historically standard version of the cosmological constant problem, the grounds for pursuing solutions to the problem in the course of quantum gravity research ultimately trace back to beliefs concerning what it is about our current, empirically successful theories that accounts for their successes. In Chapter 4, I then shift gears in order to argue against the applicability of a recent account of creativity in scientific discovery, which is spelled out in terms of the wide exploration of something very much like theory space. This picks up a thread from Chapter 2, which I expand upon in the Conclusion (Chapter 5). Ultimately,

the perspective on theory development supported by the method I articulate seems to better capture how, beginning with current theory, we may wind up with a future theory of quantum gravity that is sufficiently novel to accomplish all that we presently demand of it.

Chapter 1

Introduction

The modern proletarian class does not carry out its struggle according to a plan set out in some book or theory; the modern workers' struggle is a part of history, a part of social progress, and in the middle of history, in the middle of progress, in the middle of the fight, we learn how we must fight... That's exactly what is laudable about it, that's exactly why this colossal piece of culture, within the modern workers' movement, is epoch-defining: that the great masses of the working people first forge from their own consciousness, from their own belief, and even from their own understanding the weapons of their own liberation." (Rosa Luxemburg, "The political mass strike and the unions", 1910, common online translation)[1]

[1] *"Die moderne proletarische Klasse führt ihren Kampf nicht nach irgendeinem fertigen, in einem Buch, in einer Theorie niedergelegten Schema; der moderne Arbeiterkampf ist ein Stück in der Geschichte, ein Stück der Sozialentwicklung, und mitten in der Geschichte, mitten in der Entwicklung, mitten im Kampf lernen wir, wie wir kämpfen müssen.... Das ist ja gerade das Bewundernswerte, das ist ja gerade das Epochemachende dieses kolossalen Kulturwerks, das in der modernen Arbeiterbewegung liegt: daß zuerst die gewaltige Masse des arbeitenden Volkes selbst aus eigenem Bewußtsein, aus eigener Überzeugung und auch aus eigenem Verständnis sich die Waffen zu ihrer eigenen Befreiung schmiedet."* (Retrieved on 5 June, 2020 from https://www.projekt-gutenberg.org/luxembur/reden/chap032.html)

> *"We must be prepared to allow this 'Universe' the freedom to have contradictory properties, like we have been forced to grant to the atom, and in particular we must try to accustom ourselves to the idea that the velocity of change of the quantity of the dimension of a length, which occurs in the equations and is interpreted as the 'Radius of the Universe, has nothing to do with the rate of evolutionary change of stars and stellar systems."* (Willem de Sitter, in discussion at the centennial meeting of the British Association for the Advancement of Science in 1931)[2]

Now is a time of pandemic, anxiety, rage, and passion. Daily news reports provide a kind of smoldering, in vivo eschatology, which nonetheless manages to sublimate into optimism. Bail funds have gone mainstream. Mutual aid is trending. Prison abolition is under discussion. Activists across the country may actually be managing to defund the police. So, as I put together the dissertation enclosed here, my mind wanders. I find myself dwelling on other matters, and I return to the revolutionary fatalism that used to inspire me. I go searching for speeches delivered by utopians, who are, so often, Marxists.

Revisiting the passage by Rosa Luxemburg reproduced above, it occurs to me— with some degree of surprise— that, when I began this dissertation several years ago, I had a rather similar idea to hers in mind (at least, as regards my own subject of burgeoning expertise). In this way, the dissertation I have since completed might be considered a companion, or perhaps a distant relative, to Luxemburg's worldview.

The dissertation that has come out of that initial idea of mine— what is contained in the following pages— is focused, without apology, on the technical details of contemporary theoretical physics practice. I have been interested, in particular, in how considerations about our universe as a whole that stem from our current fundamental theories get to factor into the development of a future fundamental theory of quantum gravity. This is the sort

[2]Discussion reproduced on p. 573-610 in the archived proceedings, with the above quote appearing on p. 586.

of subject that risks being dismissed as minutiae, when a Black-led, multiracial rebellion chants “No justice; no peace”, and is met each evening with tear gas. But such a view is, I must continue to believe, myopic in the extreme. If it is my privilege that I may be optimistic in this regard, it is likewise my obligation to broadcast such optimism. So I say: there will be fruit found in the systematic reimagining of our physical world, right alongside the systematic reimagining of our social one. And, in the meantime, I am happy to let metaphors stream freely back and forth between such projects. (“We must try to accustom ourselves,” says de Sitter, to impossible, contradictory things...)

So it goes, then, that the kernel of my enclosed argument seems to me to depend on a very similar insight to that voiced by Luxemburg. Namely, in each of our cases, there is some collective vanguard, which, it is taken for granted, proceeds inevitably toward a more desirable future state. In this respect, the accounts are each utopian: that the procession is inevitable means that our absolute capacity to get to the more desirable future is not what is questioned. Instead, the focus is on the manner by which we may acquire, presently, the means to get there. And crucially, in each case, the vanguard is taken to succeed via a spontaneous act. The members of the vanguard reach forward into the future and, in doing so, proceed to pull themselves toward it.

It is this metaphorical act of reaching forward that intrigues me most, in the narrative. Because it seems to me that the means for doing so *can only ever be* that which is provided by the vanguard’s own history, as well as their understanding of the present, evinced by it. I am therefore left to conclude that the future is only ever built *out of* the current. But if that is so, *into* what do the members of the vanguard reach?

I cannot say much more in detail about Luxemburg’s case, nor am I suitable to speak more in detail about the currents of social change happening now. But I can say much more in detail about my case, at the frontiers of fundamental physics research.

To begin with, I think it is essentially correct to state that what follows in this dissertation turns on a single insight, which I have tried to consider from several angles. Namely: in the course of theory development at the cutting edge of fundamental physics, theoretical physicists have only the means provided by their current best theories, as we all presently have come to understand them: in terms of their empirical successes. Thus, to understand the relationships future physical theories will have both to the world and to us, it is imperative that we consider, with care, how those theories are to come to be, *out of where physicists, collectively, have happened to come so far.*

In other words, I have focused in this dissertation on a kind of sober-minded reasoning that may be said to operate at the brink of theoretical discovery— those moments in which we might otherwise be tempted to assume that mere whim, genius, or luck dominates all else. Contrary to that impulse, I have tried in earnest to consider what should constitute a theoretical physicist's best effort to develop a theory of quantum gravity out of current theory— that which simultaneously provokes their desire to develop such a new theory, in the first place.

But the project that I have described so far is unspecific as to its means. This is not true of the dissertation, which considers in explicit terms how a certain conflict between two grandiose narratives we have developed of the underlying structure of our universe comes to matter in quantum gravity research. The next two chapters stand as two independent projects related to this subject, which nonetheless build on each other, in turn.

In Chapter 2, I consider the weathered subject of the 'cosmological constant problem'. This problem is one of mutual interpretation, concerning what our universe would be like, were it left forever empty and unexcited, based (on one hand) on our current fundamental, *classical* theory of gravity and (on the other hand) on our current fundamental, *quantum* theories of particle physics. In the context of each of these, we feel entitled to certain understandings of what our universe as a whole would be like, in such a case (i.e. were that the end of the story). The trouble is that the two understandings obviously disagree with one another— so much so that it is difficult to imagine how they could disagree more.

Still, that we possess two understandings that are in disagreement with one another is not obviously a problem— our scientific worldview is filled with such disagreements. As I argue in the chapter, the interesting character of this particular disagreement is that it becomes an empirical problem— one in need of solution—, in virtue of various assumptions laid out about what will surely be the case, in the theory of quantum gravity to come. Of course, having made such assumptions, the problem is now one to solve, in due course. This observation paints out a certain method, which is plausibly descriptive of, and certainly suitable for, quantum gravity research.

In Chapter 3, I consider the epistemology behind such a method, as just described. The essential conceptual tool offered here is the notion of an anticipation about, e.g., future fundamental theory. Using, as a case study, a historically standard version of the cosmological constant problem, I argue that certain bets about the future theory of quantum gravity give rise to anticipations about it, which further shape the contours of ongoing research toward that future theory. As a consequence, the grounds for how the future theory comes to take its form thereby trace back to the reasons that may justify a theorist committing to such bets. Those reasons concern our current interpretations of our current best theories, given that we regard those theories— so understood— as empirically successful.

Finally, I turn in Chapter 4 to a more general question, which concerns the relationship between creativity and revolutionary theorizing. In particular, I argue against the applicability of a recent account of creativity in scientific discovery, in virtue of its failing to countenance the role creativity plays when revolutionary theorizing is important.

My entry into this debate is through remarks made by the quantum gravity theorist Carlo Rovelli, which I also refer to in Chapter 2. In the earlier chapter, I use those remarks as evidence of a metaphor, which I take to be commonly used in our understanding of theory development in fundamental physics: that of exploring theory space. In that chapter, I emphasize how the method I have outlined does not require the use of any such metaphor. But, plausibly, the appeal of the metaphor is supported by the account of creativity I critique in Chapter 4, in a manner that is more or less independent from my particular critiques. I pick up the discussion of this metaphor in the Conclusion. Ultimately, it seems to me that the perspective on theory development that does away with this metaphor, which is supported by the method I distinguish in the earlier chapters, is better poised to clarify how we may ever go about creating truly transformative research, in light of all that we have achieved so far.

Chapter 2

What's the problem with the cosmological constant?

2.1 Introduction

The 'Cosmological Constant Problem' (CCP) has occupied the attention of theoretical physicists for several decades, often glamorized in popular science as "the worst prediction in the history of physics". According to INSPIRE, the High Energy Physics online information system, Weinberg's 1989 paper that first formally declared it as a crisis has been cited over 4000 times.

Despite all of this activity, what is actually meant by the CCP is ambiguous. In most cases, the CCP concerns a tension between the measured value assigned to the cosmological constant ('Λ') in the standard model of cosmology and the values of certain quantities predicted in quantum field theory (QFT) that are usually thought to influence it. But the status of the CCP as a problem is made complicated by two observations. First, it is unclear by what criteria the quantities from QFT are considered to be relevant to the standard model of cosmology built on classical (i.e. not quantum) physics. Second, even if such quantities were made relevant, the current standard model of cosmology can accommodate them via an unconstrained parameter, called the 'bare cosmological constant' or, in context, the 'bare term' (c.f. Carroll (2001); Martin (2012)), usually denoted as Λ_0. As such, the CCP cannot be understood as a problem within the current standard model of cosmology. (Section 2.2 will formally develop this claim.)

The situation is quite different when one leaves behind present, well-evidenced physical theory and speculates about what a quantum theory of gravity will entail for the future of cosmology. But competing semiclassical intuitions about what our current theories of general relativity (GR) and QFT imply about future theories of quantum gravity make it unclear how to assess that conflict: different assumptions about what quantum gravity will ultimately entail give rise to different characterizations of the CCP. For instance, whether or not one believes that the future theory includes some feature responsible for a non-vanishing bare term Λ_0 in the classical gravitational limit influences what one expects ought to be so in regards to the matter limit of that same theory (in which gravitational effects are disregarded). Consequently, any proposed 'solutions' offered to the CCP are inevitably segregated into several pairwise incompatible categories that differ in their assumptions about what the future nature of the problem will be.

The immediate goal of the present chapter is to explain how it is that such incompatible theoretical proposals can all simultaneously count as plausible solutions to the CCP, when the CCP is itself understood as referring to multiple, mutually exclusive problems that only arise in mutually exclusive possible future contexts. As will be argued in section 2.4, these mutually exclusive 'versions of the CCP' contour a (perhaps surprising) amount of structure within quantum gravity research, suggesting a richer understanding of theory development than that which is conveyed by metaphors of principled explorations of some 'theory space'. Generalizing from the case of the CCP, I will argue that the primary function of certain 'problems', at least in theoretical physics, is to sketch out new avenues of theoretical research. Under this view, what constitutes a possible 'solution' to such a problem is any proposal for the next generation of theory that is well-suited to reproduce the virtues of the present theory and improve upon it in at least one precise regard: by providing an explanation for how it is that the given problem (newly articulable as a unique problem) is already handled in the new theory.

2.2 The 'Cosmological Constant Problem' is not a problem (for current theory)

This section ought to begin with a concise statement of the CCP, but to avoid introducing any basic misunderstandings, its statement first requires sufficient familiarity with the physical theories it (allegedly) implicates. Consequently, the section will instead begin with some preliminary details about the standard model of cosmology and the theory of GR upon which the standard model is built.

In the context of GR, one may understand a cosmological model to consist in three objects: a manifold M, a Lorentzian metric g_{ab} defined on M, and a tensor field T_{ab} on M, each of which satisfy certain technical conditions that will not be important here. As a pair, (M, g_{ab}) describes a general relativistic spacetime (c.f. Malament (2012)); meanwhile, T_{ab} encodes the stress-energy of a collection of matter fields over that spacetime. As was said above, the standard model of cosmology is built on the framework of GR, which means that the geometry of the spacetime wholly determines the distribution of matter throughout spacetime (as represented by T_{ab}) via the Einstein Field Equation (EFE). What is the EFE is a matter of minor controversy, but start by considering it in the following form (in geometrized units where $c = G = 1$):

$$R_{ab} - \frac{1}{2} R g_{ab} = 8\pi T_{ab} \tag{2.1}$$

where the left-hand side describes the spacetime geometry (i.e. the curvature of spacetime) via quantities determined by the metric g_{ab}, and the right-hand side defines the distribution of the stress-energy of matter across the spacetime. The local conservation of energy-momentum is satisfied by the purely geometrical fact that the covariant divergence of the left-hand side vanishes.

The source of controversy concerning the EFE arises in the following way. It is not difficult to show that the left-hand side of the EFE can be modified to include an additional 'cosmological' term $\mathring{\Lambda}\, g_{ab}$ without losing the conservation law over T_{ab}, as long as $\mathring{\Lambda}$ is a fixed constant:

$$R_{ab} - \frac{1}{2} R g_{ab} + \mathring{\Lambda}\, g_{ab} = 8\pi T_{ab} \tag{2.2}$$

In this version of the EFE, the cosmological term is naively interpretable as the inherent elasticity of the spacetime. Historically, when Einstein first presented his own model of cosmology, he used equation (2.2), but he did not need to. One can easily modify equation (2.2) to return to an expression like that of equation (2.1), where the cosmological term $\mathring{\Lambda}\, g_{ab}$ is absorbed into the stress-energy tensor that governs the right-hand side of the expression:

$$R_{ab} - \frac{1}{2} R g_{ab} = 8\pi \left(T_{ab} - \frac{\mathring{\Lambda}}{8\pi} g_{ab}\right) = 8\pi T_{ab}^{(total)} \tag{2.3}$$

The two versions of the EFE are thus equivalent, but in the original version we can understand the collection of matter fields represented by the stress-energy tensor $T_{ab}^{(total)}$ to also include the possibility of a vacuum term: a constant stress-energy defined everywhere on the spacetime.[1] Even when all of the ordinary matter contributions represented by T_{ab} go to zero (i.e. in vacuum regions), there is a constant quantity of stress-energy that characterizes the region. In summary then, a positive-valued 'cosmological' term $\mathring{\Lambda}\, g_{ab}$ is understood as corresponding to a negative-valued vacuum term given by the expression $-\frac{\mathring{\Lambda}}{8\pi} g_{ab}$.

Whether one understands the EFE as the expression given in equation (2.2) or as the expression given in equation (2.3) is a matter of free choice about what one considers to be the theory of GR upon which the standard model of cosmology is built. Since the 1990s, the standard model of cosmology has included a non-zero effective term like $\mathring{\Lambda}$, here denoted Λ, to recover empirical observations about the accelerating expansion of the universe.

That the standard model of cosmology requires a non-zero Λ is normally taken as evidence that there is such a thing as an 'effective' vacuum term, like that included in equation (2.3) but precisely characterized by Λ. Moreover, since the vacuum term resembles a constant multiple of the spacetime metric, it is easy to interpret the effective term as arising due to a classically available and gravitating energy density of the vacuum, with a pressure of opposite sign and zero shear (c.f. Zel'dovich (1968), Frieman et al. (2008)). This sets the stage for theoretical considerations about the phenomenology of matter fields that could, in principle, be thought to give rise to such a 'vacuum energy'.

At this point, it is crucial to note that nothing in the theory of GR and, as a consequence, nothing in the standard model of cosmology prohibits hypotheses concerning a new classical term Λ_0 which behaves like a contribution to the effective vacuum term in the form $T^{(0)}_{ab} = -\frac{\Lambda_0}{8\pi} g_{ab}$. This term is often referred to as the 'bare cosmological constant' or, in context, the 'bare term' (c.f. Carroll (2001); Martin (2012)). Note that, by the same arguments as given above concerning the 'cosmological' term, Λ_0 may be placed on either side of the EFE. In its

[1] Curiel (2016) offers an argument based on a novel uniqueness proof for the EFE that the 'cosmological' term ought only to be understood in this latter context (returning to the original version of the EFE). Although relevant to the subject matter of the present chapter, Curiel's argument should not affect the particular claims about methodology and theory development that I will make. For this reason, I leave off a study of its potential implications on desired solutions to the CCP for another time.

ordinary usage, for instance as the quantity appears in the Lagrangian/effective field theory formalism popular in contemporary high energy physics, Λ_0 is regarded as a 'zeroth order' term in the gravitational sector of GR, which is just to say for present purposes that, unlike the rest of the contributions to the effective vacuum term that are ordinarily interpreted as *due to matter*, Λ_0 is ordinarily thought to be grouped together with the various curvature terms determined by the metric. Hence, it is natural to think of a non-vanishing Λ_0 as introducing a 'cosmological' term on the left-hand side of the EFE as in equation 2.2, even under the interpretive convention that puts all other contributions to the effective vacuum term together on the right-hand side. In this way, the value of Λ_0 in the standard model of cosmology can be understood, on the standard view, to modify which geometries one properly identifies as 'vacuum regions' according to the underlying theory of GR.

Why one would entertain such a term is secondary to the main point at hand, which is that the standard model of cosmology is built in such a way that it already includes the term, which could either represent a new constant in the underlying theory (i.e. the standard view), or else it could represent a new kind of (classical) matter field in our universe that resembles the other contributions to the effective vacuum term. No matter its interpretation, its function in the model is that of a fit parameter to help match the model to empirical data (even in the case where $\Lambda_0 = 0$). The value given to Λ_0 is wholly determined by the following expression:

$$T^{(0)}_{ab} = -\frac{\Lambda_0}{8\pi}g_{ab} = T^{(\Lambda)}_{ab} - T^{(\gamma+\delta+\zeta+\cdots+\theta)}_{ab} = -\frac{\Lambda}{8\pi}g_{ab} - T^{(\gamma+\delta+\zeta+\cdots+\theta)}_{ab} \tag{2.4}$$

where $T^{(\Lambda)}_{ab}$ is the effective vacuum term and $T^{(\gamma+\delta+\zeta+\cdots+\theta)}_{ab}$ represents any other classically available contributions to that effective term.

While there are no classical sources arising in any domains of physical theory that resemble the contributions $T^{(\gamma+\delta+\zeta+\cdots+\theta)}_{ab}$, there are quantities in the standard model of particle physics that could be candidates for such contributions, if only our theory of gravity (and subsequently our model of cosmology) were built in such a way as to handle quantum phenomena. Those quantities are the expected zero-point energies that emerge from the fields described by the standard model of particle physics (i.e. in the framework of QFT on flat spacetime) when those fields are arranged in their vacuum states.[2] The presumed relationship between these zero-point energies and Λ may be understood as follows. First, interpret the zero-point energies as the energy densities of the various vacua that arise in QFT; there exist arguments that these zero-point energies have well-defined non-zero and non-infinite expectation values that are computable (in principle, and to some approximation) within the framework of particle theory. Moreover, for such quantities to be invariant with respect to the symmetries of the implicit background Minkowski spacetime of standard QFT, the corresponding (expected) stress-energy must resemble a constant multiple of the Minkowski metric (and so, the zero-point energies must come coupled with a quantity of isotropic pressure of exact proportions and with opposite sign, with zero shear). Insofar as the background Minkowski spacetime can be naively understood as an arbitrary spacetime of GR, one has found a stress-energy tensor of just the form one needs to account for Λ.

Nonetheless, it is important to emphasize that these quantities are quantum mechanical expectation values, which is to say that they are not quantities of energy in the classical sense of the term (Saunders, 2002). Any treatment of these expectation values as classical quantities applicable to a general relativistic model of cosmology depends on semiclassical approximations that model quantum mechanical quantities as classically gravitating substances. But lacking a mature theory of quantum gravity which can be shown to reduce to GR in the appropriate limiting cases, it is unclear why the predicted zero-point energy values from QFT on flat (Minkowski) spacetime can be assumed to gravitate like otherwise classical energy contributions in curved spacetime. (As will hopefully become clear in the following section,

[2] This is not the place for an extended conversation about zero-point energies. For slightly more, see section 2.3; otherwise, for more careful treatments of zero-point energies, see Rugh et al. (1999); Rugh and Zinkernagel (2002); Martin (2012); and Kragh (2012).

such an assumption amounts precisely to a belief in advance about what will turn out to be true, in light of the future theory of quantum gravity that we presently lack.) Moreover, even understanding Minkowski spacetime as a 'vacuum spacetime' (the sort of spacetime on top of which quantized matter fields ought presumably to be defined) requires that one commit by stipulation to the belief that GR is a theory in which the bare cosmological constant Λ_0 vanishes. Naturally, this stipulation undermines the extent to which zero-point energies can be expected to explain or predict any observed value of Λ.[3]

Regardless of these considerations, physicists regularly equivocate between the effective term Λ and the interpretation of any term like $\mathring{\Lambda}$ as a vacuum term belonging on the right-hand side of the EFE and due primarily to the expectation values of the zero-point energies that arise in QFT.[4] Different heuristics to compute the total value of the collection of zero-point energies suggest that the quantity is between 60 orders of magnitude (Frieman et al., 2008) and 120 orders of magnitude (Weinberg, 1989) larger than Λ. If this is understood as a rough but direct prediction of Λ, it truly is the biggest (finite) disparity between prediction and observation in the history of physics. When the collection of zero-point energies is

[3]I have assumed, in parenthetical, that the sort of spacetimes on which one defines matter fields are vacuum spacetimes, where 'vacuum spacetime' gains its meaning via consideration of the EFE in the theory of GR. In the context of GR, only were Λ_0 to vanish does Minkowski spacetime pick out a geometry that is appropriately identified as vacuum. As one reviewer rightfully points out, there is a sense in which, in the context of the so-called 'semiclassical EFE', the situation changes: spacetimes on which one may appropriately define a particular species of quantized matter field are exactly the 'semiclassical' vacuum spacetimes, where 'semiclassical' vacuum spacetimes (with respect to some species of quantized matter field) are defined as those spacetimes for which the zero-point energies of that species of quantized matter vanish (if Λ_0 is assumed to vanish) or otherwise source a particular non-zero 'cosmological' term characterized by Λ_0. Hence, in the case of the semiclassical EFE, zero-point energies feature in a new additional consistency constraint concerning the identification of vacuum states in the theory defined on Minkowski spacetime (if Λ_0 vanishes) or another spacetime of constant scalar curvature (if Λ_0 does not vanish).

I have some reservations about using the semiclassical EFE in this way (and any use of the semiclassical EFE certainly takes us beyond our well established theories), but putting those aside, the issue seems to me now to be even more stark: why should quantities derived in QFT formulated on Minkowski spacetime (for instance) say anything at all about what happens in theories formulated on other spacetimes, for which plausible vacuum states (supposing they exist) are subject to entirely different consistency constraints?

[4]This equivocation is not entirely unmotivated; the hope, at least for some, is that all sources of classical spacetime curvature according to the EFE might ultimately be explained by quantum phenomena in the framework of QFT. But I would like to caution against such optimism. It is far from obvious, for instance, why one should expect that the resources available in the present standard model of particles exhaust all possible quantum phenomena that could come to explain Λ in future theories of quantum gravity.

instead considered as the 'matter contribution' to Λ, then (by the outline sketched above) Λ_0 assumes the difference between the empirically determined Λ and its total computed 'matter contributions'.

Finally, I am in the position to concisely characterize the CCP according to its three common variations, without fear of perpetuating any odd understandings of the terms involved. The 'old' CCP suggests that there is a problem in cosmology that Λ in the standard model is effectively 0, despite the seemingly overwhelming presence of 'matter contributions' that arise as consequences of the standard model of particle physics. The 'new' CCP suggests that the problem is that Λ is not precisely 0, while also being many orders of magnitude smaller than the total computed value of those 'matter contributions'. The 'cosmic coincidence problem' suggests that the 'new' CCP is made more peculiar than it may otherwise have been by the fact that the particular value of the not-quite-zero Λ resembles the (classical) mass-density of the universe in the present cosmic epoch.

But we have already seen that none of these variations of the CCP constitute strict problems within the standard model of cosmology, because the standard model permits the inclusion of a new term Λ_0, whose assumed value provides, by construction, the difference between the effective vacuum term and whatever other classical vacuum contributions are identified. Moreover, according to current theory, the 'matter contributions' are zero-point energies that emerge in the context of QFT, which are precisely not the sorts of classical quantities that readily act as contributions to Λ.

Nonetheless, in framing the subject in these intertheoretic terms, one begins to get the sense that the CCP is foremost an issue of how it is that the consequences of other theories that are, at face value, incompatible with the theory at hand can be leveraged in the physical interpretations of particular terms within the theory at hand. If Λ is interpreted as a vacuum term (i.e. the total gravitating energy of the vacuum), and if the quantum 'matter contributions' to Λ arising in QFT on flat spacetime are thought to be exhaustive of all suitable (i.e. gravitating) vacuum physics, then considering the remaining contribution to Λ (i.e. Λ_0) as anything besides a vacuum term seems wrong-footed. On the other hand, if one simply wishes to assign a physical interpretation to the remaining contribution Λ_0 as a vacuum term, such as by postulating the presence of a constant, classical field, then one is wielding the standard model of cosmology including the physical interpretation of Λ as a criticism of the completeness of another theory entirely. In other words, while the CCP does not describe a problem within our current theories as such, it does suggest problems that may soon arise, as one begins to speculate about the contours of a future theory of physics that is meant to straddle the overlapping domains of the current ones.

Pausing, for just a moment, our discussion of the particulars of the CCP, I suspect it will be helpful to meditate on one of the general themes of this section.[5] The point has been made that the CCP is about a putative conflict between an empirically loaded quantity, i.e. Λ from the standard model of cosmology, and a theoretical entity arising in our premier theory of matter, i.e. zero-point energies. Recalling Laudan (1978), this situation has all the hallmarks of an empirical problem ("...anything about the natural world that strikes us as odd, or otherwise in need of explanation..." [p. 15]), perhaps best regarded as an unsolved empirical problem, given that it has indisputably directed future inquiry.

On the other hand, I have stressed repeatedly that there is no sense in which the theoretical entity arising in our premier theory of matter *decidedly matters* in conversations about Λ in standard model cosmology, and the freedom of the bare term Λ_0 in current theory insulates our current standard model of cosmology from any similar problem in principle. Any conflict here must therefore be understood in exclusively intertheoretic terms; this is the conclusion that we have just arrived at. For Laudan (1978), this would seem to place the CCP as a conceptual problem— an external one, in fact. It is true that Laudan regards empirical and conceptual problems as co-existing along a spectrum, or a "continuous shading" [p.

[5] I thank an anonymous reviewer, as well as an editor, for encouraging me to make these general morals more explicit.

48], of problem types intermediate between two extremes— my intention is therefore not to disparage his taxonomy. To the contrary, what I find so curious about the CCP, and this will become the primary focus in the following section, is that despite being, in some sense, a conceptual problem that straddles our current theories of matter and gravity, it nonetheless seems *most aptly described* as a first-order, empirical problem whose domain or context of inquiry just so happens to be located within some future theory that we simply do not yet have. Since we do not yet have that theory, there will be genuine expert peer disagreement about the precise nature of the problem, which will subsequently motivate radically different paths of research; this is the subject I will take up extensively in the following section. For now though, the slogan to repeat is this: the CCP is *about* future physical theory, identified in advance via the details that we today suspect will come to feature in that future theory.

2.3 Making sense of the problem(s)

It might be tempting to dismiss the CCP outright, and declare the physics community's reactions to it philosophically lazy, misguided, or even agenda-driven. This is the view implied by Earman (2001, p. 207), who writes: "Steven Weinberg (1989), who believes that physics thrives on crises, has been instrumental in promoting this problem to the status of a crisis for contemporary physics. I want to explain why this 'crisis' needs to be viewed with some skepticism."

I do not dispute Earman's claims in these remarks, but I confess that Earman's general attitude here strikes me as backward. The philosophical point to be made is not that the physics community might be in error about a concern that is central to their field (i.e. that "this 'crisis' needs to be viewed with some skepticism"). Rather, the point to be made is that the physics community's persistent worry about the CCP *as a problem* (alternatively, as a 'crisis') suggests that there might be something else going on. There is something interesting, that is, about the fact that the depiction of the CCP as a problem for contemporary cosmology has not diminished over the past several decades in light of all of the potential 'solutions' that have been offered for it.

There is an explanation for this state of affairs, which was hinted at in the final remarks of the previous section. If the zero-point energies that arise in QFT are understood as predictions today about what vacuum energy sources will be present in some future physical theory which unifies GR and QFT (i.e. the theory of quantum gravity), and if, moreover, vacuum energies in that future theory contribute to some effective term that reduces to a classical vacuum term, then that future theory (but *not* any present theory) is thought to have to reckon with the apparent disparity between the quantum contributions in current QFT and Λ in the standard model of cosmology. Subsequently, solving the CCP consists in endorsing a possible framework for future physical theory and subsequently demonstrating that within such a framework, the CCP is an ordinary empirical problem with an ordinary solution already built into the future theory. Insofar as any such solution depends on conjectures about what future theory will entail, the CCP has remained (and will continue to remain) an open problem until some of those conjectures gain sufficient empirical grounding to warrant adoption into the corpus of established physical theory. But this entire conversation presupposes some resolution to a fascinating methodological conundrum: how do we talk about what future, as-of-yet undiscovered theories *ought to entail*, and on what grounds do we make such claims?

In the case of quantum gravity, in which the goal is to find a theory that formally interprets both the gravitational properties of energy sources and the physics associated with quantum states, the standard move is to introduce a semiclassical theory of gravity that is meant, in principle, to approximate such an unknown future theory. Unfortunately, since the eventual theory of quantum gravity is unknown, it is unclear in what respects and what domains of applicability this procedure is meant to approximate that eventual theory, and a well-evidenced classical model of cosmology (e.g. the standard model of cosmology) does not necessarily transform into a well-evidenced semiclassical model when the underlying theoretical frameworks are swapped. Nonetheless, this is the sort of architecture available to theorists interested in considering possible eventual consequences of quantum gravity in cosmological contexts.

The point here is that claims made on the basis of semiclassical variants of classical gravitational models are not consequences of current physical theory. Rather, they are claims about what we might infer from contemporary physics about approximations of future physics. This answers the first part of the conundrum voiced above (how do we talk about what the future theories ought to entail: by semiclassical alternatives to the current theories), but it does not answer the second part of the conundrum (what are the grounds for these claims about such future theories). Notice that justifications for the methods used to predict consequences of future theories can only ever consist in claims which derive from the current theories, which (by setup) are silent precisely in those contexts where the semiclassical theory is needed. Thus, the assumptions that go into these methods are open to expert disagreement: there are reasonable disputes about what semiclassical gravity entails for the future of cosmology, which means that there can be no obvious answer to the second part of the conundrum. This conclusion is not surprising. Nonetheless, I believe it is helpful to think about theory development explicitly in these terms, i.e. as cases of genuine disagreement about the details of the future theory sought, stemming from disagreements about what it is appropriate to broadly infer about that future theory from our current best theories that we have developed so far.

In fact, competing assumptions along these lines give rise to three broad categories of solutions to the CCP:

1 Assume that vacuum quantities (i.e. quantities that resemble constant multiples of the metric) do not gravitate as ordinary sources in the EFE. Then the CCP becomes:

what gives rise to the effective vacuum term characterized by Λ in the standard model of cosmology?

2 Assume that zero-point energies gravitate as tensorial quantities on the right-hand side of the EFE and exhaustively source the effective vacuum term characterized by Λ. Then the CCP becomes: how does one account for the discrepancy between the currently computed values of the zero-point energies and the observed Λ?

3 Assume that Λ is not exhaustively sourced by vacuum energies. Then the CCP becomes: what other physical mechanisms can contribute to that which is presently understood as the effective vacuum term characterized by Λ?

Note that both the second and third categories expect (or at least permit) zero-point energies to contribute to the effective vacuum term that arises in the standard model of cosmology. (Where they disagree is about what the status of the bare term Λ_0 will be after all vacuum sources have been considered.) In this way, they are both incompatible with category **1** solutions, which assume that the vacuum term represented by Λ in the EFE simply cannot consist in vacuum stress-energies. To get a sense of what this means, consider some examples of category **1** solutions: unimodular gravity (c.f. Earman (2003) and Ellis et al. (2011)) and gravitational aether (c.f. Afshordi (2008)). Both approaches consider trace-free restrictions of the right-hand side of the EFE, and hence vacuum terms are decoupled from spacetime curvature.[6]

[6]The inference here in the case of unimodular gravity has been criticized in an unpublished manuscript by Earman (see, in particular, §4.1 therein). It is perhaps more appropriate to say: assuming first that any vacuum terms are taken to be fixed externally, once and for all, across all physical possibilities intended to be circumscribed by the models of the resulting theory, unimodular gravity decouples those terms from spacetime curvature. Whether or not the antecedent assumption is viable as an attitude in theorizing about cosmology is beyond our purposes here, but will surely inform the sense in which unimodular gravity constitutes a satisfying category **1** solution to the CCP. Per arguments in the same paper, it may also be the case that unimodular gravity, understood as a different theory whose interpretation precludes the assumption just articulated, winds up as a solution to the CCP belonging in a different category— such difficulties seem to have less to do with how the categories are mutually contrasted and more to do with the nuances inherent in delineating theories in terms other than their interpretations (and particularly when those interpreted theories have not been extensively studied).

By contrast, both categories **2** and **3** assume that vacuum sources, where they can be assigned suitable stress-energies, gravitate in the usual way. But whereas category **2** accepts the semiclassical assumption that zero-point energies in QFT will come to (fully) source Λ, category **3** disputes it. More formally, category **2** assumes the most common semiclassical variant of the standard model of cosmology, in which the term that reduces to Λ is considered to be explained by those terms that reduce to the zero-point energies in QFT. Given this as a starting point, approaches in category **2** focus on undermining the plausibility of the particular values handed over from QFT as contributions to Λ. Approaches in category **3**, meanwhile, begin with the assumption that there are other mechanisms that contribute (perhaps exhaustively, perhaps not) to Λ that cannot be understood as stress-energies of vacuum sources. Subsequently, work done in category **3** is spent seeking out alternative physical explanations for Λ's particular non-zero value. Hopefully it is clear how approaches in these two categories, insofar as they are taken to be solutions to the CCP, are thus incommensurable from the start (even though work in one category can absolutely inform work in the other category).

Theorists working in category **2** have two avenues available to them that would solve their variant of the CCP. Both avenues consist in challenging the particular computations of zero-point energies in contemporary particle physics as providing an accurate heuristic for the values of the same in the future theory. The first avenue usually involves disputing the assumption that calculations of vacuum quantities in QFT formulated on flat spacetime resemble plausible predictions of vacuum quantities in QFT formulated on a more generalized family of curved spacetimes. Instead, heuristic accounts of QFT on curved spacetime are explored to determine whether the computed zero-point energies are sufficiently suppressed so that the CCP does not appear. In other words, this avenue is motivated by the possibility that the generalization of the standard model of particle physics to particle physics in the presence of spacetime curvature, on route toward a theory of quantum gravity, will happen to undercut the CCP.[7]

The other avenue consists in suggesting potential modifications to standard particle physics on flat spacetime, even before considering how the model could be generalized to curved spacetimes. Work along this avenue explores what sorts of modifications would, in effect, cancel out the currently computed values. The idea here is that the standard model of particle physics is incomplete, even in the low-energy regime of approximately flat spacetime. Improving upon the model, even still in the framework of QFT on flat spacetime, might happen to resolve the discrepancy between Λ and the total contributions from the zero-point energies. For instance, it was first noted by Zumino (1975) that certain supersymmetric particle theories could suppress zero-point energies in precisely the right way to remove the worry of the CCP. Generalizations of this idea have become quite popular in the context of supersymmetric string theories (see, e.g., Kachru et al. (2003)).

There are many theoretical difficulties with this broad approach, especially if Λ is non-zero (Witten, 2001), but its pursuit as a possible solution to the CCP is noteworthy. The standard model of particle physics is incredibly well supported by empirical data, so the idea is not that the standard model *ought to be* modified so as to address the CCP. Rather, the idea is that there is value in determining which sorts of modifications to the standard model of particle physics happen to solve the CCP. In the case of such modifications, if in the future

[7] For example, when Wald (1994) introduces his candidate for a semiclassical stress-energy tensor that is defined with respect to the fundamental observables of a quantum field theory understood in a setting appropriate for curved spacetimes, the uniqueness of that candidate depends on putting in by hand its value when the relevant quantum field is defined on Minkowski spacetime and arranged in its vacuum state. In other words, it is essential in his approach that one stipulate what is the value of the zero-point energies that are thought to arise in standard model particle physics. Depending on one's arguments about how to generalize QFT to curved spacetimes, one may find themselves stipulating different values, including those that would mimic the effects of the effective vacuum term in the standard model of cosmology. On the other hand, as one reviewer points out, the inflexible nature of this framework as a means by which to resolve the CCP can easily lead to over-corrections, and therefore to a new instantiation of the CCP, as which seems to occur in an article by Hollands and Wald (2004).

it were to turn out that such modifications to the standard model were justified, then the community would no longer be worried about the CCP.

Along both of the theoretical avenues, notice how the CCP, not quite a problem within its present context, is transformed into an ordinary, defeasible problem in the contexts of new theoretical initiatives (e.g. QFT on curved spacetime, supersymmetry).[8] I can be more explicit. A solution to the CCP by appeal to some new theoretical initiative X can be understood as consisting in a demonstration that the CCP would be a solved problem in the next generation of physical theory, were the next generation of physical theory to include X. In category **2** (pursued by those who believe that zero-point energies ought to wholly account for the presence of Λ in the standard model of cosmology), X is any generalization of the standard model of particle physics, which is presently formulated in QFT on flat spacetime, to a new theoretical framework compatible with our present understanding of spacetime qua gravity. In such a generalization, the CCP is transformed into a particular ordinary problem within that framework: what is the total magnitude of the expected quantities of zero-point energies that arise in that framework?

This suggests that one role of the CCP in its present form is to provide a heuristic by which new theoretical initiatives are judged, or rather, explicit motivation for the pursuit and development of some such initiatives over others. Once suitably transformed, the defeasibility of the CCP (as it is rendered in the terms available to some new theoretical initiative) is viewed as a strength of that theoretical initiative: because were it the case that future physicists turn out to need the new theoretical initiative, a (future) problem has both already

[8]It is worth flagging here what I intend throughout this chapter by the term 'theory', as opposed to 'theoretical initiative' or 'framework'. It would take us too far from the central point here to discuss what ultimately ought to delineate these various constructs, properly construed. As such, except where the word appears as parceled together in a standard phrase from physics (e.g. quantum field *theory*, string *theory*, etc.), I have tried to reserve 'theory' for use just in regards to our current, well-supported theories of GR and QFT (formulated on Minkowski spacetime), as well as in regards to the future theory of quantum gravity we ultimately seek. Theoretical initiatives and frameworks, by contrast, are more preliminary in nature, as the sort of things upon which or with the help of one might hope to eventually develop successful cosmological models.

been formulated and been subsequently solved. This could partially explain the intense spotlight that has been shined on the CCP over the past several decades: for as long as theoretical physicists have been seriously pursuing a new theory of quantum gravity (and the standard model of cosmology was sufficiently mature so as to warrant talk of Λ), the CCP has provided a heuristic by which to evaluate new proposals, orienting focus in an otherwise wild theoretical field. Amongst those who think that the new theory ought eventually to respect zero-point energies as the exhaustive sources of the vacuum term that arises in present cosmology, solutions to the CCP in category **2** highlight which theoretical initiatives along the way are most attractive.

Contrast the transformation of the CCP relevant to category **2** with that which occurs in the case of category **3**. Recall that category **3** is characterized by its rejection of the common semiclassical assumption that zero-point energies fully source the vacuum term characterized by Λ in standard cosmology. For this category, X is any semiclassical theory of gravity in which the geodesic structure of a spacetime is no longer thought to uniquely determine the distribution of stress-energy across it. Beginning with the assumption that this is to be the next new theoretical initiative, the CCP is transformed into a different ordinary problem with a different ordinary solution: what new physics can at least partially account for Λ, which is presently understood as wholly characterizing a gravitating vacuum term? Approaches in this category generally consider the addition of new fields on spacetime that do not couple to spacetime geometry as traditional energy sources, or else more fundamental modifications to how we think of the relationship between spacetime geometry and matter.

The first of these two avenues takes seriously the popular interpretation that Λ is, at least partially, due to a 'dark energy' that couples in atypical ways with the other matter fields involved in standard particle physics. In other words, one expects there to be an effective classical field theory that can characterize the accelerating expansion of the universe. If it turns out to be a constant scalar field, then it is empirically indistinguishable from the interpretation of the effective vacuum term as due to vacuum energy contributions in the classical regime, but should be empirically distinguishable from predictions of the standard model of particle physics in a high-energy regime described by quantum field theories. If it turns out to be any other sort of field (e.g. the model of quintessence given by Zlatev et al. (1999), or else that of an exotic fluid as given by Kamenshchik et al. (2001)), then it is empirically distinguishable from the current theories in both regimes. Either situation carries implications for an eventual theory of quantum gravity.

The second avenue explores what it would take to capture the accelerating expansion of the universe (whether constant or ultimately dynamic) as a cosmic-scale consequence of certain local geometrical properties of spacetime, independent of the matter content of the universe. In these approaches, the quantity Λ that arises in the current standard model of cosmology is understood as an effective term that captures the magnitude by which GR is inadequate as a framework on which to build cosmic-scale spacetime models. $f(R)$ theories of gravity are phenomenological examples of this approach, but there are others examples as well. For instance, MacDowell-Mansouri gravity (Wise, 2010) and the related projects in 'doubly special relativity' or 'de Sitter relativity' consider spacetime theories like GR, except where the most relevant tangent space at each point is de Sitter-like, instead of Minkowskian (Kowalski-Glikman and Nowak, 2003; Aldrovandi and Pereira, 2009; Almeida et al., 2012).

Another example along this second avenue (but of a very different character to those already mentioned) is found in cascading gravity models, which realize *degravitation*, the theoretical notion that the presence of additional large spacetime dimensions can effectively degravitate vacuum energies (De Rham et al., 2008). These approaches take seriously the question "why does the vacuum energy gravitate so little?" (Dvali et al., 2007, p. 1) and try to answer it by demonstrating feasible mechanisms (i.e. different embeddings of four-dimensional spacetime in a higher-dimensional space) by which Λ can be decoupled from any computed values of vacuum energies. The goal of these projects is to dilute the relationship between the observed value of Λ and any computed stress-energies of vacua in precisely the right way to reconcile observations and/or predictions of the two in conjunction.

In all of these cases, note that the modifications to GR obviously influence considerations about semiclassical gravity by providing a new classical theory that semiclassical gravity must reduce to. This in turn guides which strategies toward a theory of quantum gravity are considered most viable. In this way, considerations of the many different sorts of solutions to the CCP that appear in category **3** reveal a richer understanding of the role that the CCP plays in theoretical research: more than providing a guiding hand as to which theoretical initiatives can do more than others in the wild field of quantum gravity research, the CCP suggests possible physical constraints on future theories of quantum gravity that are both motivated by our most well-evidenced theories and that would otherwise be entirely unforeseen. Just like in the case of the category **2** solutions to the CCP, both avenues of research in category **3** pave the way for the eventual theory of quantum gravity to have a built-in solution to a technical problem that formally reduces to the CCP. But now I can say a bit more: category **2** solutions focus on pushing the boundaries of what we know about the quantum nature of matter based on appeal to cosmological evidence, while category **3** solutions focus on pushing the boundaries of what we know about the relationship between matter and spacetime geometry (i.e. geodesic structure) based on appeal to cosmological evidence. Naturally, each of these is relevant in pursuit of quantum gravity.[9]

[9]There is a fourth category of solutions to the CCP in the context of future theory that focuses on explaining away the CCP by heuristic arguments concerning probability measures that emerge in the context of several speculative theoretical initiatives, which render the observed value of Λ antecedently probable. Since the backbone of the fourth category, anthropic reasoning, is a source of its own philosophical battles, the details of solutions in this category have not been included in the present work. For more on the subject, see Smeenk (2013) and Curiel (2015) to get a sense of the arguments that plague the general approach.

2.4 Discussion

The goal of this chapter was to discuss the odd state of affairs surrounding the CCP, in the hopes of drawing philosophical lessons about scientific methodology on the cutting-edge of theory development. To that end, some care was given to how the CCP was introduced: as a quirk of our standard model of cosmology but not technically a problem within it, which only starts to look more like a problem when one turns one's attention to what future physical theories might entail. How the CCP transforms into an ordinary problem in light of what future theories are thought to entail became the focus of section 2.3, whereupon it was noticed that the CCP transforms into different problems for future physical theory depending on the assumptions one is prepared to make about what the future theory will look like.

This unifying feature of all three categories– that the CCP of today will eventually take the form of an ordinary problem with an explicit solution already provided– suggests a general understanding of why it is that the CCP of today is treated by the theoretical physics community as a problem at the forefront of their field. Simply put, it is the reduction to the language of our present theories of (one of) the first empirical problem(s) that the next physical theory (but *not* any current theories) will be expected to solve, above and beyond all of the other problems that it must also already solve (namely, all those that reduce to already-solved problems in our current best theories). Moreover, it provides a way by which the consequences of our best current theories can be leveraged for a glimpse at what may be necessary in the development of the next generation of physical theories. That the solutions to the CCP— those glimpses at what may be necessary— can sometimes seem entirely incommensurable as solutions to the same underlying problem turns out to be an artifact of the quite reasonable observation that there exists genuine expert disagreement about what we know today concerning the future of physics.

In regards to the first of these points, Weatherall (2011) identifies a similar, although retrospective, story concerning the equivalence (according to Newtonian physics) of inertial and gravitational mass, where it was viewed as a problem that the two Newtonian quantities were empirically indistinguishable (even though it technically was not a problem for the theory). The solution (that is, an explanation of their observed equivalence) came in the context of GR, even though there is no equivalent problem in the language of GR to be solved (because there are no such things as inertial or gravitational mass). Nonetheless, in the reduction of GR back to Newtonian physics, one may derive the equivalence of the two Newtonian quantities, which seems to count as a satisfying explanation in response to the original Newtonian problem. In this way, the development of GR is retrospectively characterized as having made progress in regards to the purported problem, even though there was no actual problem in the framework of Newtonian physics and the theory of GR is such that there could never be any such problem, coherently stated.

Weatherall (2011) uses this case to come to the conclusion that at least some problems in physics have the effect of shaping the next generation of research in the field, because their solutions involve departing from present theory and engaging in new theoretical developments. Neglecting the details of his particular argument, I am inclined to make a related but slightly stronger claim on the basis of the CCP: the primary function of at least some problems in physics is to shape the next generation of research in the field, because for researchers to merely entertain them *as problems to be solved* requires new theoretical initiatives with which to properly articulate them. The distinction between Weatherall's claim and my own is subtle but important: my claim is not that problems like the CCP are wanting for solutions in the next generation of theory, whereupon today's theory will be viewed as inadequate in regards to solving such a problem. To the contrary, recall the slogan from the end of section 2.2: the CCP is *about* future physical theory. That is to say, the broad strokes which characterize research in frontier physics are laid out in accordance with the different ways in which intertheoretic worries like the CCP can be converted into specific articulable problems to be solved, whereupon by the time the details of the next generation of theory are worked out, those brand new specific problems will already have been taken care of.

I contend that the intense focus on the CCP in contemporary theoretical physics supports this claim. As a consequence of the community worrying about what can only really be described (at worst) in present theory as the setting of a fit parameter according to empirical data, multiple independent lines of advanced theoretical inquiry have been developed, each of which features a sketch of future physical theory in which a newly articulated problem has been solved. Moreover, those sketches can be easily grouped, broadly, in terms of the additional, often pairwise incompatible assumptions that would warrant them as components of the next generation of physical theory.

It would be a mistake to miss what is at stake in this claim. That we have access to what are essentially first-order empirical problems within a future theory, without having prior access to the details of that future theory, is extraordinarily counter-intuitive. That potential solutions may be worked out for those problems in advance, prior to writing down that future theory, is perhaps even more so. But I hope that the case study of the CCP demonstrates how these affairs can nonetheless be so, albeit at the cost of there being genuine peer disagreement about what we know today on the basis of our current best theories of physics.

If the target of my hope here has been achieved, I believe that the present analysis runs counter to an otherwise dominating metaphor in the methodology of theoretical research. The dominating metaphor to which I refer is that which considers theory development in terms of principled explorations of something like a 'theory space'. As an instance of this metaphor in action, consider the recent essay by Rovelli (2018), in which he laments, in the context of frontier physics and quantum gravity research, what he identifies as "the current "why not?" ideology: any new idea deserves to be studied, just because it has not yet been falsified; any idea is equally probable, because a step further ahead on the knowledge trail there may be a Kuhnian discontinuity that was not predictable on the basis of past knowledge..." [p. 487]. The influence of the 'theory space' metaphor is, I think, quite obvious: there exists some total collection of ideas from which the next theory is to be selected, according to some or other principle of selection. Rovelli's charge against the frontier physics research community is that the relevant principle of selection employed is something like random sampling with respect to a principle of indifference. Against this principle of selection, Rovelli goes on to write that "When we consider ourselves to be "speculating widely", we are mostly playing out rearrangements of old tunes: true novelty that works is not something we can just find by guesswork". In other words, the criticism of the "why not?" ideology is (at least in part) one of internal inconsistency: that the sampling mechanism utilized in its name does not conform to the specifications that would otherwise justify it. But Rovelli also points out that the "why not?" ideology suffers, moreover, from that total collection of ideas being too large to meaningfully sample.

Rovelli's conclusion is that we ought instead seek a method that has science proceeding "through continuity, not discontinuity". My implicit suggestion in this chapter has been that there are such other methods that may already be found in practice, and that the role of the CCP in quantum gravity research witnesses one of them. Namely, I have suggested that the persistent identification of the CCP as a problem in need of solution exhibits a method of theory development in the frontiers of physics, whereby future theory is built in the first place as the sort of thing which already solves what may today be recognized as empirical problems arising within it. Note the departure from the 'theory space' metaphor: my claim is not that theory development be understood, in this case, in terms of the selection of a new idea from the space that happens to solve the CCP; to the contrary, the new idea is built in the first place by supposing future theory to already be constrained in certain ways (by the successes of current theory), and subsequently solving the problems that would arise within the future theory given those constraints.

Taking a step back, it is clear that by virtue of being entertained as empirical problems in the context of some or other future theory, intertheoretic worries like that of the CCP in our current best theories provide a way to leverage those current theories as justifications for particular new theoretical initiatives in the pursuit of future physics. That is to say, the CCP is best understood as a means by which the frontier physics research community may illuminate the possible paths forward from current theories of spacetime and matter to future, more sophisticated ones.

Chapter 3

Betting on future physics

3.1 Introduction

From the vantage point provided by our current best physical theories, how do we anticipate the details of future theories that we do not yet have? Having not yet developed those future theories, the situation is apparently precarious. Common sense skepticism seems to council against even asking such a question: the world has, time and again, surprised us in dramatic ways and defied our expectations of it; the future should likewise come to surprise us still.

But beliefs regarding the contents of future theories are of a different kind than are beliefs about the world, because we are the ones, generally speaking, who are responsible for bringing those future theories about. It should not be surprising, then, that our beliefs today about some of the details of those future theories may fair better in regards to (future) states of affairs than our beliefs today about the world. As such, we plausibly *can* anticipate details of future theories in advance, in which case it becomes prudent to ask how we might indeed go about doing so. And an answer to this question should be of great interest to philosophers of physics and physicists alike— to the former because the answer explicitly concerns the future stability of our present knowledge about the physical world, and to the latter because it is, in a strict sense, the professional intention of that community to carve out the path that ultimately gets us to those future theories, from where we are at present.

Nonetheless, my goal in what follows is not to provide an answer to this question in the most general case; it is, comparatively, much more modest. I will focus (very narrowly) on the circumstances surrounding what I take to be the historically standard version of the "Cosmological Constant Problem" (CCP), a "crisis" in contemporary theoretical physics that has motivated a large amount of speculative work in the past half century of quantum gravity research. In this setting, I will consider what it takes to justify the belief that such a problem is *worth taking seriously* (in the development of the future theory of quantum gravity). While the emphasis here is on motivating paths of research in theory development, I will argue that this issue is essentially about what justifications are available for anticipating that certain details will feature within that future theory which is eventually to be developed.[1] The details in this case will concern a resolution to the historically standard version of the CCP, which one must first bet is well posed within the context of that future theory. It is in this sense that the present study becomes a special case of the much more general project.

Meanwhile, in the context of the more general project, the virtue of focusing narrowly on the case of the CCP is also clear. As already mentioned, the situation in the general case is precarious. But in the special case of the CCP, that the details of the future theory under scrutiny concern a resolution to an explicit problem gives us traction in developing

[1]The suggestion here, more generally, is that there is sometimes a connection between the pursuitworthiness of paths of research toward a future theory and what we anticipate about the details that will turn out to feature within that future theory. In these cases, arguments in regards to the former can be found in the justifications that are available for beliefs pertaining to the latter. Of course, as was alluded to already, it may turn out that we are ultimately wrong about future theory, perhaps in surprising ways. Since the justifications I will discuss below are only ever anchored in the features of our current theories that we *today* believe are responsible for their empirical successes, it is by no means my intention that everyone within the physics community pool their efforts along a single "most pursuitworthy" path (and I also make no claims as to there being anything like a total order on such paths, defined by the extent to which they are each pursuitworthy). The point, rather, is to better understand how disagreements amongst theoretical physicists about how to develop future physics bottom out in principled disagreements about what we may presently anticipate about the end product of that development.

one answer to the general question, at least in the context of quantum gravity research. In particular, I will argue that a belief along the lines above depends, one way or another, on commitments to particular interpretations in our current theories of gravity and matter, according to which there are *reasons to bet* that the problem will indeed arise. To then say that such a belief is grounded in the successes of our current best theories so far will depend on the extent to which we are willing to view our commitments to those interpretations as themselves grounded in the same.

3.2 The 'Cosmological Constant Problem'

In the previous chapter, I argued that the CCP is best understood as a collection of mutually exclusive problems, indexed by the mutually exclusive assumptions about future physics that give each of them their force. I do not wish here to take a side for or against the claim argued there. What I will do here is adopt, largely uncritically, the thought that when one speaks of the particulars of the CCP, what one really means is to speak of the particulars of one (or another) *version of the CCP*. Moreover, in doing so, one inherits the baggage of whatever assumptions about future physics are implicated in the precise statement of that version.

In particular, I will focus here on (what I take to be) the *historically standard version* of the CCP, or the version of the CCP that corresponds to historically standard assumptions about future physics, made by appeal to our current theories of general relativity (GR) and quantum field theory (QFT).[2] The statement of that historically standard CCP is as follows:

> How does one account for the (enormous) discrepancy between the observationally constrained value of the cosmological constant that arises in the standard

[2]For a physicist's pedagogical introduction to this version of the problem, see Carroll (2001). Weinberg's (1989) paper first declared it a crisis, but one finds a much longer history in the article (and thorough philosopher's primer) on the subject by Rugh and Zinkernagel (2002).

model of cosmology and the currently computed values of the zero-point energies that arise in standard model particle physics?[3]

As a slogan, the version of the CCP stated here is apparently quite compelling, lending itself as decades-long motivation for an impressively vast literature of speculative physical mechanisms that, pending their inclusion in the next generation of physical theory, would either resolve the problem or sidestep it altogether. Philosophers of physics have not ignored this. Earman (2001), for instance, has emphasized that, despite all of the attention the CCP is given in the theoretical physics community, we should be skeptical of its status as a problem. Much more recently, Koberinski (2021) has argued that, given such a skepticism, we should be hesitant to continue working on projects done in its name.

Within the physics community, meanwhile, in light of all of this activity, a criticism of a slightly different kind has been echoed by a vocal minority. The criticism is this: what we know already about spacetime, or perhaps about fields on spacetime, or perhaps about defining quantum fields, renders the problem stated above as *simply not worth taking seriously*. On this view, the future theory of quantum gravity is just not the sort of thing that one anticipates will feature, in its details, a resolution to this problem. (And if one even thought that it might, then that is already to have conceded that the problem is well posed.) Meanwhile, believing otherwise requires a dogmatic refusal to accept some or other lessons we have already learned about the physical world, on the basis of our current best theories so far.[4] This is the sort of criticism that I think it will be fruitful to focus on here. In particular, one might rightly ask: by appeal to what is one permitted to argue that physicists ought not to take a problem seriously on the basis of our current best theories so far? Conversely, by

[3]Prior to the late 1990s, the standard model of cosmology was consistent with a vanishing cosmological constant, and so earlier commentaries on the CCP frame the problem as a mystery about the vanishing observational signature of the otherwise enormous zero-point energies (e.g. Weinberg (1989)). Nonetheless, the key insight is that putative cosmological observations of the one quantity are, on the historically standard view, taken to be observations in some sense of the sum total of the values of the other quantities.

[4]By way of examples, I take Carlo Rovelli, George Ellis, William Unruh, and Robert Wald to be four influential physicists who would all agree with this criticism, but each for different reasons and to very different ends. I will return to each of these four figures below.

appeal to what is one permitted to argue otherwise, such that we might anticipate of future theory that it will wind up built as the sort of thing that includes a resolution to that very same problem?

In what follows, I will develop an argument in defense of taking the historically standard version of the CCP seriously, which is to say an argument in defense of the anticipation that the future theory of quantum gravity will come to feature, in its details, a resolution to it. As has already been mentioned, the argument will turn on particular assumptions about standard interpretational commitments that physicists have in our current theories of gravity and matter.

To forestall the possibility of confusion, I want to stress that in developing such an argument, I do not mean to endorse its conclusion. Indeed, I am sympathetic to the skepticism broadly shared by those whom I have already mentioned. Nonetheless, I think it is productive to trace back the locus of that skepticism to particular disagreements with the standard view about what we may claim to know on the basis of the successes of our current best physics. Moreover, for those who are not swayed by such skepticism, the argument I will develop stands as proof of concept: the belief that quantum gravity will come to feature a resolution to the historically standard version of the CCP can be justified on principled grounds, via commitments to particular interpretations in our current, empirically well-supported theories. (And consequently, per the remarks in footnote 1, those paths of speculative research that aim to resolve the problem are themselves worthy of pursuit.)

3.2.1 The physics that anticipates 'vacuum energy'

There are, upon inspection, two distinct statements contained within the slogan given above. The first is that there is a discrepancy between the value of the cosmological constant, on one hand, and the values of the zero-point energies, on the other. The second is that each of these quantities pick out the same physics, or the same features of the world. Put aside the first of these statements to focus on the second: according to what arguments are we supposed to understand the two quantities as picking out the same features of the world? Call this question "Q".

Q is not so straightforward of a question, because the two quantities implicated in it gain their significance in radically different theoretical frameworks (this is, after all, why they are referred to by different names). The cosmological constant arises in the context of GR, our premier framework for understanding the large-scale dynamics of spacetime qua gravity, whereas zero-point energies arise in the context of QFT (on Minkowski spacetime), our premier framework for understanding the small-scale phenomenology of matter. Reconciling the two competing pictures of the world gotten from GR and QFT is precisely what an adequate theory of quantum gravity is meant to do (Huggett and Callender, 2001). Depending on whether we are supposed to understand the two quantities as picking out the same physics or not, different adequacy criteria for a theory of quantum gravity will present themselves.

One answer to this question, perhaps the favorite pedagogical answer (see, e.g., Rugh et al. (1999) or Carroll (2001)), is that the two quantities each describe, in their respective theoretical frameworks, an energy level associated with the vacuum, or perhaps 'empty space', such as can be expected to dominate in the vast regions of void found within our observable universe. On this view, there is some pre-theoretic feature of the world, 'vacuum energy', to which the theoretical terms "cosmological constant" in GR and "zero-point energies" in QFT each correspond.

I would like to caution against taking this view as providing an answer to Q, and rather suggest that it amounts to a restatement of the question. The recent history of philosophy of science has taught us to be distrustful of accounts of physics that depend essentially on there being some language of the world, to which our theoretical terms straightforwardly correspond. That we can postulate a pre-theoretic feature of the world called "vacuum energy" and insist that the theoretical quantities of the cosmological constant in GR and zero-point energies in QFT each correspond to it does not explain why those quantities in each of the theories can, in any case, be expected to pick out the same feature of the world. Particularly in the present case of interest, both 'vacuum' and 'energy' are incredibly theory-laden concepts whose characteristics differ dramatically depending on which theoretical environment, GR or QFT, one considers. It seems like the choice of label, i.e. "vacuum energy", is unjustly performing much of the purported explanatory work.

I have said why the favorite pedagogical answer is not, in fact, an answer to Q (or at least not obviously so), but I have not yet said how it is a restatement of the question. This is more subtle. First, recall the more general question at the beginning of this chapter: how do we anticipate the details of any future theory, given only the resources provided by our current theories? Suppose now that "vacuum energy", rather than directly referring to a feature of the world, refers instead to a quantity in the future theory of quantum gravity that is meant to bring together the domains of GR and QFT under one framework.[5] Such a future theory ought to explain how it is that the two current theories are separately so successful in their appropriate domains. And fortunately, lacking the details of that future theory— for instance concerning the quantity already picked out as "vacuum energy"— does not preclude us from making various bets about those details (to the contrary, it makes the activity of betting more interesting). Here is one such bet: from that which is appropriately identified as "vacuum energy" within the future theory, we ought to be able to recover, in the appropriate limits, both the cosmological constant and zero-point energies. This is, I

[5]I take this to be more a general assumption than that which was assumed in the favorite pedagogical answer. After all, if one believes (as in that case) that the term refers to a feature of the world, then that feature of the world is presumably, in virtue of the properties we are already want to ascribe to it, something that the future theory is expected to describe. But one may also believe the term to refer to some or other mathematical construction relevant to the future theory (which need not possess definite physical content), which stands in certain relationships to the mathematical constructions that populate our current theories of physics. In this case, the explicit identification of such a construction, which indeed stands in all of those relationships, is just part of the job in writing down the future theory.

claim, the historically popular bet about future physics, in the sense that the slogan given above constitutes the historically popular version of the CCP.

Naturally, given this bet, any (enormous) discrepancies between the values of the two quantities in the current theories presents a problem for which the future theory needs, in its details, some or other resolution: how is it that in the appropriate limiting cases, the same quantity in the future theory corresponds to such dramatically different-valued quantities as are familiar in our current theories? Put slightly differently: how can one or another of the theories that represent limiting cases of the future theory of quantum gravity be so wrong in their descriptions of the quantity of 'vacuum energy' thought to arise in that future theory, in particular when each of those limiting theories is so successfully applied toward other ends? Lacking a resolution to this problem renders empirically inadequate any theory that otherwise satisfies the terms of the above bet. In this way, the act of committing to that bet today gives rise to the expectation that the future theory will, in its details, resolve this problem.

Consequently, justifications for committing to the bet today are also justifications for such subsequent expectations— i.e. beliefs— today along these lines. In this way, we may anticipate that the future theory will detail a resolution to this problem, so long as the bet itself is (presently) reasonable. Of course, I have said nothing about why such a bet is reasonable in the first place; this is, at its heart, Q restated: how does one defend the bet that the future theory of quantum gravity will include such a quantity as just described? Arguments in favor of this bet are thus tantamount to arguments in favor of certain anticipations about details of the future theory.

Here is the form of one such argument. If one can show that the cosmological constant, understood always in the framework of GR, entails a physical quantity in GR that wants for a description in QFT as zero-point energies, and if one can show that zero-point energies, understood always in the framework of QFT, entail a physical quantity in QFT that wants for a description in GR as the cosmological constant, then one has reason to bet that some single quantity in the future theory will reduce in appropriate contexts to each of the two. The sense in which arguments of this form carry force is hopefully somewhat transparent: exactly all we have to work with in the development of future theory are our empirical observations, as they are understood in the contexts of our best physical theories to date. If, in each of our theories of GR and QFT, specific quantities can be shown to be left wanting for a description in the latter as one another, then our best way of understanding the empirical world today is precisely in such a way as there appears to be a feature of the world that admits each of the quantities from GR and QFT respectively as descriptions in suitable limits. This feature of the world is then exactly the sort of thing that the future theory of physics ought to describe.[6]

In the next two sections, I will argue one sense in which the cosmological constant, understood always in the framework of GR, entails a physical quantity in GR that wants for a description in QFT as zero-point energies, and vice versa in the framework of QFT. In doing so, I can make explicit which assumptions about the physics within each of the two current frameworks together motivate the historically standard version of the CCP, or equivalently which assumptions about our current physics warrant the bet just proposed. Such assumptions therefore also warrant, on the basis of alleged facts about our current best theories so far, the anticipation that the future theory of quantum gravity will, in its details, resolve the problem that follows from having committed to that bet. In order of their presentations below, these assumptions are:

1. The cosmological constant is regarded on the "matter view" in GR.

2. The cosmological constant is "minimally coupled to matter" in GR.

[6]On the other hand, these are, after all, just arguments in favor of bets about the future theory; they are not deductive and we may turn out to be radically wrong. The point, in that case, is to understand what exactly we are wrong about now, when our bets later prove in retrospect to have led us astray. (This "retrospective" view is taken up again in section 3.3.)

3. Zero-point energies characterize semiclassical “zero-point stress-energies” in QFT.

4. Semiclassical stress-energies are “minimally coupled to curvature” in QFT.

As I will suggest, these assumptions are only sensible as premises in the context of particular interpretations regarding the physics within the two theories respectively.[7]

3.2.2 Getting to zero-point energies

GR is, from one perspective, a framework for particle dynamics in the presence of spacetime curvature. One traditionally begins with a relativistic spacetime, the points of which are interpreted as events, whereupon curves through the spacetime describe the possible trajectories of test particles of various kinds. For this reason, GR lends itself to the practice of model building in observational cosmology, where the physical scales of interest are such that the observed, reconstructed trajectories of whole galaxies may be regarded as the trajectories of test particles of a certain kind in some background spacetime.

But GR also provides a dynamical theory of the constitution of four dimensional spacetime, where the specification of suitable initial data along a three dimensional spatial submanifold (where such suitable data exists) uniquely determines the nearby characteristics of the four dimensional spacetime. This local determinism is governed by Einstein’s equation, which

[7] I also suspect that these various assumptions are where the four physicists mentioned in footnote 4 find cause to depart from taking the historically standard CCP seriously in the development of future theory. See footnotes 9, 10, 11, and 12 in section 3.3 below for more on this point. As one reviewer points out, what I am not doing here is arguing the necessity of these four assumptions in justifying the historically standard version of the CCP. It is enough for my purposes that such assumptions are sufficient to justify it, and that these assumptions are likely to be those which are held by the relevant communities that do indeed take it seriously. Sufficiency ensures that interpretations in our current theories which entail these assumptions likewise entail (by the above argument) taking the historically standard version of the CCP seriously. That these assumptions are likely endorsed by the relevant communities makes it promising that my analysis has identified the loci of disagreement, in practice, about whether physicists ought indeed to take it seriously. Concerning this latter point, I take the implicit satisfaction of these assumptions in the popular effective field theory approach (see §3.4) to provide at least some evidence that this is so. That the four physicists named previously each seem to depart from the standard perspective in virtue of disagreeing on each of these four respective assumptions provides further evidence along these lines.

may be written globally (with respect to a fixed manifold M) as

$$R_{ab} - \frac{1}{2}Rg_{ab} - \Lambda g_{ab} = 8\pi T_{ab}$$

where g_{ab} is the metric of the spacetime, R_{ab} and R are functions of that metric and its derivatives, T_{ab} is a symmetric tensor field on the manifold that is divergence-free with respect to g_{ab}, and Λ is a Real-valued constant.

T_{ab} is taken to encode the total stress-energy of some collection of matter fields defined on M. Where those matter fields admit of descriptions as fluids or (other) classical fields (e.g. the electromagnetic field), one may study the solution space of Einstein's equation to identify classical *cosmological models*, understood as spacetimes for which it is the case that the corresponding curvature is sourced by the matter defined on top of it. One commonly studied class of cosmological models is that whose constituents all feature the spatiotemporal structure of Minkowski spacetime, the geodesically complete spacetime whose manifold structure is $\mathbb{R}^4$ and whose metric is flat (so, notably, $R_{ab} = \mathbf{0}$ and $R = 0$).

As it turns out, when one commits in cosmology to modeling the universe at the largest scales as if all of the matter were (perturbed) perfect fluids (and one adopts certain guiding principles about the global structure of the universe), the observed, reconstructed trajectories of galaxies, understood as test particles in that spacetime, can be rendered as measurements of the constant Λ in Einstein's equation. This is the cosmological constant. Observe that it is present even in the "vacuum sector" of GR, when $T_{ab} = \mathbf{0}$. In what follows, we will only concern ourselves with that sector of the theory, i.e. situations in which Einstein's equation may be expressed as

$$R_{ab} - \frac{1}{2}Rg_{ab} - \Lambda g_{ab} = \mathbf{0}$$

Note the notion of 'vacuum' being deployed here: it applies whenever T_{ab} vanishes, irrespective of the statuses of whatever other fields one wishes to define on the spacetime. In particular, in cosmological contexts, one is free to entertain the possibility of vacuum models of various matter fields. For instance, any free (i.e. non-interacting) scalar field on a spacetime whose field values are everywhere-vanishing is standardly understood to carry everywhere-vanishing stress-energy. Hence, such a scalar field, understood as a (very boring) classical matter field, when defined on a spacetime for which the vacuum sector of Einstein's equation is satisfied, defines a vacuum model of cosmology.

In the vacuum sector, observe that we may just as well express Einstein's equation as

$$R_{ab} - \frac{1}{2}Rg_{ab} + [0]g_{ab} = \Lambda g_{ab}$$

where we have kept a vanishing multiple of the metric on the left-hand side in brackets. Since Λ is a constant and g_{ab} is (as a metric) symmetric and divergence-free with respect to itself, the term Λg_{ab} may just as well be regarded as a stress-energy in its own right, which is present even in the vacuum sector of the theory. Call this the "matter view" of the cosmological constant, because it interprets the (entire) measured cosmological constant as characterizing a quantity of stress-energy, and hence a kind of matter, that is associated with vacuum spacetimes.

On the face of things, all that has happened here in adopting the matter view is an algebraic manipulation. But the organization of Einstein's equation doubles as a helpful bookkeeping device, where the left-hand side defines a tensor with a geometric interpretation— the "Einstein$_{\Lambda_0}$" tensor for fixed $\Lambda_0 \in \mathbb{R}$ (written presently as the quantity in brackets)— which is dynamically constrained by the tensor on the right-hand side, interpreted as the total stress-energy associated with that geometric structure. From this perspective, one may understand the matter view of the cosmological constant to be diametrically opposed to another interpretation, call it the "gravitational view", in which the measurement of Λ— as a measurement of Λ_0— is understood as modifying the spacetime geometries that one associates with vacuum spacetimes. Note, though, that there exist a continuum of stances between these two, where the term in brackets in the most recent expression takes on any Real value. Each of these stances amounts to a different view about what GR tells us to conclude about the geometries of vacuum models. In exactly one of these stances, namely the matter view of the cosmological constant, Minkowski spacetime counts as a suitable vacuum geometry, characterizing inertial motions in a model where gravitational interactions with matter have been "switched off".

Now, observe that the matter view of the cosmological constant— understood as fixing a standard of inertial motions, absent gravity, that is independent of measurements of Λ— gives rise to certain expectations about the details of the theories that govern the matter fields within any particular vacuum model. Consider any such vacuum model, where we may simplify matters by assuming that there is only one species of matter in it. The stress-energy characterized by the cosmological constant on the matter view, within the context of such a model, implies that within the theory meant to govern that species of matter, there is a quantity of stress-energy that arises just by virtue of the matter being defined at all. This quantity of stress-energy does not depend on any dynamical features of that matter. Instead, it depends only on the structure of the background spacetime on which the matter is defined, as well as on some constant characteristic scalar that must show up within the theory itself.

What is happening here is that the presence of new terms on the right-hand side of Einstein's equation has created new obligations in the theory that is meant to govern the matter implicated in any particular cosmological model. Given, in particular, a vacuum model and the theory of matter implicated by it, it is clear that the cosmological constant must arise as a characteristic term within that implicated theory in such a way that is irrelevant to the (possibly degenerate) dynamics of the matter. In this way, one assumes that the cosmological constant arises as a term in the matter theory that, according to that matter theory, does not depend on the dynamics ascribed to that matter. That is, one assumes the cosmological constant in that theory to be "minimally coupled" to the matter.

But our best theory of matter to date as would be implicated in the standard model of cosmology is not given in terms of classical field theories; it is rather formulated in the framework of QFT (on Minkowski spacetime). In this way, the cosmological constant, on the matter view, is left wanting for a treatment in QFT, our premier theory of matter, as a characteristic term that may be associated with a quantized field, even when that quantized field is arranged in its degenerate or vacuum state. As we will see along the way in the next section, this is precisely what zero-point energies appear to be.

3.2.3 Getting to the cosmological constant

In broadest strokes, given— for pedagogical purposes— a free (i.e. non-interacting) scalar field on Minkowski spacetime, the quantization of that field consists in replacing field values at points in the spacetime with "smeared field operators" that are passed, as input, test functions taken from the collection of smooth fields on the spacetime that feature compact support. In the limit about a given point (in the sense of distributions on the spacetime), one may understand the smeared field operator in the quantized theory to correspond, semi-classically, to the map that takes that point in the spacetime to the value of the classical field evaluated at that point. That is, just as one may define various other classical observables (besides 'field value') by specifying certain pointwise operations on the classical field, one may define various other quantum observables (besides 'smeared field operator') in the case of the quantized field by specifying certain operations on those smeared field operators. As is the case for any (free) quantized system, states of the quantized field can be thought of as certain maps (satisfying various technical properties) from a set of observables to some abstract field, for instance the complex numbers. The image of a given observable in such a map is interpreted as the expectation value of the corresponding observable, when the quantized field is arranged in that state.

In this broad setting (and defined with respect to Minkowski spacetime), zero-point energies may be understood as a measure of the expected energy of the quantized field, when it is arranged in its vacuum state. More precisely, zero-point energies are the expectation values of some particular observable, call it energy$_Q$, when the field is arranged in a particular state for which it is the case that (globally) there are no excitations of the field and, moreover, given any two timelike-separated points in the spacetime, the correlation functions between the values of the field peaked over those two points depend only on the geodesic distance between them.[8] The latter of these two conditions witnesses the tight connection between the vacuum state of a quantized field and the inertial structure of the background spacetime; the former condition, meanwhile, picks out a global standard for lowest energy field configuration which is preserved under the global symmetries of that background.

I have not said anything of the construction of the observable energy$_Q$ in the quantized theory; here, it suffices to say that its construction bears strong analogy to the construction, in the corresponding classical field theory, according to which one assigns an energy density to the field. Unfortunately (and quite famously), constructions such as this in QFT, i.e. those which require taking products of smeared field operators in the coincidence limits about single points, lead to divergences. For this reason, various methods are introduced to regularize those sums, so as to recover finite expectation values. In the context of effective QFT, this is done by introducing a cutoff parameter based on high energy regimes where one takes the theory to no longer be applicable (for independent reasons); in more systematic approaches, this is done according to one or another renormalization prescriptions that are tailored so as to cancel out what may be quarantined as just the singular component of the particular sum. For our purposes, we may remain agnostic on the calculational point and

[8]See Redhead (1994) for a discussion about the vacuum state of a field as it relates to the physics witnessed by inertial observers.

merely assume that the resulting zero-point energies, however they are gotten, are formally well-defined and finite.

It is perhaps worth stressing here that there is no a priori reason for the (classical) energy density associated with a quantized field to agree with the expectation value of any particular observable constructed within the quantized theory. Given the above framing, this is easy to see: the only way we have to understand classical energy density is as a feature, namely an observable, of a particular physical theory that happens to be classical. Observables in a quantized theory are simply a different sort of mathematical object. Any relationship between observables in a quantized theory with observables in a classical theory must be put in by hand; an insistence that *some particular observable* in the quantized theory stand in such a relation is nothing short of an intertheoretic demand on the practice of physics— it is not a demand that any candidate quantized theory automatically satisfies (Feintzeig, 2017). Nonetheless, one generally assumes that such an identification can be made. By dint of the analogy between the constructions of various observables in QFT and the constructions of observables in classical field theory, one often assumes, as an interpretive resource in the quantized theory, what might be regarded as a "semiclassical translation dictionary." According to this dictionary, for instance, the expectation value of energy$_Q$ in the quantized theory is interpreted as an expression (perhaps averaged in some sense) of the classical energy density *associated with* the quantized system.

To sum up what has been said so far: via the availability of a semiclassical dictionary as just described, when a quantized field is arranged in its ground state, zero-point energies are expected energies gotten via measurements of a particular operator that corresponds, in the semiclassical dictionary, to the classical energy density associated with that quantized field in that ground state.

But we know that the appropriate, observer-independent geometric quantity that encodes the classical energy density of a field is a stress-energy tensor, so to be more precise in our semiclassical translation dictionary, we are looking for a "semiclassical" stress-energy tensor that is defined with respect to those zero-point energies. That is, we want a stress-energy tensor that, when acted on by unit four-vectors at any point in the underlying spacetime (that correspond to arbitrary inertial observers at that point), gives a measure of the "semiclassical" energy density associated with the field in that state. As it turns out, the only family of suitable stress-energy tensors that are invariant under the symmetries of Minkowski spacetime (and hence yield the same energy density with respect to arbitrary inertial observers) is the Minkowski metric, up to a constant factor. In other words, on the assumption that zero-point energies so characterize "zero-point stress-energies", up to a scaling constant there is only one way possible to define a semiclassical stress-energy tensor in those terms. Since that scaling constant amounts to a scaling of units in proper time along an inertial observer's worldline, we have the following: via the semiclassical dictionary mentality, the constant factor in the semiclassical tensor that encodes those zero-point energies is equal in value to the zero-point energies themselves. Recall now from the previous section that the cosmological constant also appears as a constant factor modifying the spacetime metric. In Minkowski spacetime, these two tensors are syntactically equivalent.

But we have already seen that the cosmological constant appears in Einstein's equation modifying an arbitrary metric, which means it is not enough that we have constructed a semiclassical stress-energy tensor syntactically equivalent to that term in Minkowski spacetime. For zero-point energies to want for a description as the cosmological constant in the relativistic theory, the semiclassical tensor that we have constructed needs to be the sort of thing that is defined in the same way, given any arbitrary metric. This modification has a name: the minimal substitution rule, and it is a familiar tool in classical field theories (see, e.g., Wald (1984, p. 70)). What this rule amounts to is a manipulation of syntax in the expression of a field equation originally defined on Minkowski spacetime: one simply replaces any explicit reference to the Minkowski metric with a reference to an arbitrary metric, and one replaces any explicit reference to the flat derivative operator (native to the Minkowski metric) with the derivative operator that is native to the arbitrary metric just introduced. In this way, one generally says that the resultant theory on the arbitrary spacetime is "minimally coupled to curvature", because it contains no explicit reference to curvature terms that would otherwise have been degenerate in the case of Minkowski spacetime.

Since the minimal substitution rule is a manipulation of syntax, its application to a particular expression therefore requires independent physical arguments: that the field in question is the sort of field for which the application of a minimal substitution rule is warranted. Note that this is a case (perhaps more so than any of the other three assumptions) where it transparently matters how one defends one's choices of interpretations. Why, for instance, should this semiclassical stress-energy tensor be understood as the sort of field that does not explicitly couple to curvature terms? On what grounds, in other words, does one assert that zero-point energies (granted already that they may be understood, semiclassically, as encoding the classical energy density of a quantized field in its vacuum state on Minkowski spacetime), when relocated to curved spacetimes, behave as fields that are minimally coupled to that curvature?

These are difficult questions that I will put aside, except to note that they are exactly the sort that philosophers of physics are already in the habit of asking. This is good! It means that philosophers of physics already engage in the sort of work on the basis of which we may anticipate details of future physics. Consequently, there is a natural way in which philosophers of physics can expect to contribute to the development of future physics: in establishing principled defenses for going about developing that future theory in some ways, rather than other ways, based on how that future theory is supposed to explain the successes of our current best theories so far.

3.3 The epistemology of physical interpretation

We have seen that the cosmological constant in GR and zero-point energies in QFT stand in a certain sort of global handshaking relationship, borne by interpretive moves concerning each of the two quantities respectively, entertained in each of the theories held separately. In particular, I have stressed that the identification of the handshake depends (precisely) on four premises: that the cosmological constant be interpreted on the "matter view" in GR,[9] that the cosmological constant be interpreted as "minimally coupled to matter" in the matter theories implicated within cosmological models in GR,[10] that zero-point energies characterize semiclassical "zero-point stress-energy" on Minkowski spacetime in QFT,[11] and that such a stress-energy tensor is appropriate for the application of the minimal substitution rule in QFT (i.e. it is "minimally coupled to curvature").[12] My central claim from section 3.2 (cf. section 3.2.1) is that if we believe these particular assertions about the physics that arises in the two separate theoretical frameworks, then the cosmological constant and zero-point energies plausibly describe characterizations of the same future physics in two different limits. In other words, such beliefs behave as warrants for the bet proposed earlier, and so render the historically standard version of the CCP as a problem worth taking seriously.

One worry about this account is that such beliefs— made on the basis of interpretations in GR and QFT held separately— are not (contrary to what was assumed in section 3.2.1)

[9]This is where Carlo Rovelli gets off board the historically standard version of the CCP: "It is especially wrong to talk about a mysterious 'substance' to denote dark energy.... It is like saying that the centrifugal force that pushes out from a merry-go-round is the 'effect of a mysterious substance' " (Bianchi and Rovelli, 2010, p. 1).

[10]This is where George Ellis should like to get off board, given his advocacy for unimodular gravity (cf. (Ellis et al., 2011)), wherein vacuum contributions in the stress-energy associated with one's theory of matter do not couple to spacetime curvature, but meanwhile the stipulation that total stress-energy be a locally conserved quantity gives rise to a cosmological constant of integration.

[11] If we understand this as the claim that zero-point energies are an appropriate quantity to expect to gravitate, this is evidently where William Unruh gets off board, given his recent (2017) co-authored work (clarified further in a (2018) note).

[12]Robert Wald must get off board at this point, as his axiomatic treatment of zero-point energies leads to curvature term ambiguities which, in order to break, he must assume behave in a certain way on Minkowski spacetime that would give rise to other values for zero-point energies in other spacetimes (cf. (Wald, 1994, pp. 85-97)).

themselves justified by all that we presently have to work with. After all, our empirical observations, as they are understood in the contexts of our best physical theories to date, are empirical observations *as interpreted by mutually conflicting theories.* If we already know that something or other from each of the relevant theories will have to give, one fears placing too much justificatory weight on those parts therein that may just as well soon be discarded.

This worry is reminiscent of a classic discussion by Belot (1998) on interpretations of electromagnetism in light of the Aharonov-Bohm effect. In that discussion, Belot (1998, p. 532) effectively begins with the following moral, or what he calls "the kernel of the common wisdom" about the Aharonov-Bohm effect: "*until the discovery of the Aharonov-Bohm effect, we misunderstood what electromagnetism was telling us about our world*" (emphasis in the original). According to Belot, what this moral says is that, regarding those aspects of our world that we can understand correctly by virtue of the theory of electromagnetism being successful in limited domains, our classically-motivated interpretation of that theory was leading us astray. In this way, it is easy to imagine our worry above as actually realized in that older episode in the history of physics. Conceivably, anticipations about the theory that came to be quantum electrodynamics, whose relevant bets depended on that misunderstanding, would have been misguided.

The prima facie puzzle that Belot goes on to discuss is how there can be such aspects of our world, concerning which interpretative work in a known-to-be-false theory could lead us to misunderstandings. In our case, this is to ask: how can such anticipations as just imagined turn out to have been misguided? He spoils at least part of his conclusion within the introduction (p. 533):

> Thus we find that the requirement that our false theories [electromagnetism, non-relativistic quantum mechanics] mesh in an appropriate way—ontologically as well as empirically—places strong constraints upon our interpretative practice.

The key point may be put like this: there is knowledge about the world to be gotten by virtue of the fact that *these two false theories in particular* co-exist, messily, in our contemporary corpus of physical theory. Considerations of only one or the other of these false theories in isolation would necessarily miss the signposts toward that knowledge (and so in our imagined case, we were misguided insofar as we, by hypothesis, missed those signposts). On the other hand, requiring, in our interpretations of each of the false theories, that those false theories mesh well provides a way of generating some of that knowledge.

In present context, Belot's conclusion appears to constitute a challenge to any claim that we warrant bets about future physics via interpretative work in our current theories held separately. Contrary to our initial expectation, there is at least some sense in which those successes of our current physics include *more* than just what is to be interpreted in each of the two theories on its own. As such, the possibility of misunderstanding what either GR or QFT on its own is telling us about the world threatens the integrity of the justificatory chain intended to ground our bet about future physics in facts about the successes of our current best physics so far. We might turn out to be misled, that is, in virtue of our interpreting each of the theories separately and disregarding all of the ways in which our current best physics ultimately includes both.

I agree with much of Belot's original argument, and so this present challenge strikes me as serious. What is at stake here is not only the justification for my committing to such bets (such as would entail my forming beliefs about the details of the future theory as discussed), but also the status of my beliefs subsequently formed. Whether those beliefs of mine constitute my anticipating future physics depends not only on my having committed (for better or worse) to the relevant bets, but on the reasonableness of my having done so.

Nonetheless, I think it is important to recognize that Belot's view is, essentially, retrospective (and, as will become clear, this is sufficient to defuse the challenge). That is to say, his view takes, as its starting point, that the goal at hand is to understand the world, *given, in* *retrospect, the best theories we have so far managed to articulate.* One way this shows up is in his assessment of what, ultimately, goes into interpretive work, understood as additional structure placed on top of the formalism of a physical theory (cf. p. 551): purely metaphysical views and our beliefs about the structure of our own world. Making no stance in regards to the first of these, the second seems to me to beg the question in many cases of interest, at least in the context of frontier physics research. This is because our beliefs about the structure of our own world are, in a very real sense, precisely that which are being formed as we go about our theorizing. Our trust in any tentative guesses on this front hardly seems to be the sort of thing that would provide sturdy foundations for interpretive work in the current theories we intend to soon discard.

A tidy demonstration of this point can be found incidentally in an article by Teh (2014) on the subject of whether lower-dimensional GR can be regarded as a gauge theory, such as could (perhaps) be quantized in ways similar to other field theories familiar in contemporary physics.[13] I have in mind a small remark therein (p. 510), where Teh points out that different interpretations in GR lead to dramatically different quantizations, such that "if we mean to interpret the classical theory in light of its quantization, then we should take into account the fact that different actions will in general give rise to different quantizations." On the other hand, Teh then goes on to offer a purely classical argument for regarding two particular actions as distinct interpretations of GR (wherein one of them permits—in the language introduced above—cosmological models that exhibit spinor fields, and the other does not).

In the context of building a theory of quantum gravity in the first place, the former approach—interpreting in light of expecting to quantize down the road— does not provide any leads as to which interpretation might best capture what we know about the physical world. On the other hand, the latter approach— interpreting via classical considerations— does so

[13]Though, see also Fletcher et al. (2018) for broader concerns about the status of physical interpretations in lower-dimensional GR.

transparently, in at least one sense: if one has arguments to the effect that the cosmological models of GR should include spinor fields as possibilities, then one of the two interpretations discussed by Teh ought to be picked out in favor of the other. That is, interpretive work in the purely classical regime (at least if our universe admitted one fewer dimensions) is what would ultimately pick out the theory that one feels is worth it to quantize in the first place.[14] Likewise, in the present case of interest, whether the historically standard version of the CCP is worth worrying about in the first place is exactly what is up for grabs.

For this reason, I would like to distinguish between the epistemological context of Belot's project and my own in the following way: whereas his is a "past-looking" project, concerned with an assessment of what we know about the world based on our best theories so far, mine is a "future-looking" one, concerned with an assessment of what we may anticipate about the contours of future theory based on the same.[15] I am quite happy, that is, to concede Belot's point that how we interpret known-to-be-false theories ought to be constrained by considerations of our beliefs about where future theory will soon take us.

On the other hand, in the practice of building toward that future theory, I contend that it is the identification of where our current theories are supposed to mesh, via their respective interpretations in isolation, that paints the contours of our frontier research. Belot writes (p. 553) that "...it seems essential to demand for every pair of overlapping theories an assurance that their empirical predictions mesh in the appropriate manner." In other words, what I care about is already assumed in his antecedent: what justifies a belief that such

[14]One might object that spinors are hardly to be understood as "purely classical", because all of the usual motivations for them are born in the wake of quantum theory. Perhaps this is so, in which case the story here is more complicated. But inasmuch as spinors are defined classically on a manifold (and relate, e.g., to familiar classical concepts like the existence of a conformally flat metric), one can nonetheless imagine— in principle— motivations from classical physics for their inclusion.

[15]Following Barrett (2008), there is another distinction that one could draw to help frame what I am doing here: between an understanding of our current theories in terms of their truthlikeness and an understanding of our current theories as snapshots in diachronic inquiry. On one particular reading of Belot (1998), wherein an interpretation in the classical theory is favored *on the basis of it being less false* (in a world whose particles turn out to behave quantum mechanically), this is perhaps the more valuable distinction to make for the purposes of framing my project. Nonetheless, I prefer the distinction made in the body of this chapter, because it seems to remain appropriate under a much broader reading of Belot's argument.

theories indeed overlap any one way, rather than any other?[16] And here, it seems clear that the answer has to be given in the terms specific to each of the relevant theories, which is to say in terms of the physical world *as it is interpreted by the lights of each such theory held separately.*

If, as I have just claimed, one must leverage each of our current theories separately from one another in the context of frontier physics research, what should we make of the initial worry, amplified to the status of a challenge by Belot's work, that we will be misled in virtue of doing so? Here is where, I believe, the past-looking/future-looking distinction is particularly useful. Imagine again the situation wherein there is some anticipation about quantum electrodynamics that follows from a bet made on the basis of a classically-motivated interpretation of electromagnetism. Today, (following Belot) we would recognize that anticipation as having been misguided, precisely to the extent that it depends on what we now (by hypothesis) recognize as a misunderstanding of what electromagnetism tells us about the world.

But note that it is only *in retrospect* that the anticipation is misguided, as a consequence of the relevant bet taking for granted an interpretation that we have *since found reason* (namely, in the discovery of the Aharonov-Bohm effect) to discard. Prior to that discovery, the bet was, by hypothesis, reasonable. That is to say, *prospectively*, the imagined anticipation about quantum electrodynamics was perfectly apt, for as long of a time as we lacked the additional information provided by the discovery. This imagined example captures the fact that our original worry— that we are placing justificatory weight on something that we may soon be compelled to abandon— is merely a reflection of how it is that, in the business of

[16] To be clear: I do not take myself to be arguing here for a distinction that Belot would find controversial. Indeed, he as good as states the distinction himself in the middle of his article (p. 546):

> One's interpretative beliefs can shape one's judgments as to the relevance of certain quantizations or approaches to quantization. Conversely, one must accept that one's interpretative beliefs are open to revision in light of the empirical success of the approaches to quantization which they suggest.

Belot clearly cares foremost about the revision of one's interpretive judgments in light of new beliefs about the world. But he is also quite happy to contend that interpretive beliefs shape one's judgment about how to move forward in the first place.

anticipating future physics, we ought always to be prepared to revise our beliefs in light of new information encountered along the way. The subsequent challenge, meanwhile, is a red herring: there is nothing wrong, epistemologically speaking, with the state of *having been* misled.

As I have said already, it is not my intention in this chapter to argue for or against the particular interpretive commitments that I have insisted would justify the historically standard version of the CCP. And so, likewise, it is not my intention to argue what discoveries have (or may) come about that would point to our having been misled by those commitments, in virtue of other commitments being on better footing. But such debates are ones that philosophers of physics would do well to have for the sake of both anticipating details of the future physical theory and working with physicists toward the wholesale development of it. In this context, I have suggested that the way to have these debates is by considering each of our current theories in isolation from each other for as long as possible. At the end of the day, with well-argued interpretations settled in each of the theories independently, conflicts in the "overlapping" areas *revealed in the first place by those interpretations* constitute justified bets about the future theory: that there are particular, articulable problems arising in the context of that future theory, which the future theory will, in its details, resolve.

3.4 The view from high energy physics

In contrast to what I have advocated so far, one finds a very different attitude in the high energy physics/effective field theory community, in which our current frameworks of GR and QFT are never held apart in mutual isolation. (In philosophical circles, the attitude of this community is defended most adamantly in the view put forth by Wallace (2006), but see also Crowther (2013).) For this reason, it is worth remarking briefly on that view in the context of the CCP. Indeed, as will be discussed shortly, from the perspective of this community, the severity of the historically standard version of the CCP appears inescapable.

Space will not permit me to elaborate on the possible virtues or vices of the effective field theory approach, which attempts to bring together all of contemporary physics under one interpretation. Instead, I must content myself with briefly sketching how the resulting interpretation implicitly engages with the four assumptions I described above.[17] To be sure: just as my claim above was neither to argue for nor against the interpretations that support the four assumptions needed to take seriously the historically standard CCP, my claim now is not that this community has been led astray regarding the CCP by virtue of their attitude opposite what I have advocated. Nonetheless, insofar as it is opposite, there is reason down the road to study the arguments in favor of that view, such as would justify (what I argue shortly) is their running together the threads that I have endeavored to keep separate. On the view I have advocated above, just such a study is what is needed in order to justify taking the historically standard version of the CCP seriously *by virtue of its inescapability on this approach.*

One way of presenting all of the predictive power of contemporary physics is by writing down a single (classical) Lagrangian for all of the matter fields implicated in the standard model of particle physics, a prescription for how to quantize the free and interacting terms within it, a list of cutoff parameters (by which the various divergent sums that would otherwise arise in the aftermaths of those quantization prescriptions are regularized to finite values), and to include, at the end of the long expression, terms familiar from the Einstein-Hilbert

[17]Indeed, for readers not already invested in the effective field theory view (nor in its relationships to other perspectives in frontier physics), this section is supererogatory to the core aims of the chapter. The short summary of the section is this: a highly popular view of the success of contemporary physics seems to support an interpretation of contemporary physics that implies each of the four assumptions I discuss above, from which we get the historically standard version of the CCP. Whether the historically standard version of the CCP is worth taking seriously as a consequence of this being so will depend on the arguments offered in favor of the popular view in the first place. That is to say: a justification for the effective field theoretic interpretation of contemporary physics should, by my arguments above, be the sort of thing that would justify taking seriously the historically standard version of the CCP.

action as well (i.e. the action from which one may recover Einstein's equation in the vacuum sector), including the cosmological constant. Just as when one varies the action defined by that Lagrangian with respect to the metric in the classical case, the cosmological constant retains its gravitational significance from GR in the quantum case by providing a zeroth-order approximation to the path-integral determined by that Lagrangian.

But meanwhile, zeroth-order terms in the rest of the action, found in the components of the Lagrangian that correspond to each of the standard model matter fields, also take on gravitational significance in the same way. These values make no difference to the predictions of QFT absent considerations of gravity, but are fixed via the regularization prescription alluded to above (in terms of the chosen cutoff parameters). When one introduces gravity, those fixed values provide an enormous contribution of opposite sign to the very small cosmological constant measured via the standard model of cosmology. Hence, the fundamental cosmological constant at the very end of that long Lagrangian has to nearly exactly compensate the effects of all of the other terms. The demand that the fundamental cosmological constant provides this extraordinarily large and exact compensation is generally regarded by this community as physically unreasonable (at least absent additional mechanisms to explain that compensation). Consequently, the discrepancy between the total sum of contributions from the matter fields and the measured cosmological constant is considered in obvious need of resolution, which is just to say that this community takes the historically standard version of the CCP very seriously.

In this framing of contemporary physics, one can get the sense that the historically standard version of the CCP is inescapable. This may be so, but it is important to see how that inescapability can be traced back to what has already been assumed of the four premises about current physics that I emphasized earlier. On the GR side, the distinction between the matter view and the gravitational view (and every other view in between) is seemingly disregarded, in favor of an insistence that large and exact compensations by the fundamental cosmological constant are physically unreasonable, absent additional mechanisms. This is the naturalness problem, and is the source of its own philosophical discussion (cf. (Williams, 2015) in the context of the Higgs sector and references therein). But from the perspective taken in this chapter, the argument against the naturalness of the cosmological constant providing large and exact compensations is, in effect, an argument that measurements of the cosmological constant cannot bear on the question of what we take to be the inertial structure of the vacuum, prior to or absent considerations about the interactions between gravity and matter.

Meanwhile, the second premise articulated above— that the cosmological constant from GR characterizes a field "minimally coupled to matter"— is stipulated outright. In taking the Einstein-Hilbert action to be the (leading) contribution to the gravitational portion of the path-integral in light of the success of GR as a large-scale theory of gravity, one has just as well regarded the fundamental cosmological constant as directly analogous to all of the other zeroth-order terms present. But this need not be the case, as is demonstrated by various dark energy proposals in the very same tradition.[18]

On the QFT side, that stress-energies associated with each of those zeroth-order terms from the matter contributions resemble cosmological constant-like terms on Minkowski spacetime, and that these terms each minimally couple to curvature, each come for free with the decision to include curvature invariants in one's derivation of the path integral. The only reason that this practice is well-defined, however, is because one has already insisted on a shift from expressing the matter theory classically as a Lagrangian, whose action is understood as taken with respect to the volume element associated with Minkowski spacetime, to a Lagrangian density, whose action is understood as taken with respect to the volume element associated with the metric that determines those curvature invariants. This decision is, in some sense,

[18]That dark energy proposals which toy with this assumption count as worthwhile projects highlights the difficulty in assessing when purported solutions to the CCP are indeed solutions as such (without having agreed in advance about the nature of the problem to be solved). This is, more or less, my observation in the previous chapter.

an assumption in its own right of minimal coupling to curvature: it supposes that the way to introduce the metric as a new interaction is, from the classical perspective, as a new independent field.

3.5 Discussion

I have argued that we can anticipate details of future physical theories in a principled way. As such, I have also suggested that the pursuit of such future theories in the first place, in particular regarding a theory of quantum gravity, may be approached systematically and on principled grounds. One consequence of the particular analysis I have given is that it is philosophical work in the foundations of our current theories that is well suited for uncovering the particulars of that systematic approach. Namely, work concerning interpretations in our current theories, understood in each of their own rights, can give rise to precise bets, the commitments to which entail particular beliefs about details of future physics. That is to say, it is via interpretations in our current theories that we may locate concrete problems today for which it is reasonable to bet that future theory will, in its details, resolve.

I have also explicitly contrasted the suggestion here with the view developed by Belot (1998), also regarding interpretations at the cross-roads of contemporary physics. With no criticism of his past-looking view, making sense of the structure of our physical knowledge given our best, albeit strictly incompatible, theories so far, I have emphasized that there is, as well, a future-looking view, which provides grounds for developing new theories *from those very present, incompatible theories.* In the future, those new theories will hopefully come to furnish us with even higher-fidelity knowledge of the physical world than that which can be gleaned from recognizing Belot's point (i.e. that there is knowledge to be gotten from the fact that it is our present theories *in particular* that are what co-exist—however messily—in our present corpus of physical theory).

In the particular case of quantum gravity research, the future-looking view articulated here has brought attention to several assumptions about interpretations in our premier theories of gravity and matter respectively, on which the popularity of the historically standard version of the CCP seems to rely. If those assumptions are not themselves well-motivated (because interpretations that do not entail them are favored in accounting for the successes of the respective theories), so be it, and physicists ought perhaps to move on to other versions of the CCP that may arise by virtue of those other interpretations (cf. Chapter 2), or other problems altogether (cf. (Koberinski, 2021)). If, on the other hand, a physicist should find them to be well-motivated by the successes of the respective theories (perhaps for reasons related to the endorsement of a view like that discussed in section 3.4), then they better bet that the future theory of quantum gravity will explicitly resolve the historically standard version of the CCP, skeptics be damned.

Chapter 4

Creativity in the social epistemology of science

4.1 Introduction

Currie (2019) argues that research in existential risk (‘X-risk’) should be more creative than it likely is. In the course of the argument, he introduces a general account of creativity in scientific discovery (hereafter, ‘creativity’). This account is intended to capture how conservatism can be detrimental to the health of inquiry in certain scientific communities, given certain aims of research. It is also advertised as complementing the use of formal landscape models in studying policy initiatives within the social epistemology of science.

Independent of Currie’s project, Rovelli (2018) decries a “why not?” ideology he reports is in vogue within his scientific community, engaged in fundamental physics research. By his reckoning, this ideology promotes a method of guesswork. His concern is that such a method is detrimental, given facts about his community and their research aims.

Here, I will argue that Rovelli's remarks, when interpreted in light of Currie's account, raise trouble for the general applicability of the latter. This is because Currie's account fails to countenance the possibility that revolutionary theorizing might be valuable, as features in Rovelli's argument. As a result, it is unclear when Currie's account is actually appropriate for studying the effects of conservatism on the health of inquiry. This threatens to undermine the use of such an account in arguments undergirding policies meant to respond to conservatism. It would be prudent to seek out means of identifying what it is about any given scientific community that could render Currie's account appropriate there.

4.2 Creativity in science

Stanford (2019) has argued that the structures and institutions of contemporary science foster conservatism in research, stifling revolutionary theorizing. Currie (2019) is concerned that the same conservatism is detrimental to inquiry within X-risk. This is because, according to Currie, X-risk is best pursued creatively. Arguing that creativity is in tension with conservatism, Currie concludes that the scientific community focused on X-risk is likely insufficiently creative. The final pages of his article raise a question about what steps, if taken, might correct that insufficiency.

As just presented, Currie's project depends essentially on his providing an explicit account of creativity within a scientific community. The remainder of this section is dedicated to describing the account he provides, as well as developing it further (where necessary) in a friendly manner.

Consider the situation wherein there is some well-posed problem, whose solution a scientific community agrees constitutes the aim of their collective research. The statement of the problem places severe constraints on what counts as viable research within that community, united by that aim. We may think of the statement of the problem as characterizing the research program pursued by that community. And associated with that problem is, following Currie, a collection of possible solutions. This 'solution space' is meant to be roughly coextensive with all professional moves available to members of that community, engaged in that research program. The researchers occupy points in the solution space, and they choose which points to occupy next.[1]

Currie introduces into this picture the following two metaphors. 'Hot searches' through solution space are energetic; 'cold searches' are the opposite. A hot search refers to a sequence of points, whose iterative selection by a theorist describes that theorist as hopping around through the solution space. A cold search refers to a similar sequence of points, except that it describes the behavior of a theorist who is nearly staying still.

To make these metaphors, Currie needs a notion of distance between points in the space. He borrows from Bayesian epistemology to develop one. (I will have more to say that is critical of this below.) By his reckoning, distances to solutions are relativized to each individual at a time, and are indexed to that individual's credences at that time. So, roughly speaking, solutions assigned low priors are far, and solutions assigned high priors are near.[2]

Currie does not elaborate on the interpretation of priors over points in the space. Evidently, he has in mind something pragmatic: "Our priors serve to set expectations across a space of possible solutions to a problem" [p. 6]. In this respect, the account is non-committal about what it is that ultimately makes a solution worth visiting. We are free to suppose that there is some unspecified constellation of virtues, possibly specific to the research program at hand, that one hopes is jointly maximized (i.e. via some method of aggregation) by

[1] In fact, there are other professional strategies that are ultimately available to researchers, regarded as decision-making agents. Whether activity gets channeled into those other strategies, rather than into moving between solutions, is an important degree of freedom in Currie's account.

[2] As will become clear, it may be that we ought to insert a ceteris paribus clause here. If so, we would say that whatever are otherwise the distances to solutions, those numerical values are then systematically deformed to reflect comparative facts about one's priors over each solution.

whatever solution is visited next. On this picture, hot searches are sequences for which the researcher's decisions are insensitive to their beliefs about where it will be prudent to visit. Oppositely, cold searches occur when the researcher's choices correlate strongly with those beliefs.

Currie then defines an agent's creativity in terms of their propensity for hot searches, given this picture. In other words, an agent is creative in proportion to the unconditional likelihood that they attempt a distant, low-credence solution. A community's creativity, meanwhile, is defined to correspond with what would generally occur if the members of the community were all individually creative. The upshot is that a community's creativity is defined as proportional to the efficiency with which they explore solution space widely. (What it means to explore widely is, of course, agent-relative. Here, we might assume that a community explores widely when it does so by the lights of most of its members.)

This behavior is in contrast with what, following Currie, we may call 'pooling'. Pooling describes a situation wherein a community fails to explore their solution space widely. Intuitively, pooling occurs when individuals within the community fail to be creative, each favoring cold searches instead of hot searches. But, as Currie notes, pooling may be avoided in such a case, provided that the community is cognitively diverse. So long as cognitive diversity is understood in terms of diverse distributions of priors, cognitively diverse individuals engaging in cold searches will, collectively, explore widely. This community would count as creative, according to Currie, even though the individuals who comprise it do not.

The creativity of a community is therefore not uniquely determined by facts about the creativity of its constituents. Their propensity for peer disagreement (and so, the social structure of science, etc.) also matters. And according to this view, a community may be made more creative in various ways. One way is by interventions to promote sustained cognitive diversity, as we have understood it here. Another is by incentivizing hot searches, or increasing creativity at the individual level. In both cases, pooling behavior is reduced, in favor of wider exploration.[3]

Building on recent work by Stanford (2019), such interventions are, according to Currie, in contrast with the unchecked effects of conservatism in professional science today. This is because, according to Currie, conservatism promotes pooling behavior, as we have understood it here. But depending on the given research program, it may or may not be detrimental that science today is, generally, conservative. This is because a research program ought to be assessed individually, according to the "local details" [p. 3] relevant to it. Those details determine, for instance, whether the community is better off investing in strategies other than those relevant to scientific discovery (cf. footnote 1 above). If so, any resulting pooling according to shared priors need not be unhealthy.

As just stated, the utility of Currie's account is ultimately going to rest on certain further facts. These are facts concerning what kinds of local details might we recognize as rendering creativity— as opposed to pooling— a standard of good epistemic health in the community. Any such local details are encoded, we may suppose, in the statement of the problem that constitutes the aim of that community's research. Recall that it is from this problem that, in principle, we may extract the parameters of the solution space we envision that community to explore. It follows that assessments of the local details of a research program will generally shape our expectations about the solution space associated with the problem. Likewise, facts about a solution space can correlate with facts about whether pooling or creativity is preferred in the corresponding research program.

[3] A third way to increase creativity, noted by Currie, is to impose on the community a diverse collection of search algorithms. But this raises a question: what distinguishes, in practice, our imposing a diverse collection of search algorithms from our incentivizing hot searches? At the level of analysis presently provided, it is unclear that there is any distinction. As suggested in footnote 2, it may be that we should ultimately think of solution space as admitting some intrinsic structure, independent of credences. In that case, search algorithms could be defined with respect to that intrinsic structure, and would generally result in searches that appear hot.

Unfortunately, Currie does not state how such a correlation would work. This omission could suggest that we ought not to regard local details as shaping our expectations about solution space (besides via shaping our priors). But this would render Currie's account in tension with the standard interpretation of formal landscape models. Currie regards the use of such models within the social epistemology of science as complementing his approach (cf. p. 11 in the article). In such models, one typically regards the intrinsic structure of the landscape as an independent variable, whose possible values encode arbitrary research environments. So too, we might conclude, the structure of a solution space should reflect facts about the corresponding research program.

In light of this, I think it is appropriate to regard Currie's discussion of X-risk as illustrating the reasoning that would shape the relevant solution space. His ultimate conclusion is that X-risk should be creative because it should be "multi-disciplinary, pluralistic, and opportunistic" [p. 26]. We might speculate, on the basis of this, that the local details relevant to the problem of X-risk render the solution space as unusually vast.[4] In a vast solution space, cold searches could seem unfruitful, no matter how cognitively diverse we may plausibly imagine are the researchers. Consequently, creativity is generally preferred in such a case, consistent with Currie's reasoning about X-risk.

To recap: treating research programs as solution spaces, creativity is a matter of how the relevant communities explore those spaces, given some priors. Conservatism in science encourages pooling according to shared priors, which is opposite creative exploration. But specific facts about the solution space at hand can determine, in a given community, whether creativity or pooling ought to be preferred. Those facts are ultimately grounded in the statement of the problem identified by that community as constituting their research program.

[4]There is room for disagreement here. For instance, Currie's discussion of X-risk places some emphasis on its normative aspect— i.e. threat mitigation— and its role in the public eye. It is not clear what these would have to do with the size of the solution space. But perhaps these aspects of X-risk shape the space in fine-grained ways other than its size. For that matter, there is considerable ambiguity in my size term "vast", given that distances in the space are indexed to credences. This ambiguity motivates a revisionist attitude toward distances in the space. (See also footnotes 2 and 3 above.)

4.3 The situation in fundamental physics

Consider now the article by Rovelli (2018). Rovelli is a theoretical physicist focused on quantum gravity, the problem that characterizes fundamental physics research today.[5] Shortly, when it becomes prudent to bring in Currie's theoretical framework, this will be the problem understood to shape the relevant solution space. It is in regard to that solution space that I take Rovelli to have both expertise and privileged access.

Rovelli's article is adversarial. Our attention is best directed to a passage that comes in the middle, immediately following his presentation of what he calls the "why not?" ideology. According to Rovelli, this uncritical ideology is responsible for the rise of a damaging method of guesswork in contemporary fundamental physics practice. The criticism of the method proceeds as follows [p. 7]:

> Arbitrary jumps in the unbounded space of possibilities have never been an effective way to do science. The reason is twofold: first, there are too many possibilities, and the probability of stumbling on a good one by pure chance is negligible; but more importantly, nature always surprises us and we, the limited critters that we are, are far less creative and imaginative than we may think. When we consider ourselves to be "speculating widely", we are mostly playing out rearrangements of old tunes: true novelty that works is not something we can just find by guesswork.

As in Currie's article, we have here a spatial account of scientific discovery. Scientists decide how to move amongst points in the space (now, of 'possibilities', rather than 'solutions'). The role of the "why not?" ideology is to support a method of guesswork. We can understand

[5] This is, of course, a massive simplification. But so too is the problem characterizing X-risk in Currie's project. Whether the simplification is tolerable despite such objections depends on the particular context of its use.

this method as a decision procedure, the repeated execution of which amounts to "arbitrary jumps" in the space. (More formally, we might think of such a method as analogous to Monte Carlo sampling, with respect to some unspecified probability distribution on the space. Based on the context surrounding the quoted passage, Rovelli clearly has in mind a distribution that is meant to be uncorrelated with one's priors.) But absent any greater detail about the account Rovelli envisages, it is unclear why such a method should be as damaging as he claims. Prima facie, Currie's account of creativity should be helpful as a means to interpret the argument.

In Currie's framework, Rovelli's 'space of possibilities' may be understood as a solution space for the problem of quantum gravity. The solutions to the problem are, then, candidates for what may turn out to be a satisfying theory of quantum gravity. Given this reading, Rovelli's principal claim about the space is that it is vast. This seems right. In other contexts, this space is taken to be synonymous with 'theory space', the collection of all possible fundamental theories (see, e.g. (Dardashti, 2019)). From here onward, I will adopt this 'theory space' language when talking about the space of solutions relevant to the problem of quantum gravity.[6]

Recall that creativity at the community level is spelled out, on Currie's account, in terms of exploring widely in the relevant solution space. I have suggested that we understand Rovelli's remarks in terms of fundamental physicists exploring the vast theory space corresponding to the problem of quantum gravity. Since the space is vast, by the argument at the end of the previous section, creativity is likely encouraged. In other words, a more creative community is likely better off, given the local details of the problem of quantum gravity. Wider exploration should be good here.

[6]In Chapter 2, I criticize the relevance of this 'theory space' view in assessing the methodology of quantum gravity research.

Meanwhile, fundamental physicists are, according to Rovelli, *uncreative* (or, at least, are "far less creative" than they may think).[7] On the present interpretation, this would suggest that fundamental physicists fail to explore widely. Increasing creativity should be desirable.

Naively, guesswork is one such method to do so. (As described above, except if the sampling is with respect to a probability distribution correlated with one's priors, guesswork will generally produce hot searches.) On Currie's account, we may thereby understand Rovelli to hold the view that the method of guesswork happens to be implemented poorly by his community. Moreover, according to Rovelli, when his community engages in guesswork, they fail to speculate as "widely" as they typically believe themselves to speculate. So: the community does not explore widely, and they fail to recognize that this is the case.

This seems to provide a sufficient reason that the method is, according to Rovelli, damaging. Because theory space is vast, creativity constitutes a standard of good epistemic health in contemporary fundamental physics. Meanwhile, the community's poor implementation of guesswork fosters an exaggerated perspective as to how healthy their inquiry really is. Our initial hunch was correct: Currie's account of creativity can help us get traction on Rovelli's argument.

Yet, there is something unsatisfying about this interpretation of the argument. Consider the reason that Rovelli supplies for his testimony that the community implements the method of guesswork poorly. The poor implementation is due to the fact that "we, limited critters that we are, are far less creative and imaginative than we may think". In other words, guesswork is implemented poorly by his community, because their being limited ensures that they *cannot implement it well.* In particular, it his community's lacking creativity (and imagination), on this interpretation, that ultimately bears responsibility for the method being damaging.

[7]What relation this testimony could bear to the broader conversation about conservatism in science is interesting to consider, but a tangent at present.

Whether Rovelli's argument is compelling, so interpreted, is therefore going to turn on whether a community's lacking creativity can be understood to intervene on the efficacy of a method they attempt to employ. And here, Currie's account provides little guidance. Facts about the community's pooling with respect to shared priors cannot obviously prohibit researchers, all of whom are willing to speculate irrespective of their priors, from doing so. In this respect, Rovelli's argument depends on creativity (or the lack thereof) playing a further role in the social epistemology of his community than is readily countenanced by Currie's account.

As I will argue presently, Rovelli's argument is ultimately compelling, provided that we attribute to him the view that revolutionary theorizing is valuable in contemporary fundamental physics. On the other hand, recognizing the importance of such a view to Rovelli's argument should make us wary about the general applicability of Currie's account. In particular, it is no longer the case that Currie's account can be assumed to capture how to assess the epistemic impact of conservatism on a research program, for which creativity is healthy. A further question about whether or not revolutionary theorizing is valuable complicates the assessment.

4.4 Revolutionary theorizing and the health of inquiry

Suppose that there exist possibilities in theory space that, for all intents and purposes, are assigned prior probabilities of zero by all members of the community. Whereas many possibilities are *accessible* to the community, in virtue of being assigned non-zero priors by someone, these further possibilities are *inaccessible*. On Currie's terms, these are possibilities that are located an infinite distance away from the community, and are regarded as infinitely less promising to visit than any accessible possibility.[8]

[8]Assignments of zero-probability priors to non-contradictions are antithetical to an orthodox Bayesian epistemology. So, it is not obvious that the present supposition, in the case of theory space, is faithful to

In such a case, no matter how creative the community is regarding the accessible possibilities, some of theory space will never be explored. So, provided that guesswork fails to be defined over inaccessible possibilities, the method could fail to spread the community as wide as might, ultimately, be desired. This idealized setup sounds promising as a means to recover why, according to Rovelli, his community cannot implement guesswork well. We need only to attribute to Rovelli two further claims. The first is the the claim that his community's lack of creativity results in there being some possibilities that are inaccessible. The second is that at least some of those inaccessible possibilities are important to the aims of his community's research.

Evidence that Rovelli would endorse each of these claims may be found within the passage already quoted. Namely, what is inadequate about guesswork, says Rovelli, is that it does not yield "true novelty that works". This is because employing it results (instead) in "playing out rearrangements of old tunes". If we interpret the rearrangements of old tunes as the accessible possibilities, his claim is this: what there is to be sought in fundamental physics— i.e. true novelty that works— in fact resides in the inaccessible part of theory space.

Suppose that this reading is correct, and what there is to be sought in fundamental physics is, according to Rovelli, presently inaccessible. Then it is a symptom of the community's not being creative, according to Rovelli, that the implementation of guesswork necessarily fails to engender wide enough exploration. This is because the relevant sampling procedures fail to be defined over the whole of what is *worth exploring*.

We have thereby found a means to articulate the lingering part of Rovelli's argument, which we were unable to do in the previous section. Namely, says Rovelli: what is worth exploring fails to be coextensive with the accessible part of theory space. As a result, guesswork is

Currie's project. Nonetheless, given some other structure to the space (cf. footnotes 2-4), we may understand zero-probability priors as an idealization that "pushes off to infinity" the corresponding possibilities. They are, in effect, disconnected from the accessible possibilities. No amount of information gleaned from work on the latter could ever reign them in.

ineffectual. Worse, employing the method misleads the community in their self-assessment of whether they are sufficiently creative, consonant with their research aims. This is because the method only promotes wide exploration of a kind that is unsuitable for assessing the health of inquiry in fundamental physics. It only countenances that which is *conceived as* worth exploring (i.e. rather than what *is*).

If this is how we are to understand Rovelli's argument, it is easy to generalize the lesson. Consider any context wherein one has reason to regard the accessible part of solution space as failing to include some of what is worth exploring. This is a context in which genuinely revolutionary theorizing is needed, which renders accessible more of the space. Exploring just the accessible part necessarily fails to engender explorations that are sufficiently wide, in the sense that is relevant to assessing epistemic health. In other words, if a community has reason to value revolutionary theorizing in their research, no amount of hot searching amidst that which is conceivable will amount to healthy inquiry. This is despite creativity remaining a standard of good health in that community, given their research aims.

But such a conclusion spells trouble for the applicability of Currie's account in arguments about policy. Currie's observation, as discussed above, is that conservatism promotes pooling with respect to shared priors. To the extent that creativity is anticorrelated with such pooling, Currie concludes research programs that ought to be creative likely suffer, in virtue of conservatism. Therefore, interventions that would promote creativity in the communities focused on those research programs would be well motivated, given the broader context of science today. (Indeed, this is just what Currie calls for in the case of X-risk.)

But now, there is cause to doubt that creativity has anything to do systematically with pooling, as defined with respect to shared priors. Creativity may, for instance, be anticorrelated with an entirely different kind of failure to explore. At least when revolutionary theorizing is valued, this seems to be the case. Plausibly, in such a case, the relevant kind of failure is one measured in terms of how little play there is, concerning what we conceive as worth exploring.

If so, interventions to promote creativity cannot be motivated against a background of conservatism, at least as Currie has presented the topic. In cases such as these, we require a different sort of reason to motivate interventions in response to conservatism (when, still, creativity is important). For instance, suppose that the conclusion is warranted: conservatism deprives the relevant community of access to much of solution space (cf. footnote 7). Then it is plausible that what is sought by the community is inaccessible, in which case revolutionary theorizing might be valuable. Policies intended to promote creativity in that community could then be motivated, given the broader conservatism of science today. (And enacting such policies would be all the more important if, following Stanford, we further regard conservatism as stifling revolutionary theorizing.)

On the other hand, we might imagine some cases (perhaps that of X-risk) in which Currie's account adequately captures the effects of conservatism on inquiry. These are cases where we regard a community's capacity for revolutionary theorizing as, antecedently, unimportant to assessing the health of inquiry therein.

Such cases may arise in practice. But if they do, it is very difficult— if not impossible— to reliably identify them as such. This is one lesson of Stanford's original project, which foremost concerned our means of evaluating the contemporary threats posed by the problem of unconceived alternatives. The upshot is that there may turn out to be no problem inherent in the applicability of Currie's account in certain cases. Yet, there is a severe problem in asserting when we are reliably in such a case. This matters for the argumentative force of any call for new incentives to promote creativity in any particular community, based on his account.

Whether Currie's account can provide insight into the effect of conservatism on inquiry will therefore require a more sophisticated understanding of creativity. Such an understanding would need to provide a reliable means of picking out those situations wherein the benefits of creativity are not to do with revolutionary theorizing. In those situations, Currie's account could give us some grasp of how to evaluate the epistemic health of the relevant community. But the grounds for that evaluation would ultimately reside in the more sophisticated account. This is because only according to that more sophisticated account could we explain in virtue of what revolutionary theorizing is, in the particular case at hand, rendered unimportant.

4.5 Discussion

I have argued that Rovelli's remarks ultimately uncover a shortcoming of Currie's account of creativity. This shortcoming concerns the possible value of revolutionary theorizing to the aims of any given research program. Lacking a more sophisticated account of creativity, it is difficult to assess a variety of claims of independent interest. For instance, what commitments does Rovelli make about the problem of quantum gravity, in order to claim that revolutionary theorizing is valuable within contemporary fundamental physics? And when is it appropriate to focus questions about creativity exclusively on just what is conceived as worth exploring? After all, Currie is unequivocal about the relevance of his more narrow account of creativity in the case of X-risk. He states: "...it is this kind of creativity which scientific study of existential risk requires" [p. 8]. So, by what reasons do the local details of X-risk entitle us to restrict our study to an account that disregards the possibility that revolutionary theorizing matters?

Currie anticipates the possibility that a more sophisticated notion of creativity might ultimately be demanded. By his reckoning, this is because his account does not capture ‘ingenuity’ (p. 8), failing to distinguish creative searches from chaotic ones. Currie then suggests that a new account of creativity, built on the notion of creative ‘flair’ developed by Gaut (2010), might capture such a distinction.

This suggestion strikes me as promising. For instance, creative searches might be those hot searches that enable the community to subsequently achieve novelty in research (e.g. at the end of some iterative process). But I would like to conclude by noting one major obstruction to developing the suggestion further. Following Currie, the first step in articulating an account of creativity would be to specify how to extrapolate from the individual to the community level. Such a move is essential to an understanding of the relationship between the social structure of science and creativity, like we have understood it here. (Of particular interest is whether conservatism can be responsible for reliably depriving us of access to much of a solution space, within the developed account.) But extrapolating from the individual to the community level is no small challenge. Creative flair is an irreducibly agential notion, concerning an individual’s familiarity with their own goals. It is unclear at present what would mark a community that, as a whole, is creative in this refined, goal-sensitive respect.

There is, it seems, still much work to be done.

Chapter 5

Conclusion

Exactly one work is known to be written by the pre-Socratic philosopher Parmenides. It is an epic poem written in hexameter verse, which is today highly fragmented. The preface to the poem announces Parmenides as a singular individual who has, through his own pursuits, come by way of revelation from a goddess to have learned more about the cosmos than that which was known prior (Palmer, 2019). What Parmenides means by "cosmos" is evidently not what we have come to call the same today. Nor is it what we had called the same before his revelation (for what a poor revelation that would be). Nonetheless, we are in the habit of regarding such a change of referent in terms of us— as a community, in perpetuity— iteratively having *learned more about* the cosmos, ever since.

In Parmenides case, the stated means for learning more about the cosmos are that of revelation. And I think a similar attitude persists today. As witnessed in the most recent chapter, it can be appealing to understand success in scientific discovery, in terms of being met with good fortune in our exploration of theory space. Though it may not always be prudent to think of theory space as structured just in terms of what we conceive as worth exploring, nothing of the argument in that chapter undermines the metaphor that, when engaged in theory development, there is a space that is being explored.

Nonetheless, I believe that such an understanding of theory development squares poorly with the view, articulated in that chapter, that strategically, revolutionary theorizing is sometimes needed (for instance, in regards to the problem that characterizes quantum gravity research). This is because, plausibly, revolutionary theorizing is needed just whenever it is the case that a community's aims of research mandate that their work be transformative. And in such a case, exploration no longer seems an apt description of theoretical research practice. More often than not, what is being called into question— when transformative research is demanded— is a community's understanding of what the problem is that ultimately characterizes their research, or structures whatever space is taken to be, figuratively, there around them, to explore.

In the first three chapters (including the Introduction), I argued, rather indirectly, for a particular perspective on theory development in fundamental physics. Namely, future theory is developed by means of solving various problems, each of which we anticipate will be solved within that future theory, based on our current understanding of the world. In my second chapter, I relate the method supporting this perspective explicitly to a view endorsed by Rovelli. Then, in the most recent chapter, that same view by Rovelli provides my primary means of raising trouble for one account of creativity that does not countenance what goes on, when revolutionary theorizing matters.

The implicit suggestion then, given this confluence of arguments, is that the perspective I have advocated in the earlier chapters may be better suited for understanding how we achieve transformative research, such as is crucial in contemporary quantum gravity research.

Indeed, I am hopeful that this is the case. Notably, this perspective on theory development seems more amenable to articulating a suitable notion of creative flair in scientific discovery.

Recall that, at the end of the previous chapter, it was noted that flair could be important, in the pursuit of a more general account of creativity that can countenance the possible value of revolutionary theorizing. There, I suggested that flair might be relevant to describing how a community may achieve novelty in research, at the end of some iterative process, which amounts to ingenuity. I will conclude by sketching out this thought just a little bit more.

Endeavoring to solve a specific problem in the course of achieving future theory, specifically because one anticipates that the future theory will include some such solution to that problem, seems to be the kind of act in theoretical research that is done with purpose, understanding, and judgment, as well as with an awareness of what constitutes the completion of the act. If so, then future theory developed on the basis of this method could constitute a theory produced with flair. (Though, it may still be difficult to spell out an analogue notion of flair, which is defined at the community level.)

What remains to be shown is that such a theory, produced with flair, bears no necessary relationship to whatever reasons we might give for why a theoretical possibility is seen as conceivably worth studying. And, while I will not claim presently that this is so, it does strike me as likely. Although the specification of what problems are to be worked on, based on this method, depends sensitively on our current understanding of our current theories, as illustrated in Chapter 2, the method does not obviously distinguish any proposed means of solving those problems as being too speculative. Yet, the claim that some such proposals are too speculative is exactly the manner of objection one can expect (and, I think, one often sees), were it inconceivable that we would ever see fit to embrace, as our future theory, anything like what is being proposed.

If such a claim rings hollow as an objection to any research conducted in accordance with the method, then that method is, plausibly, suitable for a kind of theory development that proceeds irrespective of which theoretical possibilities are seen as conceivably worth studying. In other words, provided that future theory can be produced with flair (in the sense intended here), the resulting theory may just as well be something whose potential value was, until that moment, inconceivable. This bodes well for an understanding of creativity, which is suitable even when transformative research is demanded: the account is insensitive to any judgements about whether or not revolutionary theorizing is needed. The focus, instead, is on the creative production of a novel theory, which is built so as to succeed in every regard that we presently see fit to ask of it.

Bibliography

Afshordi, Niayesh (2008). "Gravitational Aether and the thermodynamic solution to the cosmological constant problem." *arXiv preprint arXiv:0807.2639*.

Aldrovandi, R and JG Pereira (2009). "de Sitter special relativity: effects on cosmology." *Gravitation and Cosmology*, *15*(4), 287–294.

Almeida, JP Beltrán, CSO Mayor, and JG Pereira (2012). "De sitter relativity: a natural scenario for an evolving Λ." *Gravitation and Cosmology*, *18*(3), 181–187.

Barrett, Jeffrey Alan (2008). "Approximate truth and descriptive nesting." *Erkenntnis*, *68*(2), 213–224.

Belot, Gordon (1998). "Understanding electromagnetism." *The British Journal for the Philosophy of Science*, *49*(4), 531–555.

Bianchi, Eugenio and Carlo Rovelli (2010). "Why all these prejudices against a constant?." *arXiv preprint arXiv:1002.3966*.

Carroll, Sean M (2001). "The cosmological constant." *Living Rev. Rel*, *4*(1), 41.

Crowther, Karen (2013). "Emergent spacetime according to effective field theory: From top-down and bottom-up." *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, *44*(3), 321–328.

Curiel, Erik (2015). "Measure, topology and probabilistic reasoning in cosmology." *arXiv preprint arXiv:1509.01878*.

Curiel, Erik (2016). "A Simple Proof of the Uniqueness of the Einstein Field Equation in All Dimensions." *arXiv preprint arXiv:1601.03032*.

Currie, Adrian (2019). "Existential risk, creativity & well-adapted science." *Studies in History and Philosophy of Science Part A*, *76*, 39–48.

Dardashti, Radin (2019). "Physics without Experiments?." *Why Trust a Theory?: Epistemology of Fundamental Physics.* . Cambridge University Press, 154–172.

De Rham, Claudia, Stefan Hofmann, Justin Khoury, and Andrew J Tolley (2008). "Cascading gravity and degravitation." *Journal of Cosmology and Astroparticle Physics*, *2008*(02), 011.

Dvali, Gia, Stefan Hofmann, and Justin Khoury (2007). "Degravitation of the cosmological constant and graviton width." *Physical Review D*, *76*(8), 084006.

Earman, John (2001). "Lambda: The constant that refuses to die." *Archive for History of Exact Sciences*, *55*(3), 189–220.

Earman, John (2003). "The cosmological constant, the fate of the universe, unimodular gravity, and all that." *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, *34*(4), 559–577.

Earman, John (unpublished). "Trace-Free Gravitational Theory (aka Unimodular Gravity) for Philosophers." Manuscript.

Ellis, George FR, Henk Van Elst, Jeff Murugan, and Jean-Philippe Uzan (2011). "On the trace-free Einstein equations as a viable alternative to general relativity." *Classical and Quantum Gravity*, *28*(22), 225007.

Feintzeig, Benjamin H (2017). "On theory construction in physics: Continuity from classical to quantum." *Erkenntnis*, *82*(6), 1195–1210.

Fletcher, Samuel C, John Byron Manchak, Mike D Schneider, and James Owen Weatherall (2018). "Would two dimensions be world enough for spacetime?." *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, *63*, 100–113.

Frieman, Joshua A, Michael S Turner, and Dragan Huterer (2008). "Dark Energy and the Accelerating Universe." *Annual Review of Astronomy and Astrophysics*, *46*, 385–432.

Gaut, Berys (2010). "The Philosophy of Creativity." *Philosophy Compass*, *5*(12), 1034–1046.

Hollands, Stefan and Robert M Wald (2004). "Essay: Quantum Field Theory Is Not Merely Quantum Mechanics Applied to Low Energy Effective Degrees of Freedom." *General Relativity and Gravitation*, *36*(12), 2595–2603.

Huggett, Nick and Craig Callender (2001). "Why quantize gravity (or any other field for that matter)?." *Philosophy of Science*, *68*(S3), S382–S394.

Kachru, Shamit, Renata Kallosh, Andrei Linde, and Sandip P Trivedi (2003). "De Sitter vacua in string theory." *Physical Review D*, *68*(4), 046005.

Kamenshchik, Alexander, Ugo Moschella, and Vincent Pasquier (2001). "An alternative to quintessence." *Physics Letters B*, *511*(2), 265–268.

Koberinski, Adam (2021). "Problems with the cosmological constant problem." *Philosophy Beyond Spacetime*. Ed. Christian Wüthrich, Baptiste Le Bihan, and Nick Huggett. Oxford: Oxford University Press (in preparation).

Kowalski-Glikman, Jerzy and Sebastian Nowak (2003). "Doubly special relativity and de Sitter space." *Classical and Quantum Gravity*, *20*(22), 4799.

Kragh, Helge (2012). "Preludes to dark energy: zero-point energy and vacuum speculations." *Archive for history of exact sciences*, *66*(3), 199–240.

Laudan, Larry (1978). *Progress and its problems: Towards a theory of scientific growth.* Volume 282. University of California Press.

Malament, David B (2012). *Topics in the foundations of general relativity and Newtonian gravitation theory.* University of Chicago Press.

Martin, Jerome (2012). "Everything you always wanted to know about the cosmological constant problem (but were afraid to ask)." *Comptes Rendus Physique*, *13*(6), 566–665.

Otten, Lars. "LaTeX template for thesis and dissertation documents at UC Irvine.".

Palmer, John (2019). "Parmenides." *The Stanford Encyclopedia of Philosophy.* Ed. Edward N. Zalta. Fall 2019 edition. Metaphysics Research Lab, Stanford University.

Redhead, Michael (1994). "The vacuum in relativistic quantum field theory." *PSA: proceedings of the biennial meeting of the philosophy of science association.* 77–87.

Rovelli, Carlo (2018). "Physics needs philosophy. Philosophy needs physics." *Foundations of Physics*, *48*(5), 481–491.

Rugh, Svend E, Henrik Zinkernagel, and Tian Yu Cao (1999). "The Casimir effect and the interpretation of the vacuum." *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, *30*(1), 111–139.

Rugh, Svend Erik and Henrik Zinkernagel (2002). "The quantum vacuum and the cosmological constant problem." *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics*, *33*(4), 663–705.

Saunders, Simon (2002). "Is the Zero-Point Energy Real?." *Ontological aspects of quantum field theory*, 313–343.

Smeenk, Chris (2013). "Philosophy of Cosmology." *The Oxford Handbook of Philosophy of Physics.*

Stanford, P Kyle (2019). "Unconceived alternatives and conservatism in science: the impact of professionalization, peer-review, and Big Science." *Synthese*, *196*(10), 3915–3932.

Teh, Nicholas J (2014). "Gravity and gauge." *The British Journal for the Philosophy of Science*, *67*(2), 497–530.

Wald, Robert M (1984). "General relativity." *Chicago, University of Chicago Press.*

Wald, Robert M (1994). *Quantum field theory in curved spacetime and black hole thermodynamics.* University of Chicago Press.

Wallace, David (2006). "In defence of naiveté: The conceptual status of Lagrangian quantum field theory." *Synthese*, *151*(1), 33–80.

Wang, Qingdi and William G Unruh (2018). "Reply to "Comment on 'How the huge energy of quantum vacuum gravitates to drive the slow accelerating expansion of the Universe' "." *Physical Review D*, *97*(6), 068302.

Wang, Qingdi, Zhen Zhu, and William G Unruh (2017). "How the huge energy of quantum vacuum gravitates to drive the slow accelerating expansion of the Universe." *Physical Review D*, *95*(10), 103504.

Weatherall, James Owen (2011). "On (some) explanations in physics." *Philosophy of Science*, *78*(3), 421–447.

Weinberg, Steven (1989). "The cosmological constant problem." *Reviews of Modern Physics*, *61*(1), 1.

Williams, Porter (2015). "Naturalness, the autonomy of scales, and the 125 GeV Higgs." *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, *51*, 82–96.

Wise, Derek K (2010). "MacDowell–Mansouri gravity and Cartan geometry." *Classical and Quantum Gravity*, *27*(15), 155010.

Witten, Edward (2001). "The cosmological constant from the viewpoint of string theory." *Sources and detection of dark matter and dark energy in the universe.* . Springer, 27–36.

Zel'dovich, Ya B (1968). "The cosmological constant and the theory of elementary particles." *Soviet Physics Uspekhi*, *11*(3), 381–393.

Zlatev, Ivaylo, Limin Wang, and Paul J Steinhardt (1999). "Quintessence, cosmic coincidence, and the cosmological constant." *Physical Review Letters*, *82*(5), 896.

Zumino, Bruno (1975). "Supersymmetry and the vacuum." *Nuclear Physics B*, *89*(3), 535–546.